

COLLECTION: LAUNCH EDITION

ADVANCES IN COMPUTATIONAL LEARNING FOR ROBOTICS

Highlighting the roles of advanced computational
learning approaches for robotics applications



Active triggering control of pneumatic rehabilitation gloves based on surface electromyography sensors

Yongfei Feng, Mingwei Zhong, Xusheng Wang, Hao Lu, Hongbo Wang, Pengcheng Liu, Luige Vladareanu

Ultrawideband (UWB)-based precise short-range localization for wireless power transfer to electric vehicles in parking environments

Seung-Mok Lee

Neural network assisted Kalman filter for INS/UWB integrated seamless quadrotor localization

Shuhui Bi, Liyao Ma, Tao Shen, Yuan Xu, Fukun Li

SeedSortNet: a rapid and highly efficient lightweight CNN based on visual attention for seed sorting

Chunlei Li, Huanyu Li, Zhoufeng Liu, Bicao Li, Yun Huang

Musculoskeletal modeling and humanoid control of robots based on human gait data

Jun Yu, Shuaishuai Zhang, Aihui Wang, Wei Li, Lulu Song

Service humanoid robotics: a novel interactive system based on bionic-companionship framework

Jiaji Yang, Eysin Chew, Pengcheng Liu

The classification of skateboarding tricks via transfer learning pipelines

Muhammad Amirul Abdullah, Muhammad Ar Rahim Ibrahim, Muhammad Nur Aiman Shapiee, Muhammad Aizzat Zakaria, Mohd Azraai Mohd Razman, Rabiul Muazu Musa, Noor Azuan Abu Osman, Anwar P.P. Abdul Majeed

Two-stage training algorithm for AI robot soccer

Taeyoung Kim, Luiz Felipe Vecchietti, Kyujin Choi, Sanem Sariel, Dongsoo Har

Advances in Computational Learning for Robotics

Collection Introduction

Advances in Computational Learning for Robotics features selected papers presented at the 8th International Conference on Robot Intelligence Technology and Applications (RiTA 2020).

This Conference Collection, published in *PeerJ Computer Science*, is concerned with an important area that has recently attracted the attention of both Robotics and Artificial Intelligence researchers: developing intelligent robotic systems that are capable of making decisions and acting autonomously in real and unpredictable environments, to accomplish tasks and assist humans across various domains within society.

We developed this special issue around the theme of Computational Learning to highlight the critical importance of this topic for ongoing progress in robotics and artificial intelligence.

At launch, this Conference Collection is formed of eight published articles that have examined this issue from various perspectives, such as transfer learning pipelines for skateboarding skills, a novel interactive system for service humanoid robots, musculoskeletal modeling and humanoid control, an algorithm for training AI soccer robots, a visual-attention-based rapid and efficient lightweight CNN for object sorting, an improved Kalman filter for quadrotor localization, a precise localization method for wireless power transfer to electric vehicles, and active triggering control of pneumatic rehabilitation gloves. Additional articles are still under consideration at PeerJ Computer Science and will be added to this collection upon their publication, so remember to check back regularly for more research on this topic.

We hope the robotics and artificial intelligence communities will find this special issue to be an informative and useful collection of articles.

We would like to thank the PeerJ Computer Science editorial board and a long list of anonymous reviewers for their thoughtful suggestions and constructive criticisms. The contributions of the individual articles and of the Conference Collection in its entirety are stronger thanks to their hard work.

Collection Editors: Pengcheng Liu (a), Hyun Myung (b), Yu Dong (c), Guibin Bian (d), Junwen Zhong (e), Eysin Chew (f)

a Department of Computer Science, University of York, UK;

b Korea Advanced Institute of Science and Technology (KAIST), South Korea;

c School of Civil and Mechanical Engineering, Curtin University, Australia;

d Institute of Automation, Chinese Academy of Sciences, China;

e Faculty of Science and Technology, University of Macao, China;

f EUREKA Robotics Lab, Cardiff Metropolitan University, UK

PeerJ is a modern and streamlined publisher, built for the internet age. Our mission is to give researchers the publishing tools and services they want, with a unique and exciting experience. All of our seven journals are Gold Open Access and are widely read and cited, with over 500,000 monthly views and 48,500 content alert subscribers. We have published 12,844 peer-reviewed articles since 2013.



Prestigious
Editorial Board



High-Impact
Research



Quality
Peer Review



Rapid
Publishing



Optimum
Discoverability

PeerJ Computer Science



High-quality, developmental peer review, coupled with industry leading customer service and an award-winning submission system, means PeerJ Computer Science is the optimal choice for your computer science research.

Impact Factor: 1.39
Citescore: 4.8

Scimago Ranking: 0.806
SNIP: 1.895

PeerJ Conference Collections are free for conference organisers and can increase the visibility of both the event and the research presented during the meeting. If you are organising a conference and would like to consider a PeerJ Conference Collection, email us at communities@peerj.com.

Active triggering control of pneumatic rehabilitation gloves based on surface electromyography sensors

Yongfei Feng^{1,2}, Mingwei Zhong¹, Xusheng Wang³, Hao Lu³, Hongbo Wang³, Pengcheng Liu⁴ and Luige Vladareanu²

¹ Faculty of Mechanical Engineering & Mechanics, Ningbo University, Ningbo, Zhejiang Province, China

² Robotics and Mechatronics Department, Institute of Solid Mechanics of the Romanian Academy, Bucharest, Bucharest, Romania

³ Academy for Engineering & Technology, Fudan University, Shanghai, Shanghai, China

⁴ Department of Computer Science, University of York, York, York, United Kingdom

ABSTRACT

The portable and inexpensive hand rehabilitation robot has become a practical rehabilitation device for patients with hand dysfunction. A pneumatic rehabilitation glove with an active trigger control system is proposed, which is based on surface electromyography (sEMG) signals. It can trigger the hand movement based on the patient's hand movement trend, which may improve the enthusiasm and efficiency of patient training. Firstly, analysis of sEMG sensor installation position on human's arm and signal acquisition process were carried out. Then, according to the statistical law, three optimal eigenvalues of sEMG signals were selected as the follow-up neural network classification input. Using the back propagation (BP) neural network, the classifier of hand movement is established. Moreover, the mapping relationship between hand sEMG signals and hand actions is built by training and testing. Different patients choose the same optimal eigenvalues, and the calculation formula of eigenvalues' amplitude is unique. Due to the differences among individuals, the weights and thresholds of each node in the BP neural network model corresponding to different patients are not the same. Therefore, the BP neural network model library is established, and the corresponding network is called for operation when different patients are trained. Finally, based on sEMG signal trigger, the pneumatic glove training control algorithm was proposed. The combination of the trigger signal waveform and the motion signal waveform indicates that the pneumatic rehabilitation glove is triggered to drive the patient's hand movement. Preliminary tests have confirmed that the accuracy rate of trend recognition for hand movement is about 90%. In the future, clinical trials of patients will be conducted to prove the effectiveness of this system.

Subjects Human-Computer Interaction, Artificial Intelligence, Robotics

Keywords Hand rehabilitation, Pneumatic rehabilitation gloves, Surface electromyography, Active trigger control system, Back propagation neural network

INTRODUCTION

Approximately two million people suffer from stroke every year in China, and about three-fourths of stroke patients have hand movement disorders (*Heung et al., 2020*).

Submitted 24 December 2020

Accepted 27 February 2021

Published 19 April 2021

Corresponding author

Luige Vladareanu,
luige.vladareanu@vipro.edu.ro,
luigiv2007@yahoo.com.sg

Academic editor

Gui-Bin Bian

Additional Information and
Declarations can be found on
page 16

DOI 10.7717/peerj-cs.448

© Copyright
2021 Feng et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Moreover, the other neurological disorders, such as multiple sclerosis or motor neuron disease, also show abnormal hand movements. Patients with inflexible hands are unable to complete various actions in daily life due to lack of muscle strength and fine control of the fingers. Rehabilitation robot is playing an increasingly important role in training patients instead of rehabilitation physicians, which can improve the motor function of inflexible hands and reduce the possibility of permanent disabilities ([Gaia et al., 2020](#); [Yurkewich et al., 2020](#); [Lemerle, Nozaki & Ohnishi, 2018](#)). At present, the popular hand rehabilitation robots at present can be divided into finger exoskeleton rehabilitation robot ([Agarwal et al., 2015](#); [Nycz et al., 2016](#)), flexible rehabilitation robot gloves (FRRG) and end traction finger rehabilitation robot ([Bentzvi & Ma, 2014](#); [Wu et al., 2010](#)). Compared with other types of hand rehabilitation robots, FRRG has some advantages, including good flexibility, small size, large working space, light weight, safety and reliability ([Heung et al., 2019](#); [Mahdi, Charu & Muthu, 2019](#); [Matthew, Jeong & Raye, 2019](#)). [Polygerinos et al. \(2015\)](#) developed the rehabilitation gloves, which include a molded elastomer chamber and a fiber reinforcement that produces specific bending, twisting and extending trajectories under fluid pressure to match and support the different ranges of motion of a single finger. Wang, et al. proposed a pair of antagonistic pneumatic muscles which are very similar in action to human muscles, can be used for hand passive training ([Wang et al., 2020a](#); [Wang et al., 2020b](#)). A new kind soft pneumatic glove with five segmented PneuNets bending actuators is made of elastomer, whose actuator driving the corresponding finger to bend ([Wang, Fei & Pang, 2019](#)). A new portable and inexpensive pneumatic rehabilitation glove is proposed in this paper.

Rehabilitation training, which is based on limb movement trend of patients, can improve the efficiency of recover ([Pichiorri et al., 2015](#)). The methods for trend recognition of human limb movement include biomechanical signal ([Sangwoo et al., 2018](#)) and bioelectrical signal ([Leonardis et al., 2015](#)). However, due to the structure and wearing characteristics of FRRG, it is expensive to install biomechanical sensors on the gloves, which make it difficult to use for patients with financial problems in their families. For patients with finger dysfunction caused by stroke, biomechanical sensors are not suitable for them and not easy to collect the biomechanical signals of their hands ([Leonardo et al., 2018](#)). On the contrary, bioelectrical signals are generated before movement, and the corresponding relationship between signals and movement can be obtained by collecting and decoding bioelectrical signals of human body, which provides an extremely important means for the prediction of human limb movement trend. There are many mature methods of limb movement intention recognition based on bioelectrical signals, including electrocorticogram (ECoG), electroencephalogram (EEG), magnetoencephalo-graphy (MEG) and electromyography (EMG). Due to the high cost of collecting ECoG, EEG or MEG signals, EMG is chosen as the bioelectrical signal for hand movement trend recognition in this paper.

EMG signals can be divided into two types; surface electromyography (sEMG) and needle in electromyography (nEMG). Compared with nEMG, sEMG has the advantages of noninvasive and simple operation. The signal collected by sEMG sensor is the sum of the potential generated by muscle activity in the area where the electrode is located on the skin surface. Selecting the appropriate muscle group of arm is very important and

different muscle groups have different effects, which is reflected in the amplitude change of sEMG signals (Dai & Hu, 2020). The larger the amplitude change, the more conducive to the identification of hand movement trend. The control based on bioelectrical signal from patient muscle, mainly includes sEMG trigger control (Meng et al., 2014) and sEMG continuous control (Song et al., 2008). In this paper, a new pneumatic glove trigger control system for paralysis patients' hand is developed. The trigger control is used to identify the movement trend of the patients, and then the assisting to complete the rehabilitation training is realized.

Construction of pneumatic rehabilitation glove trigger control system based on sEMG

The pneumatic rehabilitation glove trigger control system based on sEMG consists of one pneumatic gloves, an air pump, a Stm32f103 microprocessor equipped with an ARM chip, two electric relays, a Myoware sEMG sensor, two-position three-way solenoid valves and a host computer as shown in Fig. 1. The pneumatic rehabilitation gloves can well wrap the patients' fingers, palms and hand back. Air pump provides power for pneumatic gloves. sEMG sensors are used to collect patient's sEMG signals. The Stm32f103 microprocessor equipped with an ARM chip is used to process the original sEMG signals collected by sEMG sensors. It is also used as the driver of air pump and transmits the processed sEMG signals to the host computer. The host computer is developed with QT software (Cross-platform software development framework for the development of apps and devices, developed by QT Group) as the development environment. It judges the movement trend of the hand by analyzing the collected sEMG signals. According to the movement trend of the hand, it also sends related instructions to the air pump driver. Then the air pump driver controls pneumatic rehabilitation gloves to flex and extend. The above hardware platform can be divided into an acquisition layer, a decision-making layer, a driving layer and an execution layer as shown in Fig. 1. The RS232-USB (RS232 to USB) serial port is adopted between the acquisition layer and the decision layer, the decision-making layer and the drive layer. The high and low level control of the IO port pins is used between the drive layer and the execution layer. The host computer uses the QSerialPort component (Function pack of QT) to receive the sEMG signals through the RS232-USB serial port, and stores the received sEMG data in an Excel table to facilitate the subsequent static data processing.

Processing and selection of optimal eigenvalues of sEMG signals *Acquisition and processing of the sEMG Signals*

In order to facilitate the collection of sEMG signals, the muscle group on the forearm is selected as the collection object. The muscle groups of the forearm mainly include palmar longus, flexor carpi radialis, brachioradialis, teres pronatorus, extensor carpi radialis longus, extensor digitorum and flexor digitorum superficialis. The flexor carpi radialis is a flexor wrist muscle located on the inner side of the forearm. It starts from the medial epicondyle of the humerus and the olecranon, and ends at the proximal end of the second metacarpal bone. The flexor superficialis is mainly responsible for flexing the metacarpophalangeal joint and proximal interphalangeal joint of the 2nd to 5th fingers. The extensor digitorum

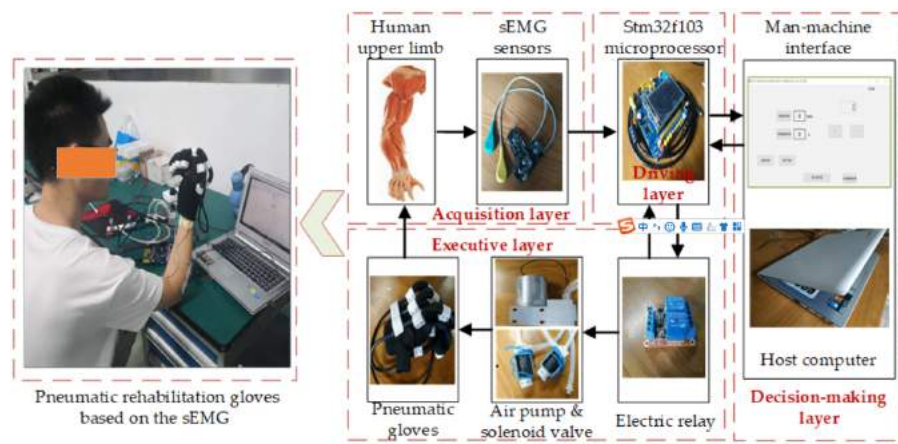


Figure 1 The composition of the trigger control system of the pneumatic rehabilitation gloves based on the sEMG.

[Full-size](#) DOI: 10.7717/peerjcs.448/fig-1

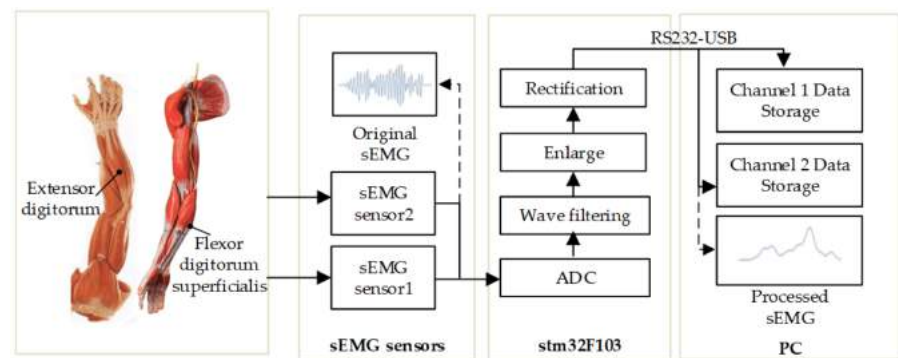


Figure 2 Flowchart of original sEMG signal acquisition and processing.

[Full-size](#) DOI: 10.7717/peerjcs.448/fig-2

can extend the metacarpophalangeal joint of the four fingers. The original sEMG signals are collected by dual-channel sEMG sensors. Each sEMG sensor has two detection electrodes and one reference electrode. The detection electrode is attached to the central part of the muscle belly of the target muscle, and the reference electrode is attached to the muscle not participating in the test exercise. The processed sEMG signal amplitude varies from 0 to 3.3V and the original sEMG signal acquisition and processing process is shown in Fig. 2. Three healthy volunteers were recruited in this experiment with the informed consents of all volunteers and the Ethical Approval (No. [2020]LLSP(12), Ethics Committee of Faculty of Mechanical Engineering & Mechanics, Ningbo University). Volunteer 1: Male, weight 64 kg, height 175 cm, 24 years old; Volunteer 2: Male, weight 73 kg, height 177 cm, 26 years old; Volunteer 3: Male, weight 75 kg, height 180 cm, 20 years old. Using sEMG sensors and Stm32f103 microprocessor, the original sEMG signals are digitally filtered, amplified, rectified and smoothed (Lyu et al., 2020; Shi et al., 2020). After repeated experiments and comparing the amplitudes of the sEMG signals of different muscle groups

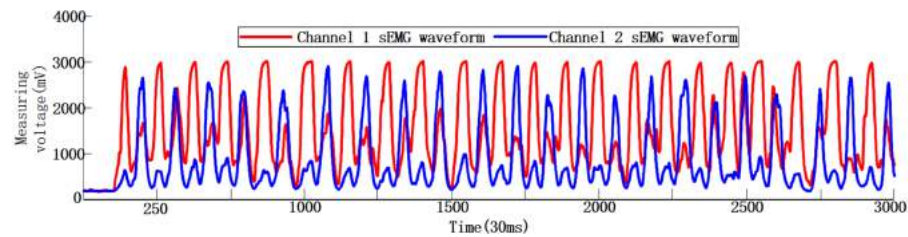


Figure 3 Dual-channel human hand motion sEMG waveforms in 90s.

[Full-size](#) DOI: 10.7717/peerjcs.448/fig-3

collected during the same hand action, the extensor digitorum and flexor digitorum superficialis are finally selected as the muscle groups for sEMG signal collection. Volunteer 1 uses dual-channel sEMG sensors to collect the actual sEMG signals during the flexion and extension movement of his hand, as shown in Fig. 3. The total signal collection duration is about 90 s, of which the sEMG signal curves do not fluctuate much in the first 3 s, as the volunteer is in a state of inactivity. During the movement of the subject's hand, the corresponding to the hand sEMG signal curves have changed, and the waveform in the figure appears to be convex. By observing the sEMG signals of the two channels, it can be seen that the signals of the two channels fluctuate synchronously when the subject hand is moving, but there are certain differences in the waveforms of each channel.

Selection of optimal eigenvalues of the sEMG signals

Figure 4 shows the obtained eigenvalues of sEMG sensor's channel 1. It is necessary to use the law of statistics to find the accurate physical quantities that best represent the essence of the surface EMG signal, that is, the extracting eigenvalues of sEMG signals. The original sEMG signal after amplification, rectification and rectification integration loses a lot of frequency domain characteristics of the original signal. By directly analyzing and processing the sEMG signal in the time domain, it will be intuitive and accurate. In the time domain, the sEMG signal can be approximated as a Gaussian distribution. At present, the most commonly used time domain eigenvalues of the signal are the root mean square value (RMS), peak value (PV), mean value (MAV), wavelength average (WAV), form factor (FF) and Willison amplitude (WAMP) (Liu & Cheng, 2018). The number of eigenvalues selected is positively correlated with the accuracy of the information representation contained in the sEMG signals, but too many eigenvalues will affect the speed of the computer to make decisions, which is manifested in the deterioration of the follow ability of the pneumatic gloves to the patient's intention. On the contrary, if the selected number of eigenvalues of the sEMG signal is too small, the pneumatic rehabilitation glove control system cannot accurately recognize the patient's movement intention. x_i represents the amplitude of the signal, and n represents the extracted step size. First, N ($N = 30$) groups of sEMG signals are extracted to form sEMG samples with empirical steps $n = 100$, $n = 150$, $n = 200$ in the continuously collected sEMG signals respectively as W1, W2, and W3. And then the above-mentioned 6 eigenvalues with each segment length as the unit to form an eigenvalue

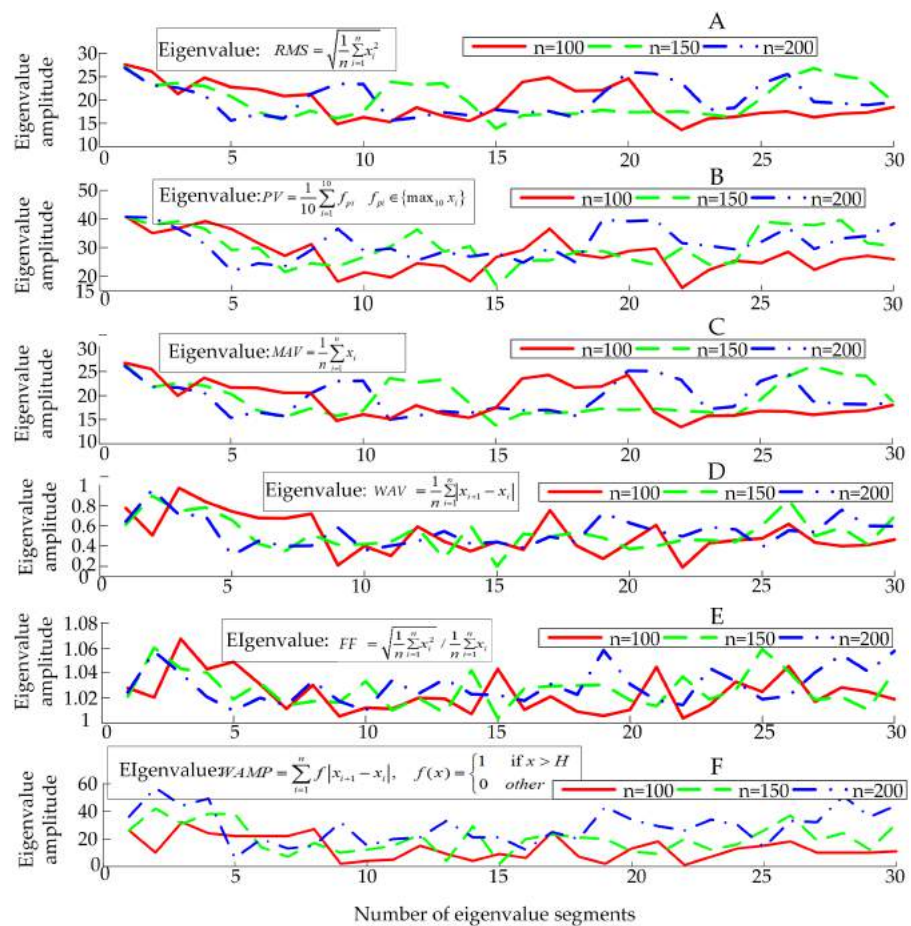


Figure 4 Sample of sEMG eigenvalues.

[Full-size](#) DOI: 10.7717/peerjcs.448/fig-4

sample E_i ($6 \times N$) is calculated, where $i = 1, 2, 3$ corresponds to the sEMG samples W1, W2, W3, respectively.

The patient's hand movement trend will be expressed as fluctuations in sEMG signals. The eigenvalues of the signals reflect the nature of the signals over a period of time, so the fluctuation of the sEMG will also be specifically reflected in the fluctuation of sEMG eigen-values. According to prior knowledge, it can be known that the greater the degree of dispersion of eigenvalues, the more conducive the neural network to the recognition of the movement trend based on eigenvalues. Based on the six eigenvalues, three eigenvalues with a large degree of dispersion will be selected as the parameters of the next action classification, participating in the training and testing of the neural network for intention recognition. Since a single dispersion index is not sufficient to fully characterize the degree of dispersion of the signals, 4 dispersion indicators will be used to process the 6 eigenvalues that have been obtained, namely range (R), interquartile range (Q), and variance (V) and fourth-order center distance (K).

Range is the difference between the maximum and minimum values between data. The greater the range, the greater the degree of dispersion, namely:

$$R = \max(s_i) - \min(s_i). \quad (1)$$

The interquartile range represents the range of the middle half of the data. The larger the interval, the greater the degree of dispersion. Arrange a set of data in ascending order. The number in the $x\%$ position is represented by P_x . The lower quartile and upper quartile are P_8 and P_{23} respectively, namely:

$$Q = P_{23} - P_8. \quad (2)$$

Variance describes the degree of dispersion of data mathematical expectation, that is, the greater the variance, the greater the degree of dispersion, namely:

$$V = \frac{1}{N} \sum_{i=1}^N \left(s_i - \frac{1}{N} \sum_{i=1}^N s_i \right)^2. \quad (3)$$

The fourth-order center distance is a cumulative numerical statistics reflecting the distribution characteristics of random variables. The larger the fourth-order center distance, the smaller the degree of dispersion, namely:

$$K = \frac{\frac{1}{N} \sum_{i=1}^N \left(|s_i| - \frac{1}{N} \sum_{i=1}^N s_i \right)^4}{\left(\frac{1}{N} \sum_{i=1}^N s_i^2 \right)^2}. \quad (4)$$

In Eqs. (1) and (4), S_i represents the data amplitude and N represents the data length. The process of determining the optimal eigenvalue is shown in Fig. 5.

By observing the sorting results of the data dispersion degree in Table 1, three eigenvalues with the largest dispersion degree are selected, which are WAMP, PV and RMS. For further verification, the dispersion index of E_2 and E_3 are calculated by the same method, and a comprehensive ranking is performed according to the magnitude of the dispersion index, as shown in Tables 2 and 3.

Research on hand movement trend recognition based on BP neural network

Using the collected sEMG signals to achieve the purpose of identifying the patient's finger movement trend is the main problem in the design of the pattern recognition classifier. The back propagation (BP) neural network model was chosen to construct the motion recognition classifier, as the BP neural network model has good self-learning, nonlinear mapping and adaptation, generalization and fault tolerance (Wang et al., 2020a; Wang et al., 2020b). It could be an ideal movement trend pattern recognition tool.

Construction of BP neural network classifier

BP neural network is an adaptive nonlinear dynamic system composed of a large number of interconnected neurons. It can learn and store the mapping relationship of multiple input-output modes without describing specific mathematical equations in advance. The



Table 1 E_1 dispersion index magnitude ordering.

Table 2 E_2 dispersion index magnitude ordering.

quality of neural network classifiers is closely related to the number of neural network layers, the number of nodes in each layer, the transfer function of the hidden layer, and the learning algorithm. The training algorithm flow chart of constructing BP neural network under QT software development environment is shown in Fig. 6.

Feng et al. (2021), *PeerJ Comput. Sci.*, DOI 10.7717/peerj-cs.448

Table 3 E_3 dispersion index magnitude ordering.

Values index	MAV	PV	RMS	WAMP	FF	WAV	Dispersion index order
V	10.6075	27.7494	11.0703	164.1194	0.0002	0.0191	4, 2, 3, 1, 6, 5
R	11.1623	18.9855	11.2990	51.0000	0.0487	0.6539	4, 2, 3, 1, 6, 5
Q	3.4440	4.6856	4.1505	16.0000	0.0130	0.1500	4, 2, 3, 1, 6, 5
W	0.0014	0.0015	0.0013	0.0524	0.0004	0.0121	3, 4, 2, 6, 1, 5

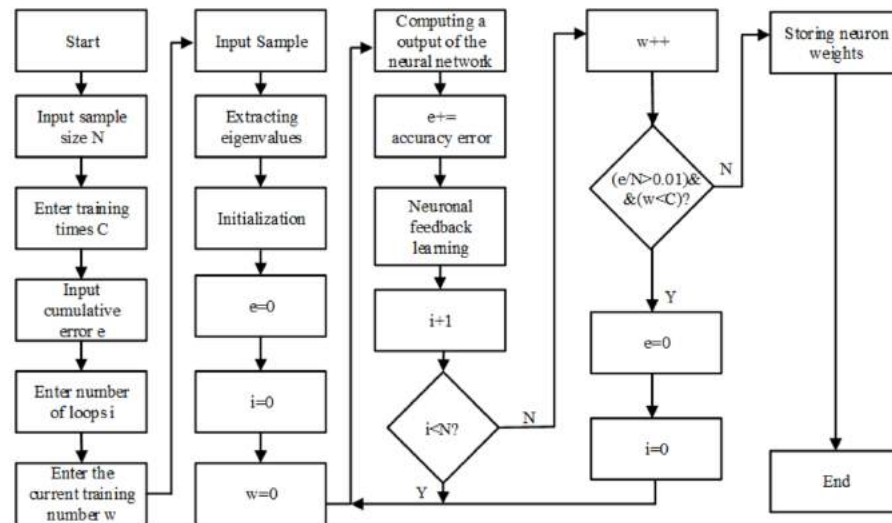


Figure 6 Flowchart of training algorithm for BP network.

Full-size [DOI: 10.7717/peerjcs.448/fig-6](https://doi.org/10.7717/peerjcs.448/fig-6)

and m-dimensional output, so in order to simplify the calculation, a three-layer network is adapted (Hecht-Nielsen, 1992).

Hidden layer transfer function:

$$y_i = \frac{(x_i - MinValue + A)}{MaxValue - MinValue + A} \quad (5)$$

Transfer function of the output layer:

$$y_k = x_k \times (MaxValue - MinValue + A) - A + MinValue \quad (6)$$

Logsig activation function is used:

$$y_j = \frac{1}{1 + e^{-x_i}} \quad (7)$$

Levenberg-Marquart (L-M) learning algorithm is used:

$$\Delta\omega = (J^T J + \mu I)^{-1} g J^T e \quad (8)$$

In Eqs. (5) and (6), *MinValue* is the minimum value of the input layer value; *MaxValue* is the maximum value of the input layer value; constant *A* represents the denominator from being zero; x_i represents the eigenvalue extracted from the sEMG signals; y_i represents the

Table 4 Action coding.

Action type	Action encoding
Action	1
No action	0

normalized feature value of the input layer; x_k represents the output value of the hidden layer, and y_k represents the final output value of the output layer. In Eq. (7), x_i represents the sum of the product of the output value and the weight of each neuron in the previous network, y_j represents the output of the j neuron in the current layer network. In Eq. (8), J represents the Jacobian matrix of the derivative of weights from network error, e represents the error vector, μ is an adaptive constant, which is greater than 0. The input layer is a 6×1 vector composed of the optimal eigenvalues of the 2-channel sEMG signals, so the number of nodes in the input layer is 6, set the number of nodes in the output layer to 1, and use the output result of the output layer to determine the triggered action. The action code is built as in Table 4.

The number of hidden layer nodes is determined by the following empirical formula (Sheela & Deepa, 2013):

$$n_1 = \sqrt{n+m} + a \quad (9)$$

where, n is the number of input nodes; m is the number of output nodes; n_1 is the number of hidden nodes; a is a constant between 1 and 10.

The number of hidden nodes gradually increases, and the training error of the neural network is observed during this process. As the number of hidden layer nodes increases, the training error gradually decreases, but after a certain number of nodes, the test error will fluctuate greatly. Therefore, considering the trend of training and test error changes, the number of hidden layer nodes is finally determined to be 12.

Training and testing of BP neural network

In order to realize the mapping function of the input matrix and the output matrix, the BP neural network needs to be trained. The feedback mechanism of BP neural network includes two parts. One is that the BP neural network produces prediction results. The other is to compare the prediction results with sample results, and then correct the neuron error until the error meets the specified requirements or reaches the specified number of training sessions. 160 sets of data are used as training samples to train the BP neural network as shown in Table 5. Each set of data contains the input and target output of the BP neural network. The input is the optimal eigenvalues of the sEMG signals collected by the two channels of the sEMG sensors, and the output is the code value of the corresponding action.

Before training the BP neural network, the training samples need to be randomly divided into two types at a ratio of 3:1, as training samples and test samples separately. After the BP neural network uses the training sample to complete each iteration, it is judged whether the average error value meets the accuracy requirements ($e < 0.01$). If the accuracy requirements are met, the training is completed. Otherwise, the prediction results

Table 5 Part training sample data.

Channel	Eigenvalue	Sample1	Sample 2	Sample 3	Sample 4	Sample 5
Channel 1	WAMP	32	0	24	5	22
	RMS	21.3020	10.4378	24.7514	19.4136	22.7687
	PV	36.8406	11.8223	39.2490	29.2650	36.4782
Channel 2	WAMP	45	0	46	0	32
	RMS	26.6754	16.1020	26.8007	17.8136	24.4910
	PV	41.85703	16.8022	41.1995	20.2604	37.9465
	Action encoding	1	0	1	0	1

are compared with the sample target results, and then start neural Meta-feedback learning, repeat the above steps until reaching the specified number of training times or meet the accuracy requirements to complete the training.

Considering that BP neural network is prone to over training and lack of generalization ability, the training samples input into the neural network training algorithm are divided into three kinds of samples: train samples, validation samples and test samples. In each epoch of training, the errors between the results of three samples and the target results are tested. When the error of validation samples does not decrease in six successive epochs, the training of BP neural network is stopped to prevent over fitting, which is caused by overtraining of BP neural network. It can be seen from Fig. 7 that the total number of epochs of BP neural network is 116. After 110 epoch of BP neural network, the error of train samples, the error of test samples and the error of validation samples no longer have a downward trend, or their downward trend is not obvious. The best validation performance is $6.293e^{-6}$. Therefore, the training of BP neural network is finished at the 116th epoch. The threshold w is set 0.98, and the trained BP neural network is used to classify and recognize patient actions, the recognition result is shown in Fig. 8. Common classification performance measures are Precision (PRE), Recall (REC), and the harmonized average of the two (F_1).

According to Table 6, the calculation formula of Pre , Rec and F_1 :

$$Pre = \frac{TP}{TP+FP} \quad (10)$$

$$Rec = \frac{TP}{TP+FN} \quad (11)$$

$$F_1 = 2 \frac{PR \cdot REC}{PR + REC} \quad (12)$$

From Eqs. (10)–(12), $Pre = 1$, $Rec = 0.818$, $F_1 = 0.8998$.

Active trigger control strategy for pneumatic gloves

The software processing algorithm of the control system mainly includes a two-channel optimal eigenvalue amplitude calculation and a BP neural network action recognition

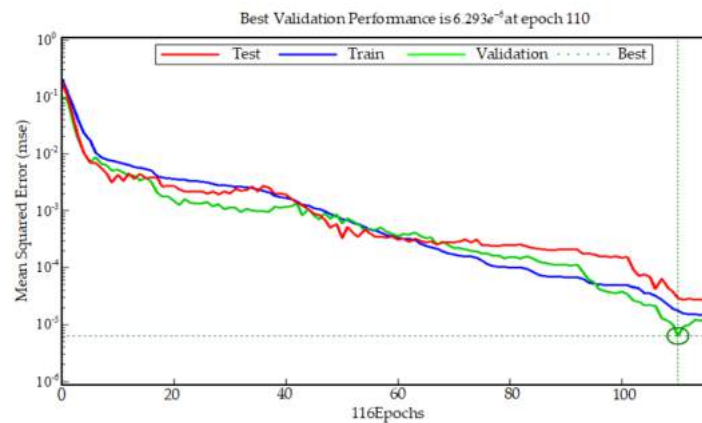


Figure 7 Best validation performance.

Full-size [DOI: 10.7717/peerjcs.448/fig-7](https://doi.org/10.7717/peerjcs.448/fig-7)

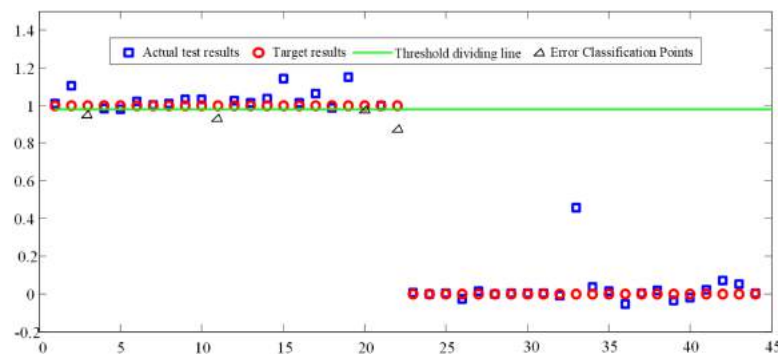


Figure 8 BP network motion classification results.

Full-size [DOI: 10.7717/peerjcs.448/fig-8](https://doi.org/10.7717/peerjcs.448/fig-8)

calculation. Among them, the same optimal eigenvalue is selected for different patients, and the eigenvalue amplitude calculation formula is unique. However, due to differences between individuals, the weights and thresholds of the nodes in the BP neural network model corresponding to different patients are not the same, so the BP neural network model library needs to be established in the actual application process. Different patients call their corresponding BP neural network models during training. When a patient conducts active training based on sEMG signals for the first time, he needs to collect sEMG signals under the guidance of a physician, and complete the training of the BP neural network, and store the required neural network in the BP neural network model library. The corresponding database will be called during a training session. The algorithm flow of active trigger control strategy for pneumatic rehabilitation gloves based on sEMG signals is shown in Fig. 9.

RESULTS

Now three male volunteers apply the above sEMG signal control strategy to identify the volunteer's hand movement trend to trigger the pneumatic rehabilitation gloves. Three

Table 6 Part training sample data.

Predetection result / real Result	Positive(+)	Negative(-)	Total
Positive(+)	18 (TP)	4 (FN)	22 (TP+FN)
Negative(-)	0 (FP)	22 (TN)	22 (FP+TN)
Total	18 (TP+FP)	26 (FN+TN)	

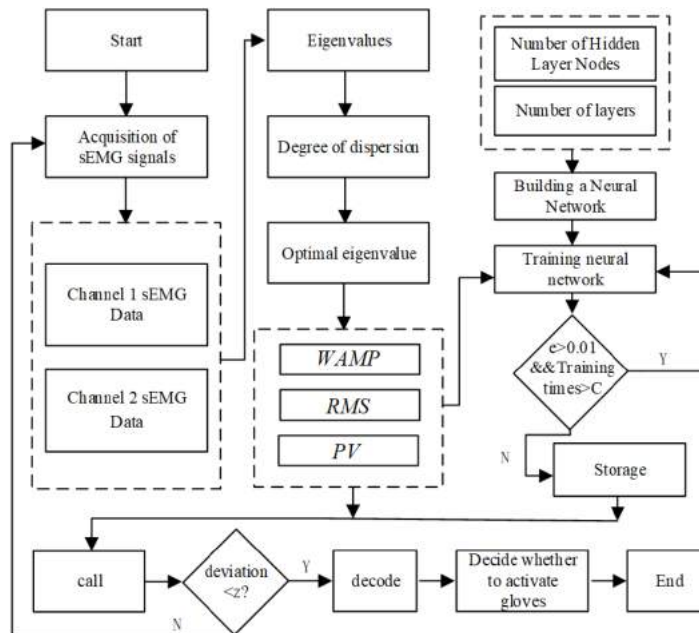


Figure 9 Algorithm flow chart of pneumatic glove trigger based on sEMG signals.

Full-size [DOI: 10.7717/peerjcs.448/fig-9](https://doi.org/10.7717/peerjcs.448/fig-9)

volunteers are required to complete the triggering of the pneumatic rehabilitation gloves six times within 100s, and the time from triggering to the completion of the training of a single pneumatic rehabilitation gloves should exceed 10s. The accuracy of the control system can be checked by completing the specified number of experiments within the specified time. The time to complete a single experiment is set to exceed 10s in order to make the extracted sEMG signal more intuitive. When the three volunteers realized the trigger control of the pneumatic gloves, the waveform diagram of the sEMG signal is shown in Figs. 10, 11 and 12. The surface EMG signal waveform without fluctuation in the figures indicates that the pneumatic rehabilitation gloves have not been triggered. At this time, the output of the control algorithm is 0. However, the combination of the trigger signal waveform and the motion signal waveform indicates that the pneumatic rehabilitation gloves are triggered to drive the patient's hand muscle movement. At this time, the output of the control algorithm is 1. All of the movement trends of the three volunteers were correctly identified, which indicates that the active triggering training based on sEMG signals may have universal applicability.

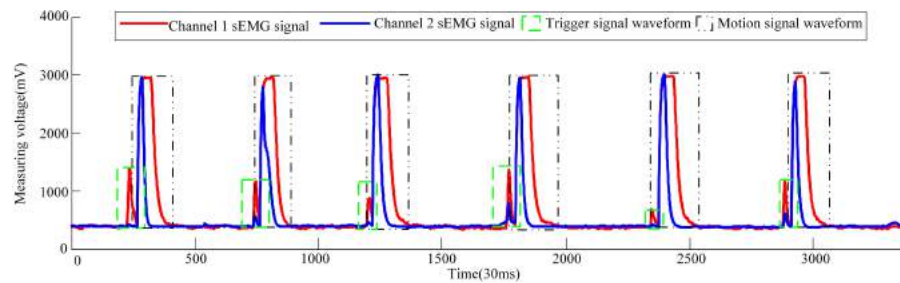


Figure 10 Volunteer 1's sEMG signal waveform when he attended pneumatic rehabilitation gloves triggering control.

[Full-size](#) DOI: 10.7717/peerjcs.448/fig-10

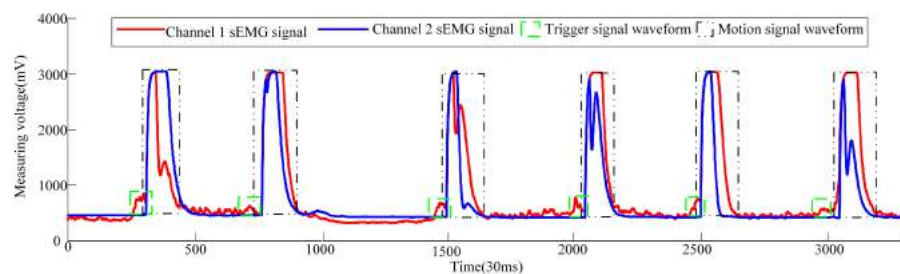


Figure 11 Volunteer 2's sEMG signal waveform when he attended pneumatic rehabilitation gloves triggering control.

[Full-size](#) DOI: 10.7717/peerjcs.448/fig-11

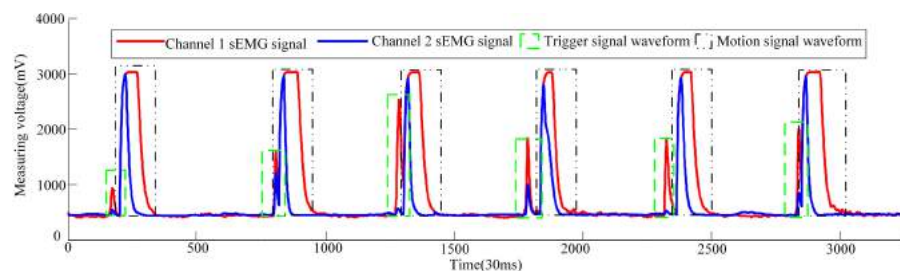


Figure 12 Waveforms of dual-channel sEMG signals when three volunteers attend pneumatic rehabilitation gloves triggering control.

[Full-size](#) DOI: 10.7717/peerjcs.448/fig-12

DISCUSSION

In order to realize active triggering training becoming possible in home rehabilitation, EMG is chosen as the bioelectrical signal for hand movement trend recognition, replacing the other high cost of collecting ECoG, EEG or MEG signals. The rehabilitation gloves' hardware platform can be divided into an acquisition layer, a decision-making layer, a driving layer and an execution layer.

The control system uses the BP neural network as a classifier for patient's hand movement trend recognition, and extracts the characteristic values of sEMG signals in the time domain:

MAV, *PV*, *WAMP*, *RMS*, *MS* and *MWL*, and then through the degree of dispersion index *R*, *Q*, *V* and *K*, the optimal eigenvalues of the sEMG signals are selected. By observing the sorting results of the data dispersion degree in [Table 1](#), three eigenvalues with the largest dispersion degree are selected, which are *WAMP*, *PV* and *RMS*. By observing [Tables 2](#) and [3](#), it can be seen that the most discrete eigenvalues extracted by samples W2 and W3 are *WAMP*, *PV* and *RMS*, which are the same as the optimal eigenvalues corresponding to the W1 sample. By comparing [Tables 1](#), [2](#), and [3](#), it can be seen that the order of the dispersion degree of each eigenvalues corresponding to different sub-samples is roughly the same. The magnitude of the dispersion index of the selected optimal eigenvalue is significantly higher than other eigenvalues. So it is reasonable to comprehensively select the optimal eigenvalues in the time domain as *WAMP*, *PV* and *RMS*.

WAMP, *PV* and *RMS* are used as the input values of the BP neural network. On the basis of the BP neural network which is used to establish the classifier of hand movement, the mapping relationship between hand sEMG signals and hand actions is finally completed by training and testing. From the [Fig. 8](#), when the actual test result is greater than w , the test result is equal to the action target result; when the test result is less than w , the test result is equal to the non-action target result. The accuracy of trend recognition is determined by judging whether the test result is equal to the corresponding target test result. A total of 44 judgments are made in the [Fig. 8](#), only 4 of which are wrong as shown by the triangle. Based on this, it can be considered that the correctness rate of BP judgment is about 90%. Based on [Fig. 7](#), the train correctness rate of BP judgment is about 99.997%. Judging the main reason for the distortion is closely related to factors such as the quality of the electrode paste, the state of the skin on the surface of the human body, and the changes in the muscle group during the sEMG acquisition process.

The pneumatic rehabilitation glove training control algorithm, based on sEMG signal, was proposed. By observing the sEMG signal waveforms of three volunteers, it can be found that when the BP neural network monitors the hand's movement trend, the pneumatic gloves will be triggered to drive the fingers to perform rehabilitation training. The difference in the amplitude and duration of the trigger signal of different volunteers in [Figs. 10](#), [11](#) and [12](#) is related to the volunteer's different physical quality, the duration and intensity of hand movement trend. Three male healthy volunteers used the control system to achieve the experimental results of the trigger experiment on pneumatic rehabilitation gloves, which preliminarily confirmed that the system has a high accuracy rate for hand movement trend recognition, and it may be useful in patient active hand training.

In the future, more healthy volunteers will be recruited to participate in this experiment. The generality and accuracy of this trigger control system for the recognition of different people's hand movement trend are tested in a larger range. Then stroke patients will be recruited to participate in the experiment to test. Comparison between the rehabilitation effect of traditional pneumatic rehabilitation robot and the ones with the trigger control system on stroke patients will be conducted. At last, the feasibility of applying the device to finger paralysis caused by different diseases will be considered. Meanwhile, we will also consider the effects of spasm, complete plegia and other factors on the accuracy of the trigger system.

CONCLUSIONS

An active trigger control system for pneumatic rehabilitation gloves, based on sEMG signals, is developed, which could achieve immediate rehabilitation movement trend to help the patient complete active hand rehabilitation training. Firstly, acquisition and processing of the sEMG signals from the human is researched, and three optimal eigenvalues of sEMG signals were selected as the follow-up neural network classification input. Then, based on BP neural network, the neural network classifier of hand movement is constructed. Moreover, the mapping relationship between hand sEMG signals and hand actions is built by training and testing. Based on the individual differences, the corresponding BP neural network model database of different people was established. At last, the pneumatic glove training control algorithm was proposed. Preliminary experiment shows that the combination of the trigger signal waveform and the motion signal waveform indicates that the pneumatic rehabilitation glove is triggered to drive the patient's hand movement. The device has high accuracy rate of trend recognition for hand movement. The above research could produce important scientific value for the development of robot technology and rehabilitation theory, provide theoretical basis and technical support for the control strategy of new hand rehabilitation robots.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This research was funded by Educational Commission of Zhejiang Province, grant number Y201941335; Research Fund Project of Ningbo University, grant number XYL19029; the K.C. Wong Magna Fund in Ningbo University; Natural Science Foundation of Ningbo City, grant number 2019A610110; Natural Science Foundation of Zhejiang Province, grant number LQ21E050008; Joint Funds of the National Natural Science Foundation of China, grant number U1913216; European Commission Marie Skłodowska-Curie SMOOTH project, grant number H2020-MSCA-RISE-2016-734875. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

Educational Commission of Zhejiang Province: Y201941335.

Research Fund Project of Ningbo University: XYL19029.

Ningbo University; Natural Science Foundation of Ningbo City: 2019A610110.

Natural Science Foundation of Zhejiang Province: LQ21E050008.

Joint Funds of the National Natural Science Foundation of China: U1913216.

European Commission Marie Skłodowska-Curie SMOOTH project: H2020-MSCA-RISE-2016-734875.

Competing Interests

Pengcheng Liu is an Academic Editor for PeerJ.

Author Contributions

- Yongfei Feng performed the experiments, performed the computation work, authored or reviewed drafts of the paper, and approved the final draft.
- Mingwei Zhong conceived and designed the experiments, performed the experiments, performed the computation work, authored or reviewed drafts of the paper, and approved the final draft.
- Xusheng Wang analyzed the data, prepared figures and/or tables, and approved the final draft.
- Hao Lu analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Hongbo Wang conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Pengcheng Liu conceived and designed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Luige Vladareanu conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

Ethical approval was granted by the Ethics Committee of Faculty of Mechanical Engineering & Mechanics, Ningbo University (No. [2020]LLSP(12)).

Data Availability

The following information was supplied regarding data availability:

Raw data and code are available in the [Supplemental Files](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.448#supplemental-information>.

REFERENCES

- Agarwal P, Fox J, Yun Y, O'Malley MK, Deshpande A. 2015.** An index finger exoskeleton with series elastic actuation for rehabilitation. *International Journal of Robotics Research* **34**:1747–1772 DOI [10.1177/0278364915598388](https://doi.org/10.1177/0278364915598388).
- Bentzvi P, Ma Z. 2014.** Sensing and Force-Feedback Exoskeleton (SAFE) Robotic Glove. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **23**:992–1002 DOI [10.1109/TNSRE.2014.2378171](https://doi.org/10.1109/TNSRE.2014.2378171).
- Dai CY, Hu XG. 2020.** Finger joint angle estimation based on motoneuron discharge activities. *IEEE Journal of Biomedical and Health Informatics* **24**:760–767 DOI [10.1109/JBHI.2019.2926307](https://doi.org/10.1109/JBHI.2019.2926307).
- Gaia VP, Jeanette P, Loïc C, Pauline R, Jean-Claude B, Elena P, Jörgen B, Pâvel GL. 2020.** Recovery and prediction of dynamic precision grip force control after stroke. *Stroke* **51**:944–951 DOI [10.1161/STROKEAHA.119.026205](https://doi.org/10.1161/STROKEAHA.119.026205).

- Hecht-Nielsen . 1992.** Theory of the backpropagation neural network. *Neural Networks for Perception* **1992**:65–93 DOI [10.1016/B978-0-12-741252-8.50010-8](https://doi.org/10.1016/B978-0-12-741252-8.50010-8).
- Heung HL, Tang ZQ, Shi XQ, Tong RKY, Li Z. 2020.** Soft rehabilitation actuator with integrated post-stroke finger spasticity evaluation. *Frontiers in Bioengineering and Biotechnology* **8**:1–10 DOI [10.3389/fbioe.2020.00111](https://doi.org/10.3389/fbioe.2020.00111).
- Heung KHL, Tong RKY, Lau ATH, Li Z. 2019.** Robotic glove with soft-elastic composite actuators for assisting activities of daily living. *Soft Robotics* **6**:289–304 DOI [10.1089/soro.2017.0125](https://doi.org/10.1089/soro.2017.0125).
- Lemerle S, Nozaki T, Ohnishi K. 2018.** Design and evaluation of a remote actuated finger exoskeleton using motion-copying system for tendon rehabilitation. *IEEE Transactions on Industrial Informatics* **14**:5167–5177 DOI [10.1109/TII.2018.2796859](https://doi.org/10.1109/TII.2018.2796859).
- Leonardis D, Barsotti M, Loconsole C, Solazzi M, Troncossi M, Mazzotti C, Castelli VP, Procopio C, Lamola G, Chisari G, Bergamasco M, Frisoli A. 2015.** An EMG-controlled robotic hand exoskeleton for bilateral rehabilitation. *IEEE Transactions on Haptics* **8**:140–151 DOI [10.1109/TOH.2015.2417570](https://doi.org/10.1109/TOH.2015.2417570).
- Leonardo C, Meyer JT, Galloway KC, Peisner JD, Ganberry R, Wagner DA, Sven E, Sabrina P, Walsh CJ. 2018.** Assisting hand function after spinal cord injury with a fabric-based soft robotic glove. *Journal of Neuroengineering and Rehabilitation* **15**:1–10 DOI [10.1186/s12984-018-0391-x](https://doi.org/10.1186/s12984-018-0391-x).
- Liu Y, Cheng L. 2018.** Spiking-neural-network based Fugl-Meyer hand gesture recognition for wearable hand rehabilitation robot. <https://ieeexplore.ieee.org/document/8489141> (accessed on 8-13 July 2018).
- Lyu MX, Lambelet C, Woolley D, Zhang X, Chen WH, Ding XL, Gassert R, Wnderoth N. 2020.** Comparison of particle filter to established filtering methods in electromyography biofeedback. *Biomedical Signal Processing and Control* **60**:1–15 DOI [10.1016/j.bspc.2020.101949](https://doi.org/10.1016/j.bspc.2020.101949).
- Mahdi HJ, Charu P, Muthu BJW. 2019.** Soft robotic bilateral hand rehabilitation system for fine motor learning. Available at <https://doi.org/10.1109/ICORR.2019.87795101> (accessed on 24-28 June 2019).
- Matthew CHC, Jeong HL, Raye CHY. 2019.** Design and characterization of a soft robotic therapeutic glove for rheumatoid arthritis. *Assistive Technology* **31**:44–52 DOI [10.1080/10400435.2017.1346000](https://doi.org/10.1080/10400435.2017.1346000).
- Meng W, Liu Q, Zhou ZD, Ai QS. 2014.** Active interaction control applied to a lower limb rehabilitation robot by using EMG recognition and impedance model. *Industrial Robot-An International Journal* **41**:465–479 DOI [10.1108/IR-04-2014-0327](https://doi.org/10.1108/IR-04-2014-0327).
- Nycz CJ, Bützer T, Lambercy O, Arata J, Fischer GS, Gasser R. 2016.** Design and characterization of a lightweight and fully portable remote actuation system for use with a hand exoskeleton. *IEEE Robotics and Automation Letters* **1**:976–983 DOI [10.1109/LRA.2016.2528296](https://doi.org/10.1109/LRA.2016.2528296).
- Pichiorri F, Morone G, Petti M, Toppi J, Pisotta I, Molinari M, Paolucci S, Inghileri M, Astolfi L, Cincotti F, Mattia D. 2015.** Brain-computer interface boosts motor imagery practice during stroke recovery. *Annals of Neurology* **77**:851–865 DOI [10.1002/ana.24390](https://doi.org/10.1002/ana.24390).

- Polygerinos P, Wang Z, Galloway KC, Wood RJ, Walsh CJ. 2015.** Soft robotic glove for combined assistance and at-home rehabilitation. *Robotics and Autonomous Systems* 73:135–143 DOI [10.1016/j.robot.2014.08.014](https://doi.org/10.1016/j.robot.2014.08.014).
- Sangwoo P, Cassie M, Lynne MW, Lauri B, Joel S, Matei C. 2018.** Multimodal sensing and interaction for a robotic hand orthosis. *IEEE Robotics and Automation Letters* 4:315–322 DOI [10.1109/LRA.2018.2890199](https://doi.org/10.1109/LRA.2018.2890199).
- Sheela KG, Deepa SN. 2013.** Review on methods to fix number of hidden neurons in neural networks. *Mathematical Problems in Engineering* 2013:1–11 DOI [10.1155/2013/425740](https://doi.org/10.1155/2013/425740).
- Shi X, Qin PJ, Zhu JQ, Xu SY, Shi WR. 2020.** Lower limb motion recognition method based on improved wavelet packet transform and unscented kalman neural network. *Mathematical Problems in Engineering* 2020:1–16 DOI [10.1155/2020/5684812](https://doi.org/10.1155/2020/5684812).
- Song R, Tong KY, Hu XL, Li L. 2008.** Assistive control system using continuous myoelectric signal in robot-aided arm training for patients after stroke. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 16:371–379 DOI [10.1109/TNSRE.2008.926707](https://doi.org/10.1109/TNSRE.2008.926707).
- Wang JB, Fei YQ, Pang W. 2019.** Design, modeling, and testing of a soft pneumatic glove with segmented PneuNets bending actuators. *IEEE/ASME Transactions on Mechatronics* 24:990–1001 DOI [10.1109/TMECH.2019.2911992](https://doi.org/10.1109/TMECH.2019.2911992).
- Wang DM, Wang YK, Zi B, Cao ZX, Ding HF. 2020a.** Development of an active and passive finger rehabilitation robot using pneumatic muscle and magnetorheological damper. *Mechanism and Machine Theory* 147:1–16 DOI [10.1016/j.mechmachtheory.2019.103762](https://doi.org/10.1016/j.mechmachtheory.2019.103762).
- Wang Y, Wu Q, Dey N, Fong S, Ashour A. 2020b.** Deep back propagation-long short-term memory network based upper-limb sEMG signal classification for automated rehabilitation. *Biocybernetics and Biomedical Engineering* 40:987–1001 DOI [10.1016/j.bbe.2020.05.003](https://doi.org/10.1016/j.bbe.2020.05.003).
- Wu J, Huang J, Wang YJ, Xing KX. 2010.** A wearable rehabilitation robotic hand driven by PMTS actuators. *Intelligent Robotics and Applications* 6425:440–450 DOI [10.1007/978-3-642-16587-0_41](https://doi.org/10.1007/978-3-642-16587-0_41).
- Yurkewich A, Illya JK, Debbie H, Rosalie HW, Alex M. 2020.** Hand extension robot orthosis (HERO) grip glove: enabling independence amongst persons with severe hand impairments after stroke. *Journal of NeuroEngineering and Rehabilitation* 17:1–7 DOI [10.1186/s12984-020-00659-5](https://doi.org/10.1186/s12984-020-00659-5).

Ultrawideband (UWB)-based precise short-range localization for wireless power transfer to electric vehicles in parking environments

Seung-Mok Lee

Department of Automotive System Engineering, Keimyung University, Daegu, South Korea

ABSTRACT

As the necessity of wireless charging to support the popularization of electric vehicles (EVs) emerges, the development of a wireless power transfer (WPT) system for EV wireless charging is rapidly progressing. The WPT system requires alignment between the transmitter coils installed on the parking lot floor and the receiver coils in the vehicle. To automatically align the two sets of coils, the WPT system needs a localization technology that can precisely estimate the vehicle's pose in real time. This paper proposes a novel short-range precise localization method based on ultrawideband (UWB) modules for application to WPT systems. The UWB module is widely used as a localization sensor because it has a high accuracy while using low power. In this paper, the minimum number of UWB modules consisting of two UWB anchors and two UWB tags that can determine the vehicle's pose is derived through mathematical analysis. The proposed localization algorithm determines the vehicle's initial pose by globally optimizing the collected UWB distance measurements and estimates the vehicle's pose by fusing the vehicle's wheel odometry data and the UWB distance measurements. To verify the performance of the proposed UWB-based localization method, we perform various simulations and real vehicle-based experiments.

Subjects Algorithms and Analysis of Algorithms, Autonomous Systems, Robotics

Keywords Electric vehicles, Localization, Ultrawideband (UWB), Vehicle pose estimation, Wireless power transfer systems

INTRODUCTION

The global electric vehicle (EV) market is growing rapidly due to the strengthening of international environmental regulations on vehicle emissions. The technical limitation that should be overcome to accelerate the popularization of EVs is the poor mileage. To this end, the capacity of the battery should be increased, but the current technology does not reach the mileage of internal combustion engine vehicles with a single charge. In addition, the charging time is too long. To compensate for this problem, a wireless power transfer (WPT) system that can easily charge EVs in a parking lot space has been proposed (El-Shahat et al., 2019; Liang et al., 2020; Machura, Santis & Li, 2020; Panchal, Stegen & Lu, 2018). When WPT systems are installed in parking lots, charging can be easily performed frequently without building a separate charging station, thereby compensating for problems caused by battery capacity limitations. In addition, when combined with an

Submitted 10 March 2021

Accepted 7 May 2021

Published 27 May 2021

Corresponding author

Seung-Mok Lee,
seungmok@kmu.ac.kr

Academic editor

Pengcheng Liu

Additional Information and
Declarations can be found on
page 19

DOI 10.7717/peerj-cs.567

© Copyright
2021 Lee

Distributed under
Creative Commons CC-BY-NC 4.0

OPEN ACCESS

autonomous parking system, an advanced driver assistance system (ADAS), the use of EVs becomes easier because these vehicles can park and charge themselves. The WPT consists of power-transmitting coils on the parking lot floor and power-receiving coils for the car, and to be charged, the EV must precisely recognize the car location within the parking area and align the two sets of coils to centimeter-level accuracy (Rozman et al., 2019). Therefore, for the WPT system to be combined with an autonomous parking system, precise localization technology is required in the parking area (Tian et al., 2020; Shin et al., 2019).

Vehicle localization technologies have been developed with different sensors and different methods for indoor or outdoor environments. Localization in outdoor environments uses the Global Positioning System (GPS) and vision sensors with high definition (HD) maps. Localization in indoor environments uses a vision sensor or lidar with prebuilt feature maps or grid maps, as GPS signals are unavailable in these environments. It is thus difficult to apply these conventional localization methods to the WPT system because a map cannot be constructed for all indoor environments.

Recently, many studies have been conducted to utilize an ultrawideband (UWB) distance sensor for vehicle localization technology (Stoll et al., 2017; Tiemann et al., 2016; Kukolev et al., 2016; Alarifi et al., 2016). In Stoll et al. (2017), extended Kalman filter (EKF)-based vehicle localization using one UWB tag mounted on the vehicle and multiple UWB anchors placed in an outdoor parking space was proposed. The method requires many UWB anchors to be installed in the outdoor space to enhance the accuracy of vehicle localization. This method shows that the average position error is approximately 0.23 m when seven anchors are used. In Tiemann et al. (2016), a UWB-based precise localization method for application to WPT was proposed. The method is based on an EKF using UWB distance measurements from one UWB tag on the vehicle and two UWB anchors placed in the corners of a parking slot. When the vehicle moves straightforward and approaches the parking slot, the localization accuracy is approximately 0.1 m near the anchors, demonstrating the possibility of applying UWB technology to precise localization for the WPT. However, since only one tag is used to estimate the vehicle's state, only the 2D position can be estimated, and the vehicle's heading cannot be estimated. In addition, since the localization method uses only UWB distance measurements, it is very vulnerable to UWB measurement noise. In Kukolev et al. (2016), a localization method based on one UWB anchor in a parking lot and two UWB tags on a vehicle was proposed. The method presented in Kukolev et al. (2016) can estimate the position while the vehicle is stationary. However, there is a limitation that an area where position estimation is not possible exists depending on the heading angle of the vehicle.

The UWB sensor can also be used for the localization of various objects, such as mobile robots (Chen et al., 2020; Shi et al., 2020), flying drones (Wang, Marelli & Fu, 2021; González-Castaño et al., 2021; Hyun et al., 2019), and users (Zhang et al., 2019; Knobloch, 2017), because it can provide precise distance measurements based on the time-of-flight (TOF) principle at short-range regions while using low power. In addition, since the UWB distance sensor is inexpensive, the UWB-based localization system can advantageously be implemented at an economical cost, even though multiple UWB sensor modules are used.

This paper proposes a novel short-range precise localization method based on a dual-anchor and dual-tag (DADT) UWB system that can be applied to WPT systems. The proposed DADT UWB-based localization method uses two UWB anchors placed in the parking area and two UWB tags mounted on the vehicle. When the vehicle approaches the parking slot where the WPT is located, the UWB anchors start to communicate with the tags, and the vehicle's pose, i.e., position and heading angle, is initialized by processing the UWB distance measurement data. Then, the wheel odometry information and UWB distance measurements are fused based on a particle filter framework to continuously estimate the pose from the initial vehicle pose. The goal of this paper is to make a precise pose estimation so that the final parking position of the vehicle has an error of less than 0.1 m, which is required for alignment between the power transmitter and receiver coils of the WPT. To verify the performance of the proposed DADT UWB-based localization method, we perform various simulations and experiments with an actual vehicle.

The preliminary results of this paper were presented in [Lee \(2020\)](#). In the preliminary results, the theoretical analysis of the minimum number of UWB modules and their placement was not performed sufficiently. Compared with the results in [Lee \(2020\)](#), the contributions of this paper can be summarized as follows. This paper provides the detailed DADT UWB localization system with rigorous theoretical analysis. It is shown mathematically that the proposed DADT method can uniquely determine the pose of the vehicle with only two anchors and two tags. Additionally, we analyze the observability of the proposed DADT method based on the Fisher information matrix (FIM). From the analysis, it is confirmed that the DADT UWB system is the minimal combination of UWB anchors and tags satisfying the condition for the DADT UWB localization system to be fully observable. In addition, more detailed simulation and experimental results are provided to show the effectiveness of the proposed method.

The rest of this paper is organized as follows: we introduce the WPT system for EVs and describe the proposed DADT UWB localization method. To verify the performance of the proposed method, simulation results with various scenarios and experimental results with a real vehicle are presented. Finally, a conclusion is presented.

LOCALIZATION FOR WIRELESS POWER TRANSFER (WPT) SYSTEMS

The basic working principle of WPT for EVs is as follows ([González-Castaño et al., 2021](#)). The WPT consists of electric power transmitter coils and electric power receiver coils, as shown in [Fig. 1](#). Power transmitter coils are installed on the floor of the parking lot, and power receiver coils are mounted underneath the vehicle. When the transmitter coil and the receiver coil are kept close to each other while maintaining a certain distance, electric power is transmitted to the receiver coils, and electric energy can be used to charge the battery. The alignment of the transmitter and receiver coils is significant for the high performance and efficiency of WPT.

Localization technologies that can be applied in parking lot environments have been developed based on mono cameras ([Hu et al., 2019](#); [Panev et al., 2019](#); [Yu et al., 2020](#)),

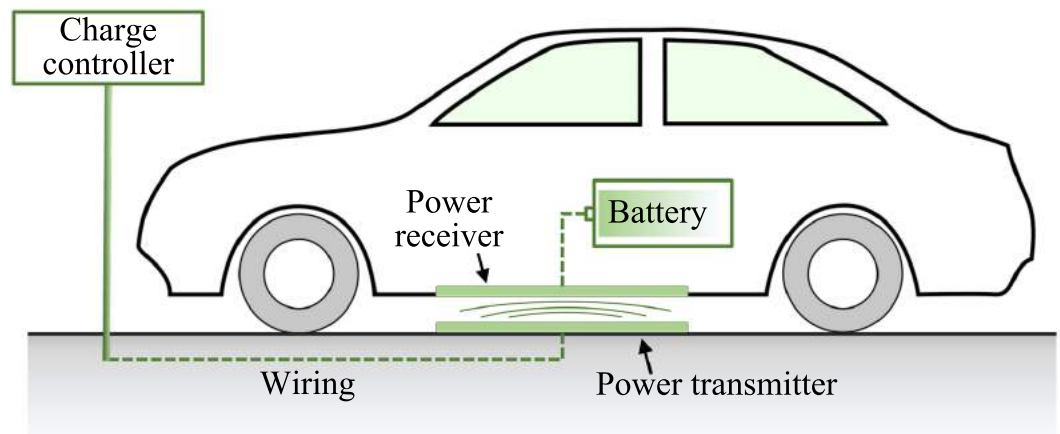


Figure 1 Wireless power transfer (WPT) for electric vehicles (EVs). When the power transmitter and receiver coils are kept close to each other, electric energy can be used to charge the battery. The alignment of the two coils is significant for the high performance and efficiency of the WPT.

Full-size [DOI: 10.7717/peerjcs.567/fig-1](https://doi.org/10.7717/peerjcs.567/fig-1)

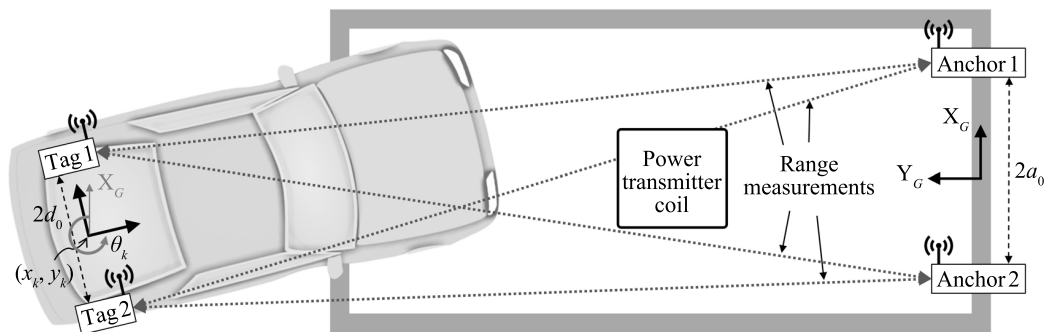


Figure 2 Dual-anchor and dual-tag (DADT) ultrawideband (UWB)-based localization system. Two UWB anchors are placed on both corners of the parking slot and two UWB tags are mounted on the vehicle. The observability of the vehicle's pose estimated by the proposed DADT UWB system is analyzed based on the Fisher information matrix.

Full-size [DOI: 10.7717/peerjcs.567/fig-2](https://doi.org/10.7717/peerjcs.567/fig-2)

depth cameras (Zhao et al., 2020), lidar (Tao et al., 2018), and radio frequency (RF) fingerprinting of WiFi signals (Gao, He & Li, 2018), but they still have some limitations. Most camera-based localization technologies are based on parking line recognition. However, parking lines are usually not standardized and may not even be drawn. These methods are also sensitive to lighting changes at night and do not operate well in dark indoor parking lots. Lidar is useful for finding vacant parking spaces, but it is difficult to estimate the relative position from the transmitting coil because it is difficult for lidar to recognize the position of the transmitting coil.

The use of UWB sensors can overcome the limitations of cameras and lidar sensors. The proposed DADT localization method only needs to know the positions of two anchors and

the transmitting coil installed in the parking slot. Thus, it does not require an inconvenient process of building a high-precision map with cameras or lidar sensors. As the UWB sensor is based on RF signals, it is unaffected by changes in lighting and robust against dynamic obstacles such as vehicles and pedestrians. In addition, if a pair of UWB sensors has a clear line of sight, the distance between them can be precisely measured with an error of approximately 0.05 m- 0.1 m. Due to the economics of UWB sensors, many automotive makers have plans to use UWB sensors in vehicles soon. Therefore, it is possible to implement a precise localization applicable to WPT for EVs with economical cost.

PROPOSED ULTRAWIDEBAND (UWB)-BASED LOCALIZATION

Dual-Anchor and dual-tag (DADT) UWB localization and observability analysis

This section describes a novel dual-anchor and dual-tag (DADT) UWB-based localization method that can precisely estimate the pose of a vehicle near a parking area. The key idea of the proposed DADT UWB-based localization is shown in Fig. 2. Two anchors are installed on the charging station, and two tags are mounted on the vehicle. The two tags should be placed so that the pose of the vehicle is always observable with only the distance measurements of the DADT UWB system. The location of the UWB anchors should be known in advance. Thus, the two anchors are placed on both corners of the parking slot so that the location of the anchor can be easily identified.

To show the effectiveness of the proposed DADT UWB sensor system, the condition in which the vehicle's pose can be uniquely determined by the DADT UWB sensor system is analytically derived. Then, observability analysis based on FIM is performed on the proposed DADT UWB sensor system.

Existence and Uniqueness Solution to DADT UWB localization

We denote the vehicle pose state vector at a time step k by $\mathbf{x}_k = [x_k, y_k, \theta_k]^T$, the two UWB anchor position vectors by $\mathbf{a}_1 = [a_{1x}, a_{1y}]^T$ and $\mathbf{a}_2 = [a_{2x}, a_{2y}]^T$, and the two UWB tag position vectors by $\mathbf{t}_1 = [t_{1x}, t_{1y}]^T$ and $\mathbf{t}_2 = [t_{2x}, t_{2y}]^T$. For simplicity, let us assume that the two anchors are placed at $\mathbf{a}_1 = [a_0, 0]^T$ and $\mathbf{a}_2 = [-a_0, 0]^T$, as shown in Fig. 2. The position vectors of the two tags can then be represented in the global frame (X_G, Y_G) as

$$\mathbf{t}_1 = \begin{bmatrix} t_{1x} \\ t_{1y} \end{bmatrix} = \begin{bmatrix} x_k - d_0 \sin \theta_k \\ y_k + d_0 \cos \theta_k \end{bmatrix}, \quad (1)$$

$$\mathbf{t}_2 = \begin{bmatrix} t_{2x} \\ t_{2y} \end{bmatrix} = \begin{bmatrix} x_k + d_0 \sin \theta_k \\ y_k - d_0 \cos \theta_k \end{bmatrix}. \quad (2)$$

The measurement model of the DADT UWB system can then be derived as

$$h(\mathbf{x}_t, a_0, d_0) = \begin{bmatrix} \sqrt{(x_k - d_0 \sin \theta_k - a_0)^2 + (y_k + d_0 \cos \theta_k)^2} \\ \sqrt{(x_k + d_0 \sin \theta_k - a_0)^2 + (y_k - d_0 \cos \theta_k)^2} \\ \sqrt{(x_k - d_0 \sin \theta_k + a_0)^2 + (y_k + d_0 \cos \theta_k)^2} \\ \sqrt{(x_k + d_0 \sin \theta_k + a_0)^2 + (y_k - d_0 \cos \theta_k)^2} \end{bmatrix}, \quad (3)$$

and Eq. (3) can be represented as

$$z_{11}^2 = (x_k - d_0 \sin \theta_k - a_0)^2 + (y_k + d_0 \cos \theta_k)^2 \quad (4a)$$

$$z_{12}^2 = (x_k + d_0 \sin \theta_k - a_0)^2 + (y_k - d_0 \cos \theta_k)^2 \quad (4b)$$

$$z_{21}^2 = (x_k - d_0 \sin \theta_k + a_0)^2 + (y_k + d_0 \cos \theta_k)^2 \quad (4c)$$

$$z_{22}^2 = (x_k + d_0 \sin \theta_k + a_0)^2 + (y_k - d_0 \cos \theta_k)^2, \quad (4d)$$

where z_{ij} is the distance measurement value by the i th anchor and j th tag. The vehicle pose as determined by the DADT UWB system can be found by solving Eqs. (4a)–(4d) for $\mathbf{x}_k = [x_k, y_k, \theta_k]^T$. From Eqs. (4a) and (4c), x_k and y_k can be uniquely determined by

$$x_k = d_0 \sin \theta_k + \frac{1}{4a_0} (z_{21}^2 - z_{11}^2) \quad (5a)$$

$$y_k = -d_0 \cos \theta_k + \sqrt{z_{11}^2 - \left[\frac{1}{4a_0} (z_{21}^2 - z_{11}^2) - a_0 \right]^2}, \quad (5b)$$

under the condition that

$$y_k + d_0 \cos \theta_k > 0. \quad (6)$$

In a similar manner, using Eqs. (4b) and (4d), x_k and y_k have another equivalent form as follows:

$$x_k = -d_0 \sin \theta_k + \frac{1}{4a_0} (z_{22}^2 - z_{12}^2) \quad (7a)$$

$$y_k = d_0 \cos \theta_k + \sqrt{z_{12}^2 - \left[\frac{1}{4a_0} (z_{22}^2 - z_{12}^2) - a_0 \right]^2}, \quad (7b)$$

under the condition that

$$y_k - d_0 \cos \theta_k > 0. \quad (8)$$

The two conditions Eqs. (6) and (8) can be satisfied the two tags are on the $Y_G > 0$ region. Subtracting Eqs. (7a) from (5a) and rearranging with respect to $\sin \theta_k$ gives

$$\sin \theta_k = \frac{1}{8a_0 d_0} (z_{11}^2 - z_{12}^2 - z_{21}^2 + z_{22}^2). \quad (9)$$

From Eqs. (5b) and (7b), $\cos \theta_k$ can be expressed by

$$\cos\theta_k = \frac{1}{2d_0} \left(\sqrt{z_{11}^2 - \left[\frac{1}{4a_0} (z_{21}^2 - z_{11}^2) - a_0 \right]^2} - \sqrt{z_{12}^2 - \left[\frac{1}{4a_0} (z_{22}^2 - z_{12}^2) - a_0 \right]^2} \right). \quad (10)$$

Using Eqs. (9) and (10), θ_k can be found as

$$\theta_k = \arctan \left(\frac{1}{4a_0} \frac{z_{11}^2 - z_{12}^2 - z_{21}^2 + z_{22}^2}{\sqrt{z_{11}^2 - \left[\frac{1}{4a_0} (z_{21}^2 - z_{11}^2) - a_0 \right]^2} - \sqrt{z_{12}^2 - \left[\frac{1}{4a_0} (z_{22}^2 - z_{12}^2) - a_0 \right]^2}} \right). \quad (11)$$

Therefore, the vehicle pose $\mathbf{x}_k = [x_k, y_k, \theta_k]^T$ can be uniquely determined by Eqs. (7a), (7b) and Eq. (11) under the condition that $y_k > d_0$.

Observation analysis based on the Fisher information matrix (FIM)

The uncertainty of the pose of the vehicle estimated by UWB distance measurements is determined by the geometric distribution of the anchors fixed on the parking lot and the tags mounted on the vehicle. To estimate the amount of uncertainty about the vehicle pose estimated by the proposed DADT UWB system, FIM-based observability analysis is performed as follows. The FIM can be defined as (Lee et al., 2015)

$$\mathcal{F} \triangleq \mathbf{H}^T \mathbf{W}^{-1} \mathbf{H} \quad (12)$$

where \mathbf{H} is the Jacobian of $h(\mathbf{x}_t, a_0, d_0)$ in Eq. (3) with respect to the vehicle state \mathbf{x}_k and \mathbf{W} is a covariance matrix of the UWB measurement noise. The sufficient and necessary condition for the DADT UWB localization system to be fully observable is that the FIM defined in Eq. (12) should be positive definite. The positive definiteness of the FIM is equivalent to the full column rank of the Jacobian \mathbf{H} . Therefore, it can be found that the DADT UWB localization system is fully observable from the UWB distance measurements if and only if \mathbf{H} has a full column rank. The Jacobian \mathbf{H} can be computed as

$$\begin{aligned} \mathbf{H} &= \frac{\partial h(\mathbf{x}_k, a_0, d_0)}{\partial \mathbf{x}_k} \\ &= \begin{bmatrix} \frac{1}{m_{11}} (x_k - a_0 - d_0 \sin\theta_k) & \frac{1}{m_{12}} (x_k - a_0 + d_0 \sin\theta_k) \\ \frac{1}{m_{11}} (y_k + d_0 \cos\theta_k) & \frac{1}{m_{12}} (y_k - d_0 \cos\theta_k) \\ -\frac{d_0}{m_{11}} ((x_k - a_0) \cos\theta_k + y_k \sin\theta) & \frac{d_0}{m_{12}} ((x_k - a_0) \cos\theta_k + y_k \sin\theta) \\ \frac{1}{m_{21}} (x_k + a_0 - d_0 \sin\theta_k) & \frac{1}{m_{22}} (x_k + a_0 + d_0 \sin\theta_k) \\ \frac{1}{m_{21}} (y_k + d_0 \cos\theta_k) & \frac{1}{m_{22}} (y_k - d_0 \cos\theta_k) \\ -\frac{d_0}{m_{21}} ((x_k + a_0) \cos\theta_k + y_k \sin\theta) & \frac{d_0}{m_{22}} ((x_k + a_0) \cos\theta_k + y_k \sin\theta) \end{bmatrix}. \quad (13) \end{aligned}$$

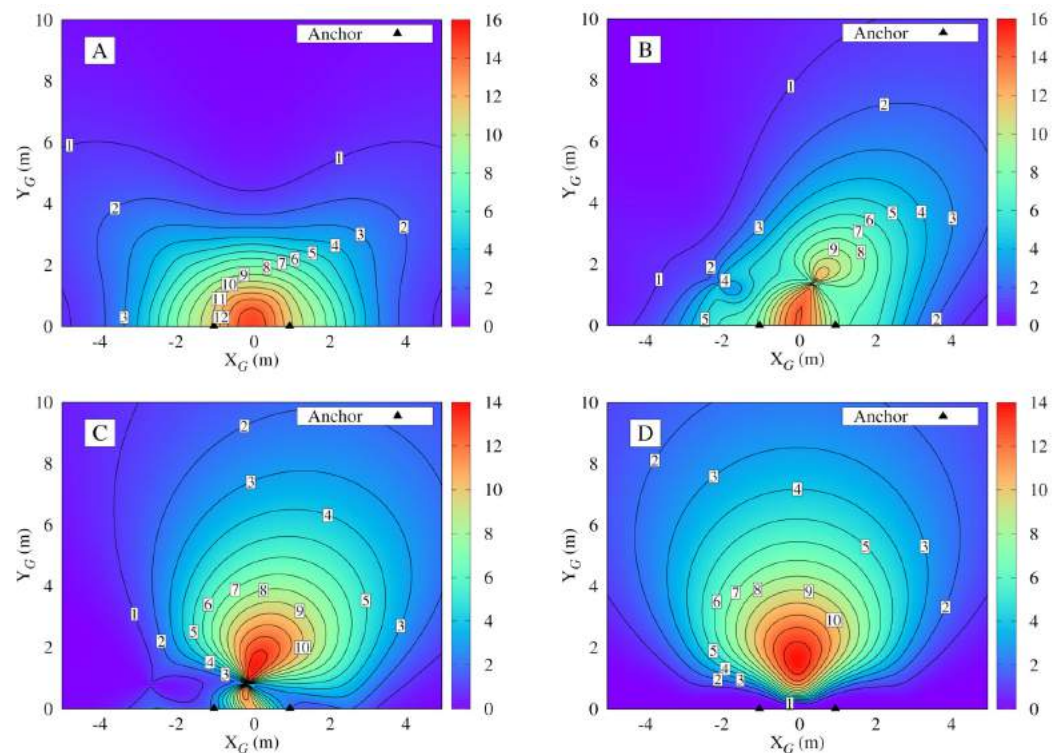


Figure 3 Numerical distribution of the determinant of the Fisher information matrix in Eq. (12) around the UWB anchors placed at $\mathbf{a}_1 = [1, 0]^T$ and $\mathbf{a}_2 = [-1, 0]^T$ with different vehicle's heading angles for every position. (A) $\theta_k = 0$ deg or 180 deg; (B) $\theta_k = 30$ deg or 210 deg; (C) $\theta_k = 60$ deg or 240 deg; (D) $\theta_k = 90$ deg or 270 deg.

Full-size [DOI: 10.7717/peerjcs.567/fig-3](https://doi.org/10.7717/peerjcs.567/fig-3)

The Jacobian presented in Eq. (13) can also be represented in a reduced row echelon form using the Gaussian elimination process as follows:

$$\begin{bmatrix} 1 & 0 & 0 & \frac{-m_{11}(y_k - d \cos \theta_k)((x_k + a) \cos \theta_k + y_k \sin \theta_k)}{m_{22}(y_k + d \cos \theta_k)((x_k - a) \cos \theta_k + y_k \sin \theta_k)} \\ 0 & 1 & 0 & \frac{m_{12}((x_k + a) \cos \theta_k + y_k \sin \theta_k)}{m_{22}((x_k - a) \cos \theta_k + y_k \sin \theta_k)} \\ 0 & 0 & 1 & \frac{m_{21}(y_k - d \cos \theta_k)}{m_{22}(y_k + d \cos \theta_k)} \end{bmatrix}. \quad (14)$$

As seen in Eq. (14), the column rank is 3, which is the full rank.

The determinant value of the FIM represents the amount of Fisher information that can be observed for vehicle pose state variables; i.e., as the determinant of the FIM increases, the vehicle pose can be estimated with higher accuracy. Figure 3 shows the numerical distribution of the determinant of the FIM around the UWB anchors placed at $\mathbf{a}_1 = [1, 0]^T$ and $\mathbf{a}_2 = [-1, 0]^T$, when the vehicle heading is $\theta_k = 0$ deg, 30 deg, 60 deg, and 90 deg for every position. In Fig. 3, the covariance matrix \mathbf{W} in Eq. (12) is assumed to be an identity matrix in order to compare the determinant value depending only on the positions. As shown in Fig. 2, when the vehicle approaches the parking slot, the heading angle θ_k is 90

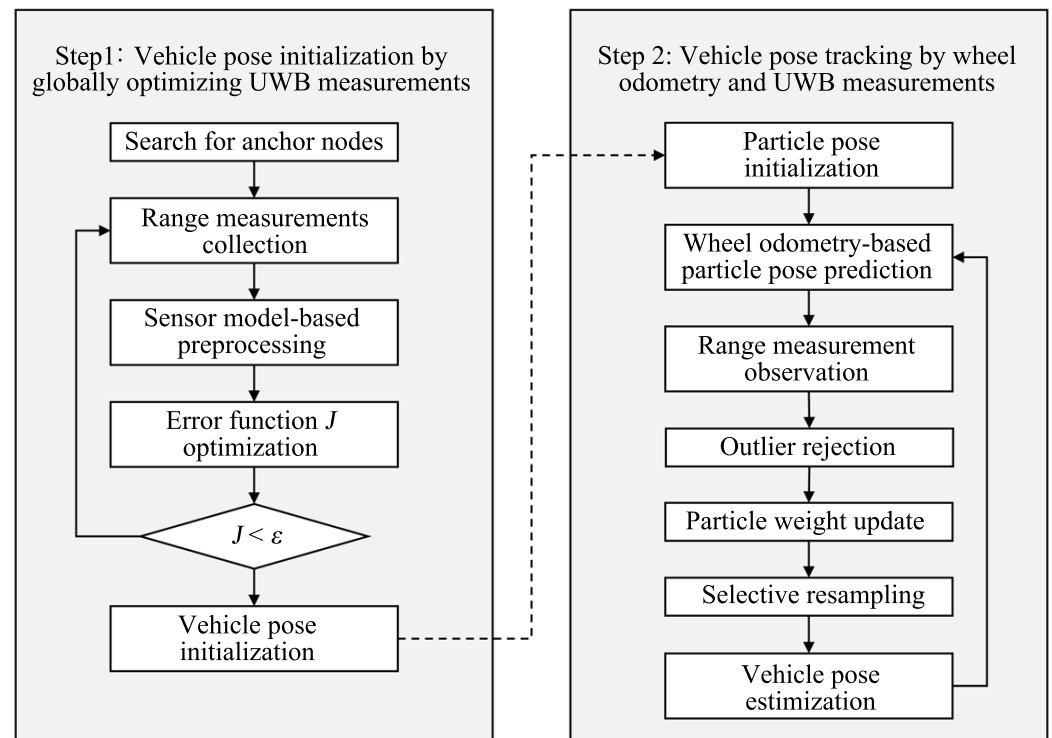


Figure 4 Flowchart of the proposed DADT UWB localization algorithm. In the first step, the vehicle's pose is initialized by globally optimizing the UWB distance measurements. In the second step, based on the initialized vehicle pose, the wheel odometry and UWB distance measurement collected as the vehicle moves are fused to estimate the pose of the vehicle in real time.

Full-size DOI: 10.7717/peerjcs.567/fig-4

deg for back-in parking and 270 deg for front-end parking. By comparing Fig. 3, it can be seen that when the heading angle is 90 deg or 270 deg, the determinant of the FIM has the largest distribution. This means that for anchors placed at both corners of the parking lot slot, installing the two tags in a direction perpendicular to the vehicle's moving direction maximizes the amount of Fisher information under the assumption that back-in parking ($\theta_k = 90$ deg) or front-end parking ($\theta_k = 270$ deg) is performed.

Vehicle pose estimation based on the DADT UWB system

The proposed DADT UWB localization algorithm consists of two major steps, as shown in Fig. 4. In the first step, when the EV approaches the parking area where the UWB tag and anchor can communicate, distance measurements between UWB anchors and tags are collected. Subsequently, the pose of the EV is initialized through global optimization of the UWB measurements. In the second step, based on the initialized vehicle pose, wheel odometry and UWB distance data collected as the vehicle moves are fused to estimate the pose of the vehicle in real time. The details are given in the following subsections.

Vehicle pose initialization by globally optimizing UWB measurements

The purpose of this step is to quickly find the approximate initial pose of the EV using only UWB distance measurements under the assumption that no prior information about the EV's pose is available. From the mathematical analysis of the proposed DADT UWB system, the EV pose can always be uniquely determined by the DADT UWB system in the area of $y > d_0$. Therefore, it is possible to predict the sensor measurement value from the measurement model as formulated in Eq. (3) and determine the initial pose of the vehicle through global optimization between the predicted value and the actual measurement value.

We propose a particle swarm optimization (PSO) (Kennedy, Eberhart & Shi, 2001)-based global optimization algorithm that can quickly search for suboptimal solutions to find the initial pose. Each particle in PSO is considered a potential solution, i.e., a vehicle pose state vector, and searches for a given solution space. The position of each particle is iteratively updated based on the experience of the particle and its neighbors and converges toward the optimal solution quickly.

The PSO-based vehicle pose initialization method is as follows. When a vehicle approaches a parking area where communication between UWB anchors and tags is possible, UWB distance measurements between each UWB tag and anchor pair are sampled. When a certain number of measurements is collected, the average value is estimated by removing outliers. The error function J for global optimization is defined as follows.

$$J = [h(\mathbf{x}_t, a_0, d_0) - \mathbf{z}_k]^T [h(\mathbf{x}_t, a_0, d_0) - \mathbf{z}_k], \quad (15)$$

where \mathbf{z}_k is a measurement vector. For the region where $y > d_0$, the PSO finds the initial vehicle pose $\mathbf{x}_0 = [x_0, y_0, \theta_0]^T$ whose error function J value is less than or equal to a specific threshold ϵ . The threshold is determined by considering the variance of UWB measurements.

Vehicle pose tracking by fusing odometry and UWB measurements

To estimate the vehicle's pose precisely as the vehicle moves, the UWB distance measurements and wheel odometry data are fused through a particle filter. The method of estimating the vehicle's pose through the particle filter is shown in the right block of Fig. 4, and the details of each part are as follows.

In particle filter-based localization, a group of particles represents the probability distribution of vehicle states, with each particle $\mathbf{x}_k^{[m]}$ representing a possible state, where $[m]$ indicates a particle index. When the initial pose state of the EV is determined by the PSO, the particles are initialized to have a Gaussian distribution. The mean of the distribution is set to the initial pose state determined by the PSO.

The motion model of the vehicle is

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k) + \mathbf{e}_u, \quad (16)$$

where \mathbf{u}_k is a control input vector and \mathbf{e}_u is normally distributed process noise with zero mean. Given the current vehicle pose \mathbf{x}_k and the positions of the anchors fixed at $\mathbf{a}_1 = [a_0, 0]^T$ and $\mathbf{a}_2 = [-a_0, 0]^T$, the observation model of the DADT UWB system can be

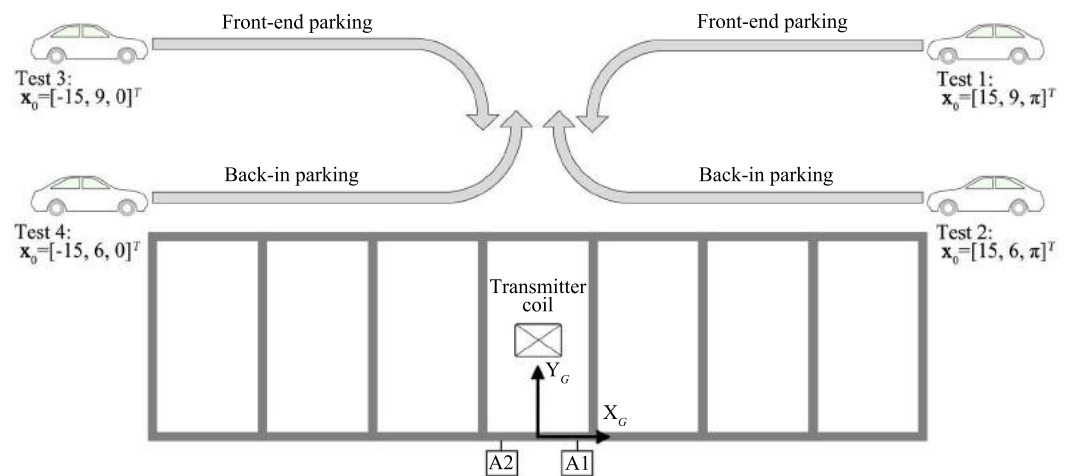


Figure 5 Four simulation tests. The vehicle starts from four different positions and moves along the path for front-end parking or back-in parking.

Full-size [DOI: 10.7717/peerjcs.567/fig-5](https://doi.org/10.7717/peerjcs.567/fig-5)

written as

$$\mathbf{z}_k = h(\mathbf{x}_k, a_0, d_0) + \mathbf{e}_v \quad (17)$$

where \mathbf{e}_v is normally distributed Gaussian noise with zero mean.

The vehicle pose $\mathbf{x}_k^{[m]}$ is predicted by taking the wheel odometry into consideration, which is denoted by $\mathbf{x}_k^{[m]} \sim p(\mathbf{x}_k | \mathbf{x}_{1:k-1}^{[m]}, \mathbf{u}_k)$. The importance weight is computed by (Grisetti, Stachniss & Burgard, 2007)

$$\omega_k^{[m]} = \frac{1}{\sqrt{2\pi}\mathbf{Q}} \exp \left[-\frac{1}{2} \left(\mathbf{z}_k - \hat{\mathbf{z}}_k^{[m]} \right)^T (\mathbf{Q}_k)^{-1} \left(\mathbf{z}_k - \hat{\mathbf{z}}_k^{[m]} \right) \right], \quad (18)$$

where $\hat{\mathbf{z}}_k^{[m]}$ is a predicted measurement and \mathbf{Q}_k is the covariance of the anchors' positions. In the process of particle filtering, the importance weight of some particles can gradually become low, and particles with lower importance weights have little effect in estimating the vehicle pose states. To prevent this effect, the particles are resampled in proportion to the weight of each particle. The number of effective particles given by

$$N_{eff} = \frac{1}{\sum_{m=1}^M \left(\omega_k^{[m]} \right)^2}, \quad (19)$$

where M is the total number of particles. When the number of effective particle is less than 50% of the total number of particles, the weights of all particles are uniformly reset after resampling particles.

SIMULATION AND EXPERIMENTAL RESULTS

Simulation results

To verify the performance of the proposed DADT UWB localization method, we perform the following simulations: (1) Initialize the vehicle's pose by globally optimizing the error

Table 1 Initial pose estimation results by the Levenberg–Marquardt and the proposed globally optimizing DADT UWB measurements for the four selected poses (unit: m, rad).

	True pose $\mathbf{x}_{GT} = [x, y, \theta]^T$	Levenberg–Marquardt		Proposed DADT	
		Estimated Pose	\mathcal{E}	Estimated Pose	\mathcal{E}
Test 1	$[15, 9, \pi]^T$	$[14.83, 9.26, 1.11]^T$	2.05	$[14.99, 9.01, 3.14]^T$	1.42E–02
Test 2	$[15, 6, \pi]^T$	$[14.88, 6.30, 0.79]^T$	2.37	$[15.01, 5.98, 3.14]^T$	2.24E–02
Test 3	$[-15, 9, 0]^T$	$[-14.74, 9.40, -0.03]^T$	0.48	$[-14.97, 9.05, 0.00]^T$	5.83E–02
Test 4	$[-15, 6, 0]^T$	$[-14.69, 6.69, -0.04]^T$	0.76	$[-14.99, 6.02, 0.00]^T$	2.24E–02

function defined in Eq. (15); (2) estimate the vehicle’s pose by fusing wheel odometry and UWB measurements. Figure 5 shows the simulation scenario. The vehicle starts from four different positions and moves along the path for front-end parking or back-in parking. Considering the actual dimensions of the vehicle and the parking space environment, the distance between the UWB tags mounted on the vehicle is set to 1.36 m, i.e., $d_0 = 0.68$ and the two UWB anchors are placed at $\mathbf{a}_1 = [1, 0]^T$ and $\mathbf{a}_2 = [-1, 0]^T$. The wheel odometry noise and UWB measurement noises are assumed to be sampled from the normal distributions denoted by $\mathbf{e}_u \sim N(0, \sigma_u^2)$ and $\mathbf{v}_u \sim N(0, \sigma_v^2)$ where $\sigma_u = 0.35$ m/s and $\sigma_v = 0.10$ m.

Table 1 shows the initial pose estimation results by the proposed global optimizing UWB measurements for the four selected poses. To show the effectiveness of the proposed vehicle’s pose initialization method, the results of the Levenberg–Marquardt method (Moré, 1978), which is a widely used method of optimization of the least square problem, are compared with the results of the proposed DADT method. The pose error \mathcal{E} between the estimated initial pose \mathbf{x}_0 and the ground truth pose \mathbf{x}_{GT} defined by

$$\mathcal{E} = \|\mathbf{x}_0 - \mathbf{x}_{GT}\|. \quad (20)$$

Since the Levenberg–Marquardt method is a local optimization method, it has a limitation in that it cannot find the initial position when it converges to a local minimum. As seen in Table 1, the Levenberg–Marquardt method fails to find the vehicle’s initial pose. However, the proposed initial pose estimation method can precisely find the initial pose for all the tests. These results are consistent with the mathematical analysis of the DADT UWB localization system.

Figures 6–9 show the results of estimating the vehicle’s pose continuously from the initial pose through the fusion of the vehicle’s odometry data and the UWB distance measurements under the particle filter framework with a fixed number of particles, $M = 100$. As seen from the results, the error of the odometry increases as the vehicle moves, whereas the trajectories estimated by the proposed DADT method match the ground truth trajectories in all four cases. Figure 10 shows boxplots for each test, including the mean, minimum, maximum, and standard deviation of the errors. Table 2 shows the numerical values corresponding to the boxplots. The results show that the proposed DADT method keeps the mean error of the vehicle’s position under 0.1 m.

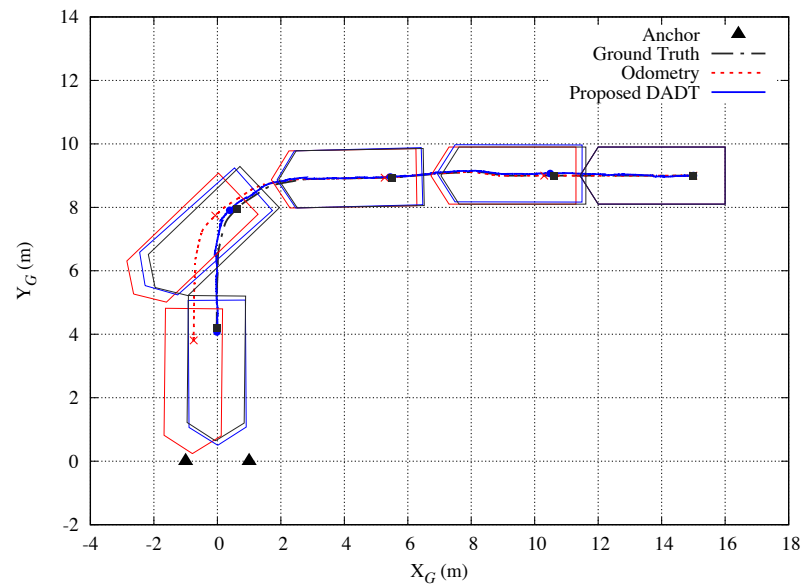


Figure 6 Simulation results of Test 1. The vehicle starts from its initial pose $\mathbf{x}_0 = [15, 9, \pi]^T$ and moves along the path for front-end parking.

Full-size DOI: 10.7717/peerjcs.567/fig-6

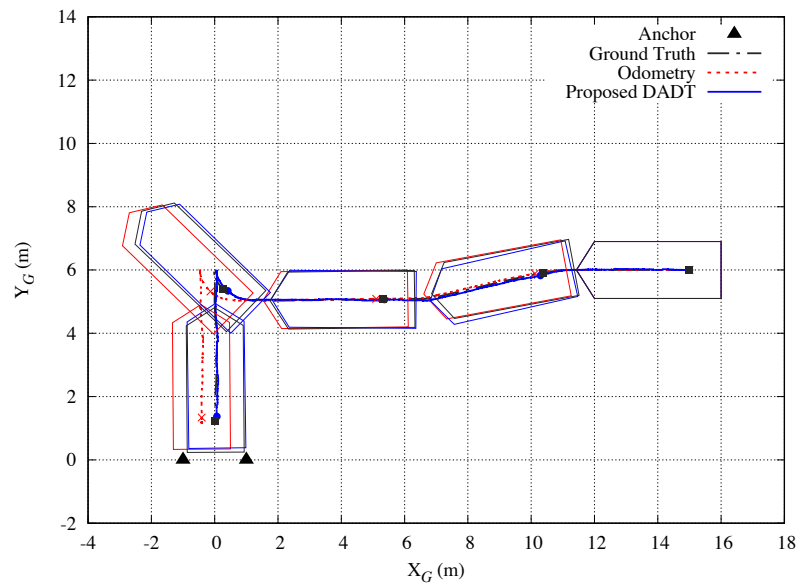


Figure 7 Simulation results of Test 2. The vehicle starts from its initial pose $\mathbf{x}_0 = [15, 6, \pi]^T$ and moves along the path for back-in parking.

Full-size DOI: 10.7717/peerjcs.567/fig-7

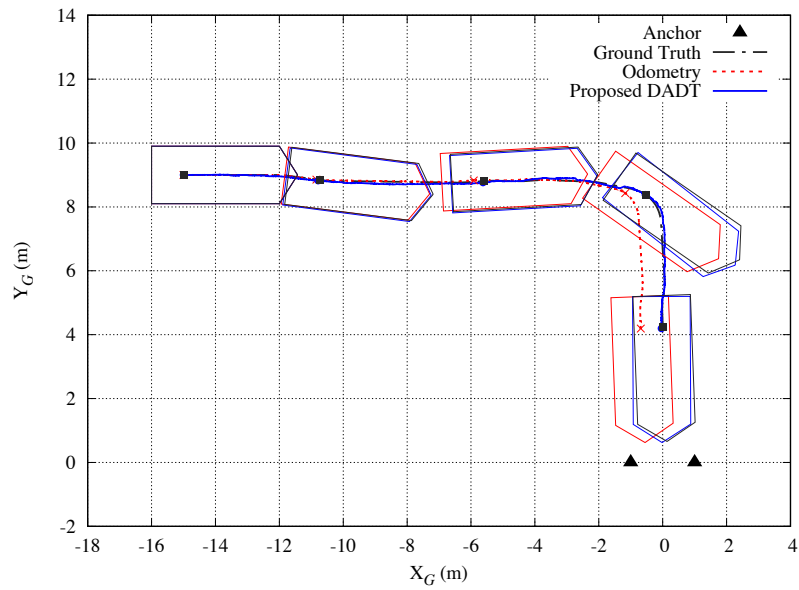


Figure 8 Simulation results of Test 3. The vehicle starts from its initial pose $\mathbf{x}_0 = [-15, 9, 0]^T$ and moves along the path for front-end parking.

[Full-size](#) DOI: 10.7717/peerjcs.567/fig-8

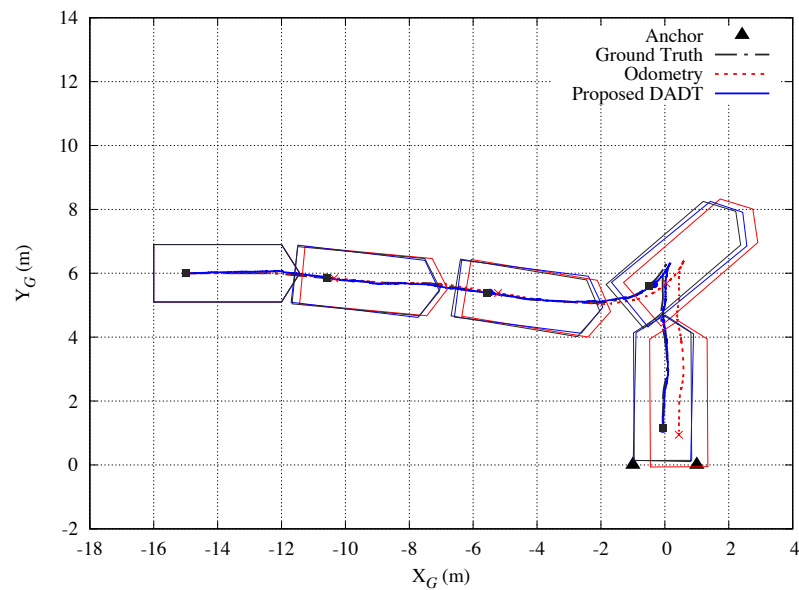


Figure 9 Simulation results of Test 4. The vehicle starts from its initial pose $\mathbf{x}_0 = [-15, 6, 0]^T$ and moves along the path for back-in parking.

[Full-size](#) DOI: 10.7717/peerjcs.567/fig-9

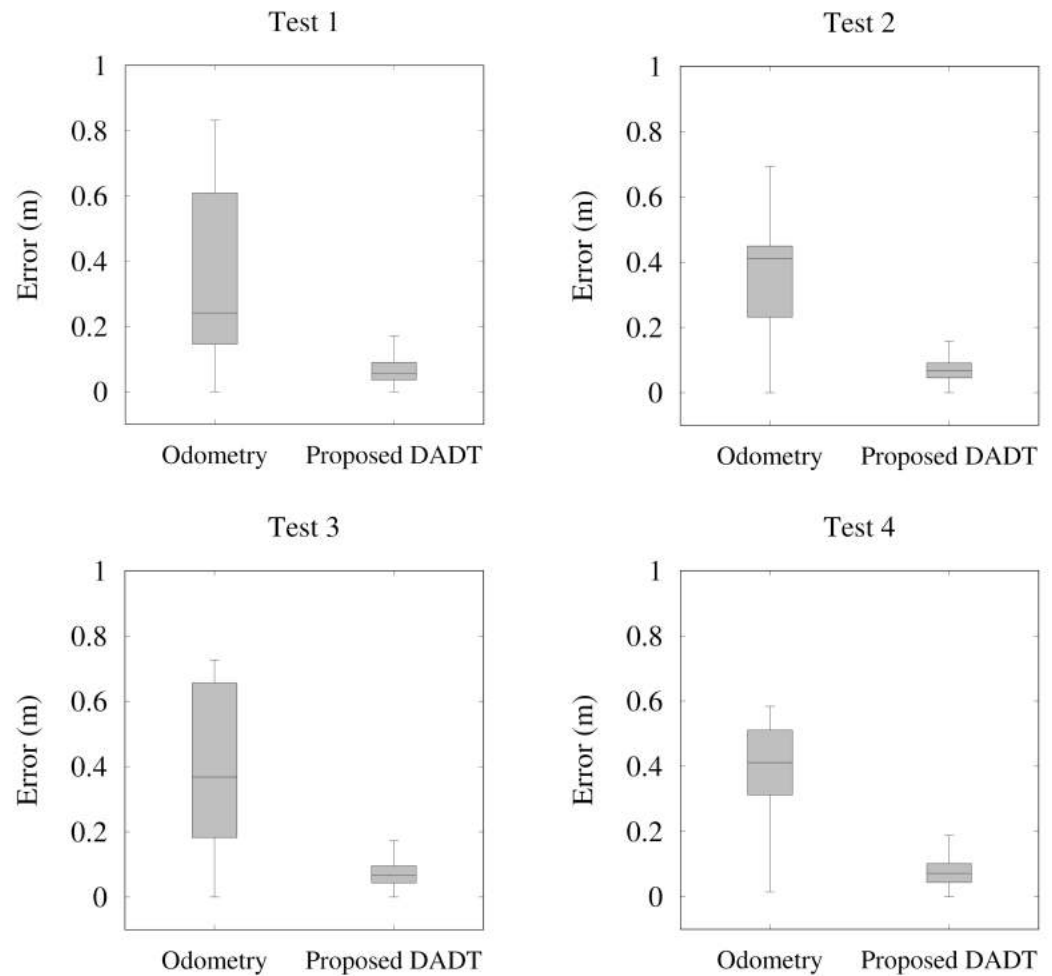


Figure 10 Boxplots of the simulation results.

[Full-size !\[\]\(4729e517bc6a7cd81c8025b9646574fb_img.jpg\) DOI: 10.7717/peerjcs.567/fig-10](https://doi.org/10.7717/peerjcs.567/fig-10)

Table 2 Comparison of pose estimation error by odometry and the proposed DADT UWB system while the vehicle moves. (Unit: m).

	Odometry				Proposed DADT			
	Mean	Min	Max	Std	Mean	Min	Max	Std
Test 1	0.3396	0.0012	0.8322	0.2613	0.0691	0.0016	0.2831	0.0476
Test 2	0.3555	0.0140	0.6925	0.1414	0.0720	0.0129	0.2220	0.0397
Test 3	0.3930	0.0034	0.7266	0.2362	0.0714	0.0035	0.2075	0.0382
Test 4	0.3831	0.0032	0.5838	0.1525	0.0732	0.0030	0.2344	0.0401

Experimental results

The proposed DADT UWB-based localization method is tested with a real vehicle. The tests are performed with the UWB modules manufactured by Pozyx ([Pozyx, 2021](#)), which have a maximum measurable distance of 30 m and an update rate of 60 Hz. As shown

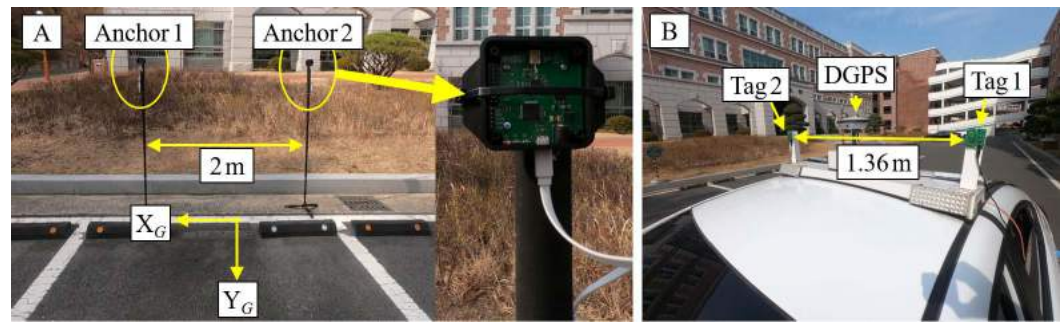


Figure 11 Experimental setup. (A) Two UWB anchors are installed near both corners of the parking slot. (B) two UWB tags and DGPS are mounted on the vehicle roof.

Full-size DOI: [10.7717/peerjcs.567/fig-11](https://doi.org/10.7717/peerjcs.567/fig-11)

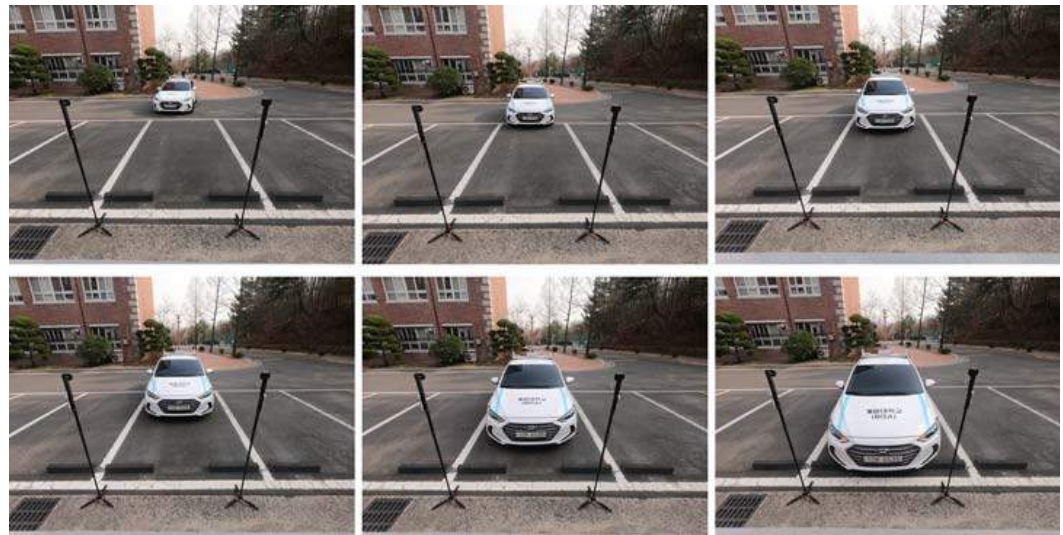


Figure 12 Snapshots of the experiment with a real vehicle. The vehicle is performing front-end parking.

Full-size DOI: [10.7717/peerjcs.567/fig-12](https://doi.org/10.7717/peerjcs.567/fig-12)

in Fig. 11, two UWB tags are mounted on the vehicle roof, and two UWB anchors are installed near both corners of the parking slot. The UWB anchors are installed at a height of 1.8 m to maintain line-of-sight communication with the UWB tags mounted on the vehicle. The positions where the UWB anchors and tags are installed are set to be the same as in the simulation. To evaluate the error of the proposed method, a differential global positioning system (DGPS) receiver module with centimeter-level accuracy is also mounted on the vehicle. The proposed localization algorithm is implemented to perform real-time computation at 10 Hz on an NVIDIA Nano Jetson board (NVIDIA, 2021). Figure 12 shows snapshots of the experiment with a real vehicle performing front-end parking.

Figures 13–14 show the experimental results with the actual vehicle. As shown in the experimental results, the error generated by the wheel odometry increases due to the slip of the wheels when moving along a curved path. However, the proposed DADT UWB-based method precisely corrects the vehicle's pose using the UWB distance measurements. Table 3

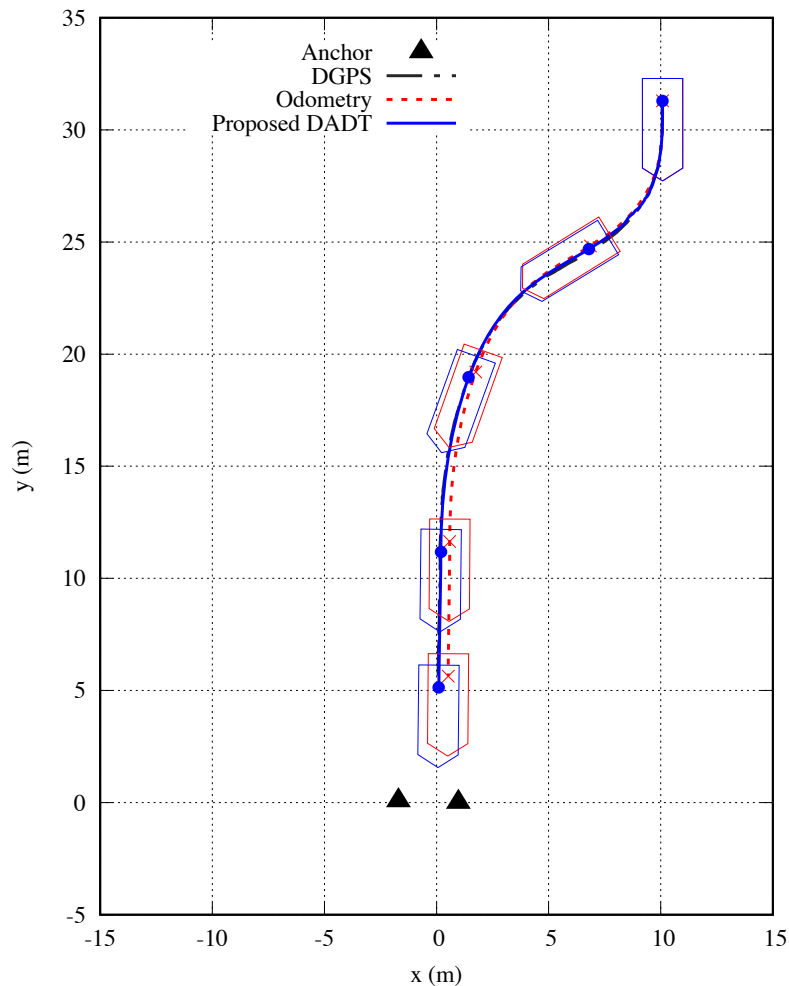


Figure 13 Results of Exp. 1. The black dashed-dotted line shows the DGPS trajectory, the red dotted line shows the odometry trajectory, and the blue solid line shows the proposed DADT UWB-based localization results.

[Full-size](#) DOI: 10.7717/peerjcs.567/fig-13

Table 3 Comparison of the final position estimation error by odometry and the proposed DADT UWB system with a real vehicle. (unit: m).

		X_G -axis error	Y_G -axis error	Distance error
Exp. 1	Odometry	0.4690	0.5826	0.7479
	Proposed DADT	0.0493	0.0763	0.0908
Exp. 2	Odometry	0.9038	0.1604	0.9179
	Proposed DADT	0.0328	0.0031	0.0329

shows that the errors are within 0.1 m at the final parked position. The proposed DADT UWB-based localization method can be sufficiently applied to WPT.

Through the experiments, the average computation time required to update the vehicle's pose at each time instant is estimated while increasing the number of particles from 20 to

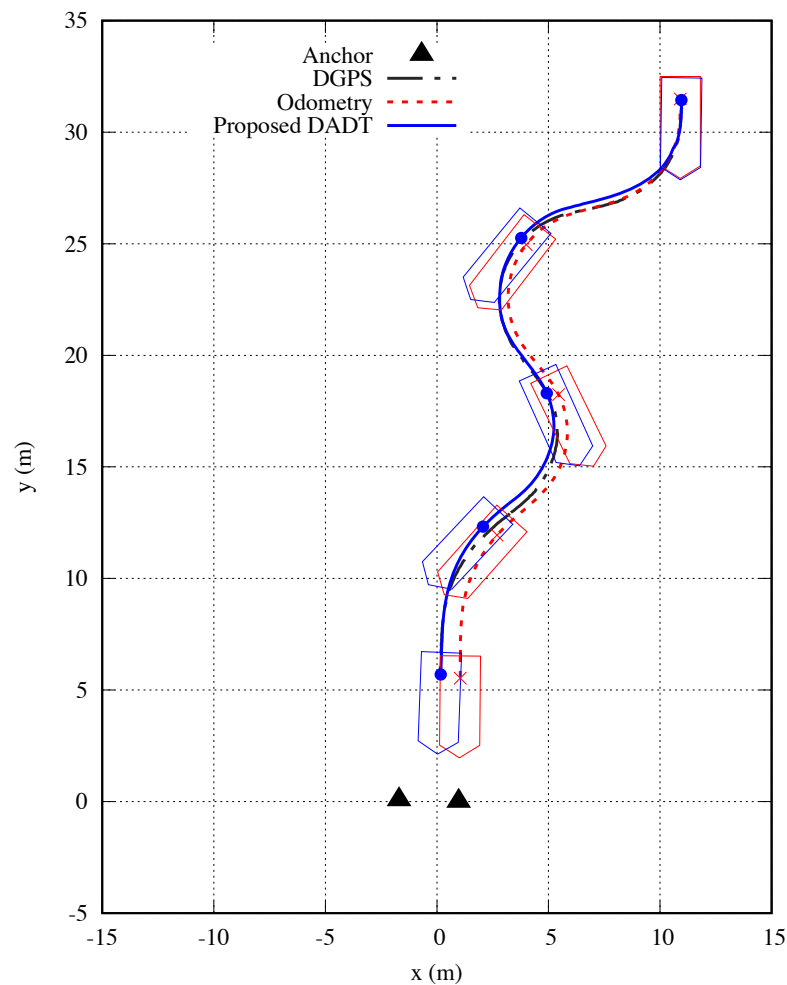


Figure 14 Results of Exp. 2. The black dashed-dotted line shows the DGPS trajectory, the red dotted line shows the odometry trajectory, and the blue solid line shows the proposed DADT UWB-based localization results.

[Full-size](#) [DOI: 10.7717/peerjcs.567/fig-14](https://doi.org/10.7717/peerjcs.567/fig-14)

100. Figure 15 shows the average computation time estimated by an NVIDIA Nano Jetson board. The computation time grows linearly as the number of particles increases. However, even when 100 particles are used, it can be operated in real time at a rate of 10 Hz.

CONCLUSIONS

This paper proposed a novel short-range precise localization method using a DADT UWB sensor system for application to a WPT system. An observability analysis of the proposed DADT UWB sensor system consisting of two anchors and two tags was performed based on the FIM. The proposed localization algorithm determines the vehicle's initial pose by globally optimizing the collected UWB distance measurements and estimates the vehicle's pose by fusing the vehicle's wheel odometry data and the UWB distance measurements.

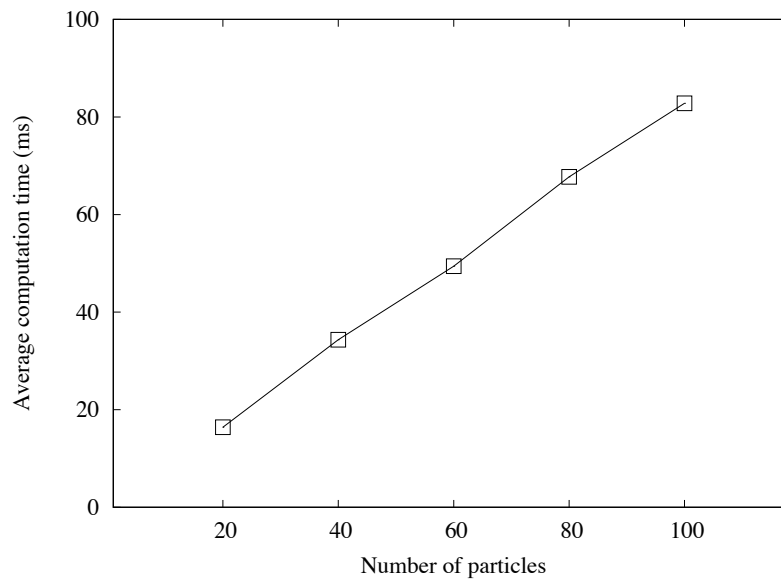


Figure 15 Average computation time required to update the vehicle's pose at each time instant.

Full-size DOI: [10.7717/peerjcs.567/fig-15](https://doi.org/10.7717/peerjcs.567/fig-15)

The effectiveness of the proposed method was confirmed through various simulations and real vehicle-based experiments.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2018R1C1B5038763). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the author:
Korea government (MSIT): NRF-2018R1C1B5038763.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Seung-Mok Lee conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:
The source code is available in the [Supplementary File](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.567#supplemental-information>.

REFERENCES

- Alarifi A, Al-Salman A, Alsaleh M, Alnafessah A, Al-Hadhrami S, Al-Amman M, Al-Khalifa H. 2016. Ultra wideband indoor positioning technologies analysis and recent advances. *Sensors* 16:707 DOI 10.3390/s16050707.
- Chen Y-Y, Huang S-P, Wu T-W, Tsai W-T, Liou C-Y, Mao S-G. 2020. UWB system for indoor positioning and tracking with arbitrary target orientation, optimal anchor location, and adaptive NLOS mitigation. *IEEE Transactions on Vehicular Technology* 69(9):9304–9314.
- El-Shahat A, Ayisire E, Wu Y, Rahman M, Nelms D. 2019. Electric vehicles wireless power transfer state-of-the-art. *Energy Procedia* 162:24–27 DOI 10.1016/j.egypro.2019.04.004.
- Gao R, He F, Li T. 2018. VeLoc: Finding your car in indoor parking structures. *Sensors* 18:1403 DOI 10.3390/s18051403.
- González-Castaño FJ, Gil-Castiñeira F, Rodríguez-Pereira D, Regueiro-Janeiro JA, Garca-Méndez S. 2021. Self-corrective sensor fusion for drone positioning in indoor facilities. *IEEE Access* 9:2415–2427 DOI 10.1109/ACCESS.2020.3048194.
- Grisetti G, Stachniss C, Burgard W. 2007. Improved techniques for grid mapping with rao-blackwellized particle filters. *IEEE Transactions on Robotics* 23(1):34–46 DOI 10.1109/TRO.2006.889486.
- Hu J, Yang M, Xu H, He Y, Wang C. 2019. Mapping and localization using semantic road marking with centimeter-level accuracy in indoor parking lots. In: *Proceedings of the 2019 IEEE intelligent transportation systems conference (ITSC)*. Auckland, New Zealand.
- Hyun J, Oh T, Lim H, Myung H. 2019. UWB-based indoor localization using ray-tracing algorithm. In: *Proceedings of the 2019 16th international conference on ubiquitous robots (UR)*. Jeju, Korea.
- Kennedy J, Eberhart R, Shi Y. 2001. *Swarm intelligence*. 1st edition. Burlington: Morgan Kaufmann Publishers.
- Knobloch D. 2017. Practical challenges of particle filter based UWB localization in vehicular environments. In: *Proceedings of the 2017 international conference on indoor positioning and indoor navigation (IPIN)*. Sapporo, Japan.
- Kukolev P, Chandra A, Mikulášek T, Prokeš A. 2016. Out-of-vehicle time-of-arrival-based localization in ultra-wide band. *International Journal of Distributed Sensor Networks* 12(8):1–11.
- Lee S-M. 2020. Ultra wideband-based precise localization of autonomous vehicle in GPS-denied environments. In: *Proceedings of the 8th international conference on robot intelligence technology and applications (RiTA)*. Cardiff, UK.

- Lee S-M, Jung J, Kim S, Kim I-J, Myung H. 2015. DV-SLAM (Dual-sensor-based vector-field SLAM) and observability analysis. *IEEE Transactions on Industrial Electronics* 62(2):1101–1112 DOI 10.1109/TIE.2014.2341595.
- Liang C, Yang G, Yuan F, Huang X, Sun Y, Li J, Song K. 2020. Modeling and analysis of thermal characteristics of magnetic coupler for wireless electric vehicle charging system. *IEEE Access* 8:173177–173185 DOI 10.1109/ACCESS.2020.3025586.
- Machura P, Santis VD, Li Q. 2020. Driving range of electric vehicles charged by wireless power transfer. *IEEE Transactions on Vehicular Technology* 69(6):5968–5982 DOI 10.1109/TVT.2020.2984386.
- Moré JJ. 1978. *The Levenberg-Marquardt algorithm: implementation and theory. Lecture notes in mathematics*, vol. 630. Berlin, Heidelberg: Springer, 105–116.
- NVIDIA. 2021. Jetson Nano Developer Kit. Available at <https://developer.nvidia.com/embedded/jetson-nano-developer-kit>.
- Panchal C, Stegen S, Lu J. 2018. Review of static and dynamic wireless electric vehicle charging system. *Engineering Science and Technology, an International Journal* 21(5):922–937 DOI 10.1016/j.jestch.2018.06.015.
- Panev S, Vicente F, Torre F. DI, Prinnet V. 2019. Road curb detection and localization with monocular forward-view vehicle camera. *IEEE Transactions on Intelligent Transportation Systems* 20(9):3568–3584 DOI 10.1109/TITS.2018.2878652.
- Pozyx. 2021. Pozyx UWB. Available at <https://pozyx.io> (accessed on 12 April 2021).
- Rozman M, Ikpehai A, Adebisi B, Rabie KM, Gacanin H, Ji H, Fernando M. 2019. Smart wireless power transmission system for autonomous EV charging. *IEEE Access* 7:112240–112248 DOI 10.1109/ACCESS.2019.2912931.
- Shi D, Mi H, Collins EG, Wu J. 2020. An Indoor low-cost and high-accuracy localization approach for AGVs. *IEEE Access* 8:50085–50090 DOI 10.1109/ACCESS.2020.2980364.
- Shin Y, Hwang K, Park J, Kim D, Ahn S. 2019. Precise vehicle location detection method using a wireless power transfer (WPT) system. *IEEE Transactions on Vehicular Technology* 68(2):1167–1177 DOI 10.1109/TVT.2018.2885942.
- Stoll H, Zimmer P, Hartmann F, Sax E. 2017. GPS-independent localization for off-road vehicles using ultra-wideband (UWB). In: *Proceedings of the 2017 IEEE 20th international conference on intelligent transportation systems (ITSC)*. Yokohama, Japan.
- Tao Z, Xue J, Wang D, Zhang S, Cui D, Du S. 2018. Accurate localization in underground garages via cylinder feature based map matching. In: *Proceedings of the 2018 IEEE intelligent vehicles symposium (IV)*. Changshu, Suzhou, China.
- Tian Y, Zhu Z, Xiang L, Tian J. 2020. Vision-based rapid power control for a dynamic wireless power transfer system of electric vehicles. *IEEE Access* 8:78764–78778 DOI 10.1109/ACCESS.2020.2989466.
- Tiemann J, Pillmann J, Böcker S, Wietfeld C. 2016. Ultra-wideband aided precision parking for wireless power transfer to electric vehicles in real life scenarios. In: *Proceedings of the 2016 IEEE 84th vehicular technology conference (VTC-Fall)*. Montreal, Canada.

- Wang W, Marelli D, Fu M. 2021.** Multiple-vehicle localization using maximum likelihood Kalman filtering and ultra-wideband signals. *IEEE Sensors Journal* 21(4):4949–4956 DOI [10.1109/JSEN.2020.3031377](https://doi.org/10.1109/JSEN.2020.3031377).
- Yu Z, Gao Z, Chen H, Huang Y. 2020.** SPFCN: select and prune the fully convolutional networks for real-time parking slot detection. In: *Proceedings of the 2020 IEEE intelligent vehicles symposium (IV)*. Las Vegas, USA.
- Zhang R, Song L, Jaiprakash A, Talty T, Alanazi A, Alghafis A, Biyabani AA, Tonguz O. 2019.** Using ultra-wideband technology in vehicles for infrastructure-free localization. In: *Proceedings of the 2019 IEEE 5th world forum on internet of things (WF-IoT)*. Piscataway: IEEE.
- Zhao H, Yang M, He Y, Wang C. 2020.** Time-of-flight camera based indoor parking localization leveraging Manhattan world regulation. In: *Proceedings of the 2020 IEEE intelligent vehicles symposium (IV)*. Piscataway: IEEE.

Neural network assisted Kalman filter for INS/UWB integrated seamless quadrotor localization

Shuhui Bi, Liyao Ma, Tao Shen, Yuan Xu and Fukun Li

School of Electrical Engineering, University of Jinan, Jinan, Shandong, China

ABSTRACT

Due to some harsh indoor environments, the signal of the ultra wide band (UWB) may be lost, which makes the data fusion filter can not work. For overcoming this problem, the neural network (NN) assisted Kalman filter (KF) for fusing the UWB and the inertial navigation system (INS) data seamlessly is present in this work. In this approach, when the UWB data is available, both the UWB and the INS are able to provide the position information of the quadrotor, and thus, the KF is used to provide the localization information by the fusion of position difference between the INS and the UWB, meanwhile, the KF can provide the estimation of the INS position error, which is able to assist the NN to build the mapping between the state vector and the measurement vector off-line. The NN can estimate the KF's measurement when the UWB data is unavailable. For confirming the effectiveness of the proposed method, one real test has been done. The test's results demonstrate that the proposed NN assisted KF is effective to the fusion of INS and UWB data seamlessly, which shows obvious improvement of localization accuracy. Compared with the LS-SVM assisted KF, the proposed NN assisted KF is able to reduce the localization error by about 54.34%.

Subjects Adaptive and Self-Organizing Systems, Data Science, Robotics

Keywords Neural network assisted Kalman filter, INS/UWB, Quadrotor, Localization

INTRODUCTION

Nowadays, the quadrotor has been widely used in many fields (*Xu et al., 2020a; Nguyen & Hong, 2019; Kou et al., 2018*). Consequently, many approaches have been proposed for the quadrotor (*Liang et al., 2019*). In order to make the quadrotor have better performance, the accurate localization scheme, which is the key technology of the quadrotor to accomplish other tasks, should be investigated (*Camci & Kayacan, 2019*).

To the localization technologies for the quadrotor, there are many approaches have been proposed. For instance, a smart quadcopter aircraft navigation system using the global positioning system (GPS) was designed, which can achieve autonomous flight control with smooth and stable maneuvering, see *Bonny & Abdelsalam (2019)*. Global navigation satellite systems (GNSS) integrating light detection and ranging (LiDAR) scheme was investigated to achieve the autonomous navigation in forests (*Chiella et al., 2019*). The indoor quadrotor localization integrated by inertial navigation system (INS) and ultra wide band (UWB) was proposed by *Xu et al. (2020b)*. A high-speed autonomous

Submitted 26 March 2021

Accepted 17 June 2021

Published 14 July 2021

Corresponding author

Yuan Xu, xy_abric@126.com

Academic editor

Pengcheng Liu

Additional Information and
Declarations can be found on
page 11

DOI 10.7717/peerj-cs.630

© Copyright

2021 Bi et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

quadrotor navigation through visual and inertial paths was proposed ([Do, Carrillo-Arce & Roumeliotis, 2019](#)). Autonomous vision-based micro air vehicle for indoor and outdoor navigation was investigated in [Schmid et al. \(2014\)](#). It should be emphasized that the basic idea of the approaches mentioned above is to replace the unavailable positioning technology with a available one.

In aggregate, the data fusion filter has played an important role in integrated navigation system ([Zhao & Huang, 2020](#); [Wang et al., 2018](#); [Li et al., 2019](#); [Liu, Yu & Shuang, 2019](#)). Moreover, the Kalman filter (KF) with its improving filters have been proposed for the data fusion ([Liu et al., 2020](#)). For example, the fading cubature Kalman filter (CKF) was designed to the initial alignment of strapdown inertial navigation system (SINS) ([Guo et al., 2020](#)). The quadrotor state estimation based on CKF was proposed ([Benzerrouk, Nebylov & Salhi, 2016](#)). An improving CKF method was investigated for the attitude determination system of missile ([Liu et al., 2019](#)). The CKF is used for the GNSS/INS under GNSS-challenged environment ([Cui et al., 2019](#)). An improved square root unscented Kalman filter was proposed for the localization of the coaxial Quadrotor ([Gośliński et al., 2019](#)). A Kalman filter/expectation maximization (EM) integrated frame was proposed in [Qin et al. \(2020\)](#). A new approach for enhancing the indoor navigation of unmanned aerial vehicles (UAVs) with velocity update applied to an extended Kalman filter (EKF) was investigated by [Zahran et al. \(2019\)](#). It should be pointed out that the outage of the data fusion filter's measurement are not considered by the approaches mentioned above. Meanwhile, in order to ensure that the data fusion filter works, some artificial intelligence (AI)-based methods have been proposed, which have been used in other fields ([Zhang et al., 2021, 2020](#)).

In this paper, we propose a neural network (NN) assisted KF, which is able to deal with the missing data in case of UWB data outage. Neural network is used to build the mapping between states and observations. The performance is verified with real data. Comparison shows that the proposed approach outperforms LS-SVM algorithm significantly in accuracy improvement.

The contributions of this work are listed in the following:

- A new NN assisted KF for fusing the UWB and INS data seamlessly is presented in this work, which employs the NN to build mapping between states and observations offline and predict the observations when the UWB is outage.
- Real tests have been done for demonstrating the effectiveness of the proposed approach.

The remainder structure of this article is sketched as follows. The description of INS/UWB integrated seamless quadrotor localization scheme is given in "INS/UWB Integrated Seamless Quadrotor Localization Scheme". "Kalman Filter" and "The Scheme of the NN" investigated the KF and the NN method for the localization scheme of INS/UWB integrated seamless quadrotor. The test is done in the "Test" section. Finally, conclusions are drawn in the "Conclusion" section.

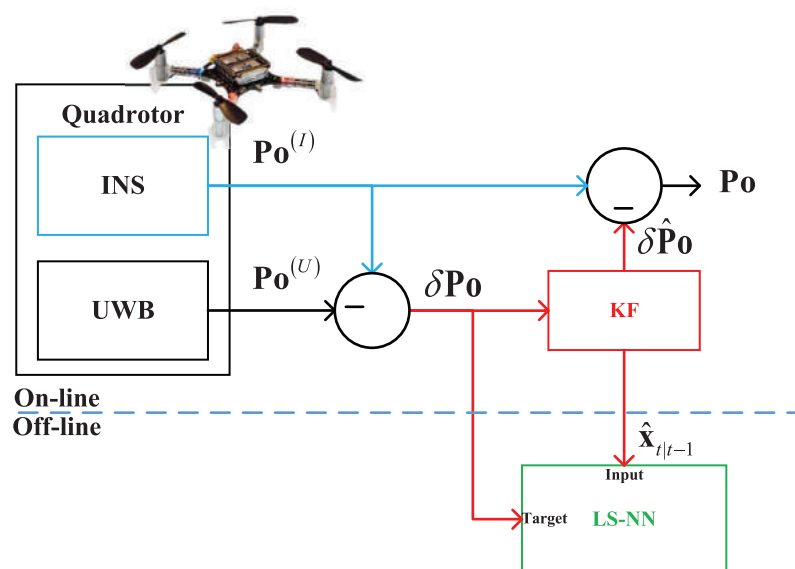


Figure 1 The data fusion scheme when the UWB measurements are available.

Full-size DOI: 10.7717/peerj-cs.630/fig-1

INS/UWB INTEGRATED SEAMLESS QUADROTOR LOCALIZATION SCHEME

In this section, the INS/UWB integrated seamless quadrotor localization scheme will be designed in two cases. The integrated seamless scheme proposed in this work are listed in the following:

- When the UWB measurements are available, the data fusion scheme is shown in Fig. 1. In this situation, the INS and UWB localization technologies measure the target quadrotor's position $\mathbf{P_o}^{(I)}$ and $\mathbf{P_o}^{(U)}$ respectively. Then, the Kalman filter (KF) estimates the position $\mathbf{P_o}$ by fusing the $\mathbf{P_o}^{(I)}$ and $\mathbf{P_o}^{(U)}$.
- Using the outputs and the measurements of the KF when the UWB measurements are available, the NN works in the training stage, it builds the mapping between the KF's measurement $\delta\mathbf{P_o}$, $t \in [1, +\infty)$ and the data filter's state vector $\hat{\mathbf{x}}_{t|t-1}$, $t \in [1, +\infty)$ after normal flight of the quadrotor. Here, the t is the time index. It should be pointed out that both the $\delta\mathbf{P_o}$, $t \in [1, \infty)$ and the $\hat{\mathbf{x}}_{t|t-1}$, $t \in [1, \infty)$ are collected when the KF works normally, and the building process of the mapping is off-line.
- When the UWB measurements are not available, the data fusion scheme can be designed as Fig. 2. In this situation, the UWB is unable to provide the $\mathbf{P_o}^{(U)}$ due to the outage of the UWB. Thus, the KF is unable to work. In this situation, the NN is employed to rebuild the measurement of the KF. It works in prediction stage, which is utilized to provide the estimated position error $\delta\mathbf{P_o}$ by using the mapping built in the above stage and the $\hat{\mathbf{x}}_{t|t-1}$. Then, the $\delta\mathbf{P_o}$ is used as the measurement of the KF, which makes the KF can work when the UWB measurement is outage.

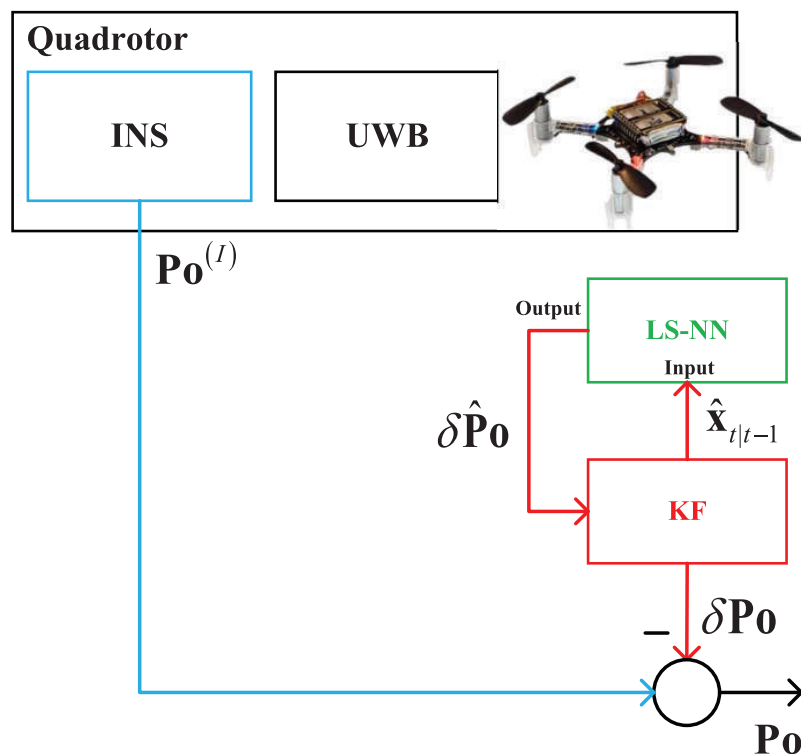


Figure 2 The data fusion scheme when the UWB measurements are unavailable.

Full-size DOI: 10.7717/peerj-cs.630/fig-2

KALMAN FILTER

Based on the seamless integrated scheme, the KF used in this work will be introduced in this section. The state equation of KF used in this work is listed in Eq. (1).

$$\underbrace{\begin{bmatrix} \delta \mathbf{P}_o_{t|t-1} \\ \delta \mathbf{V}_{t|t-1} \end{bmatrix}}_{\mathbf{x}_{t|t-1}} = \underbrace{\begin{bmatrix} \mathbf{I}_{3 \times 3} & \delta t \cdot \mathbf{I}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix}}_{\mathbf{F}} \underbrace{\begin{bmatrix} \delta \mathbf{P}_o_{t-1} \\ \delta \mathbf{V}_{t-1} \end{bmatrix}}_{\mathbf{x}_{t-1}} + \boldsymbol{\omega}_{t-1} \quad (1)$$

where the time index is denoted as t , δt means the sample time, $\delta \mathbf{P}_o_t = [\delta x_t, \delta y_t, \delta z_t]^T$ means the position error vector at the time index t , here, the $(\delta x_t, \delta y_t, \delta z_t)$ means the position error in the east, north, and up direction respectively, $\delta \mathbf{V}_t = [\delta V_{x_t}, \delta V_{y_t}, \delta V_{z_t}]^T$ means the velocity error vector at the time index t , here, the $(\delta V_{x_t}, \delta V_{y_t}, \delta V_{z_t})$ means the velocity error in the east, north, and up direction respectively, $\boldsymbol{\omega}_{t-1} \sim \mathcal{N}(0, \mathbf{Q})$ is the system noise and \mathbf{Q} is its covariance.

The measurement equation of KF used in this work is listed in Eq. (2).

$$\underbrace{\begin{bmatrix} x_t^{(I)} - x_t^{(U)} \\ y_t^{(I)} - y_t^{(U)} \\ z_t^{(I)} - z_t^{(U)} \end{bmatrix}}_{\mathbf{Y}_t} = \underbrace{\begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix}}_{\mathbf{H}} \mathbf{x}_{t|t-1} + v_{t-1}, \quad (2)$$

Algorithm 1 The KF filtering algorithm based on the model (1) and (2).

Data: $\mathbf{Y}_t, \mathbf{Q}, \mathbf{R}$

Result: $\hat{\mathbf{x}}_t, \hat{\mathbf{P}}_t$

```

1 begin
2   for  $t = 1: \infty$  do
3      $\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}\hat{\mathbf{x}}_{t-1}$ ;
4      $\hat{\mathbf{P}}_{t|t-1} = \mathbf{F}\hat{\mathbf{P}}_{t-1}\mathbf{F}^T + \mathbf{Q}$ ;
5      $\mathbf{K}_t = \hat{\mathbf{P}}_{t|t-1}\mathbf{H}^T(\mathbf{H}\hat{\mathbf{P}}_{t|t-1}\mathbf{H}^T + \mathbf{R})^{-1}$ ;
6      $\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t[\mathbf{Y}_t - \mathbf{H}\hat{\mathbf{x}}_{t|t-1}]$ ;
7      $\hat{\mathbf{P}}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{H})\hat{\mathbf{P}}_{t|t-1}$ ;
8   end for
9 end

```

where $(x_t^{(I)}, y_t^{(I)}, z_t^{(I)})$ is the INS-measured position $\mathbf{Po}^{(I)}$ in east, north, and the upside direction, respectively, $(x_t^{(U)}, y_t^{(U)}, z_t^{(U)})$ is the UWB-measured position $\mathbf{Po}^{(U)}$ in east, north, and the up direction respectively, $\mathbf{v}_t \sim \mathcal{N}(0, \mathbf{R})$ is the measurement noise and \mathbf{R} is its covariance. The KF filtering algorithm based on the model (1) and (2) is listed in [Algorithm 1](#).

THE SCHEME OF THE NEURAL NETWORK (NN)

In case of outage in complex indoor environment, due to the lack of UWB measurements, the observation vector in Kalman filter become unavailable. To provide the observation vector for the data fusion filter, the Neural Network (NN) is employed in this work.

However, it should be noticed that it is hard to model mathematically the relation between the measurements of the data fusion filter \mathbf{Y}_t and the state vector $\hat{\mathbf{x}}_{t|t-1}$. For overcoming this issue, the NN is trained to build the mapping between them using the KF's measurement $\mathbf{Y}_t, t \in [1, +\infty)$ and the $\hat{\mathbf{x}}_{t|t-1}, t \in [1, +\infty)$ collected after normal flight of the quadrotor. The input and target of the NN model are chosen as $\hat{\mathbf{x}}_{t|t-1}$ and \mathbf{Y}_t respectively. In this work, we select the simple BP neural network structure without hidden layer. Build the mapping between $\hat{\mathbf{x}}_{t|t-1}$ and \mathbf{Y}_t using the $\delta\mathbf{Po}_t, t \in [1, \infty)$ and the $\hat{\mathbf{x}}_{t|t-1}, t \in [1, \infty)$ via NN.

The NN method is summarised in [Algorithms 2](#) and [3](#). In the [Algorithm 2](#), the KF provides the $\hat{\mathbf{x}}_t$ and the $\hat{\mathbf{P}}_t$ normally. Then, the NN is used to build the mapping between $\hat{\mathbf{x}}_{t|t-1}$ and \mathbf{Y}_t using the $\delta\mathbf{Po}_t, t \in [1, \infty)$ and the $\hat{\mathbf{x}}_{t|t-1}, t \in [1, \infty)$ on off-line model.

In the [Algorithm 3](#), the KF works normally when the $\mathbf{Po}^{(U)}$ is available. Here, the KF is used to provide the estimation of the $\delta\mathbf{Po}$ using the observation vector

$\mathbf{Y}_t = \begin{bmatrix} x_t^{(I)} - x_t^{(U)} & y_t^{(I)} - y_t^{(U)} & z_t^{(I)} - z_t^{(U)} \end{bmatrix}^T$. Once the $\mathbf{Po}^{(U)}$ is unavailable, the proposed NN assisted Kalman filtering algorithm estimate \mathbf{Y}_t using $\hat{\mathbf{x}}_{t|t-1}$ and previously built mapping via NN.

Algorithm 2 NN assisted Kalman filtering algorithm (off-line model).

Data: Y_t, Q, R

Result: \hat{x}_t, \hat{P}_t the mapping between $\hat{X}_{t|t-1}$ and Y_t

```

1 begin
2   for  $t = 1:\infty$  do
3      $\hat{x}_{t|t-1} = F\hat{x}_{t-1}$ ;
4      $\hat{P}_{t|t-1} = F\hat{P}_{t-1}F^T + Q$ ;
5      $K_t = \hat{P}_{t|t-1}H^T(H\hat{P}_{t|t-1}H^T + R)^{-1}$ ;
6      $\hat{x}_t = \hat{x}_{t|t-1} + K_t[Y_t - H\hat{x}_{t|t-1}]$ ;
7      $\hat{P}_t = (I - K_tH)\hat{P}_{t|t-1}$ ;
8   end for
9   Build the mapping between  $\hat{x}_{t|t-1}$  and  $Y_t$  using the  $\delta Po_t, t \in [1, \infty)$  and the  $\hat{x}_{t|t-1}, t \in [1, \infty)$  via NN;
10 end

```

Algorithm 3 NN assisted Kalman filtering algorithm (on-line model).

Data: Y_t, Q, R , the mapping between $\hat{x}_{t|t-1}$ and Y_t

Result: \hat{x}_t, \hat{P}_t

```

1 begin
2   for  $t = 1:\infty$  do
3      $\hat{x}_{t|t-1} = F\hat{x}_{t-1}$ ;
4      $\hat{P}_{t|t-1} = F\hat{P}_{t-1}F^T + Q$ ;
5     if  $Po^{(U)}$  is available then
6        $Y_t = \begin{bmatrix} x_t^{(I)} - x_t^{(U)} \\ y_t^{(I)} - y_t^{(U)} \\ z_t^{(I)} - z_t^{(U)} \end{bmatrix}$ ;
7     else
8       Estimate  $Y_t$  using  $\hat{x}_{t|t-1}$  and previously built the mapping between  $\hat{x}_{t|t-1}$  and  $Y_t$  via NN;
9     end if
10     $K_t = \hat{P}_{t|t-1}H^T(H\hat{P}_{t|t-1}H^T + R)^{-1}$ ;
11     $\hat{x}_t = \hat{x}_{t|t-1} + K_t[Y_t - H\hat{x}_{t|t-1}]$ ;
12     $\hat{P}_t = (I - K_tH)\hat{P}_{t|t-1}$ ;
13  end for
14 end

```

TEST

In order to demonstrate the effectiveness of the proposed method, the real test will be investigated in this section.

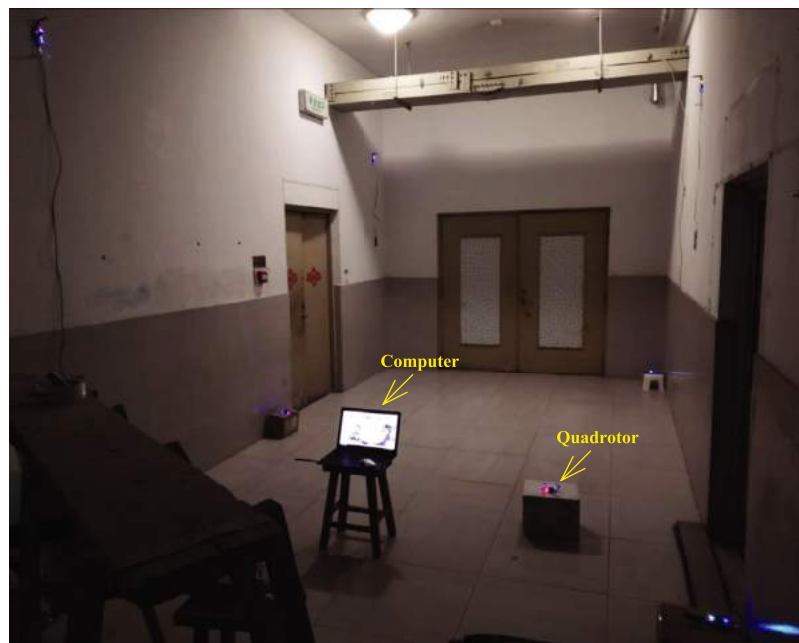


Figure 3 Test environment.

Full-size  DOI: 10.7717/peerj-cs.630/fig-3

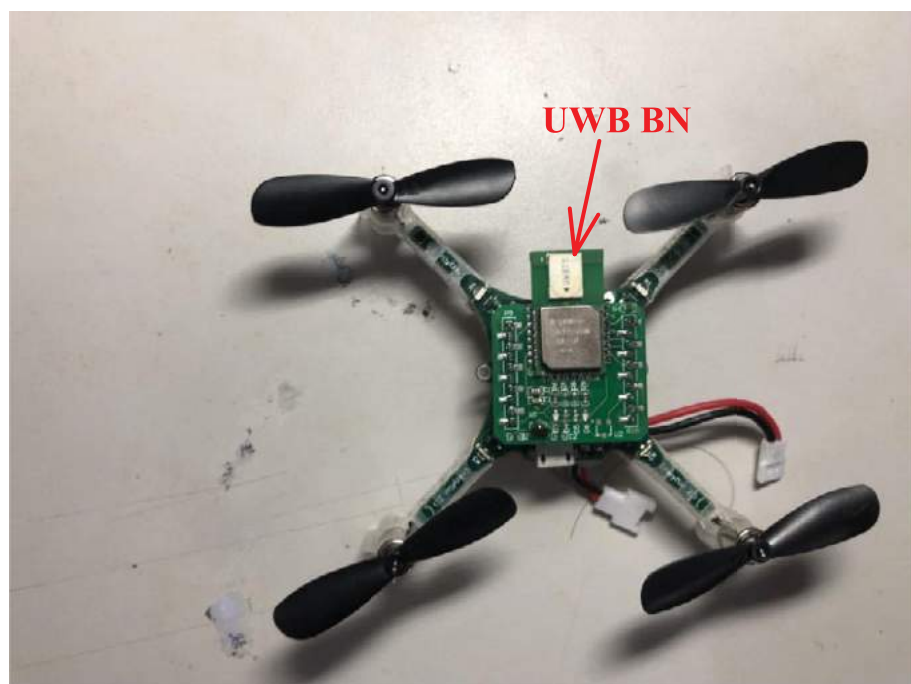


Figure 4 The quadrotor used in this work.

Full-size  DOI: 10.7717/peerj-cs.630/fig-4

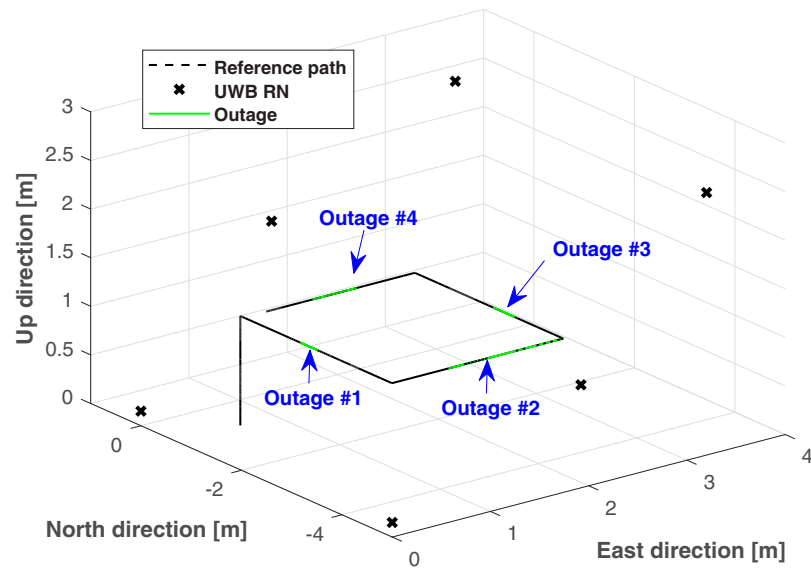


Figure 5 The reference path, UWB RNs, and the outage areas used in the test.

Full-size [DOI: 10.7717/peerj-cs.630/fig-5](https://doi.org/10.7717/peerj-cs.630/fig-5)

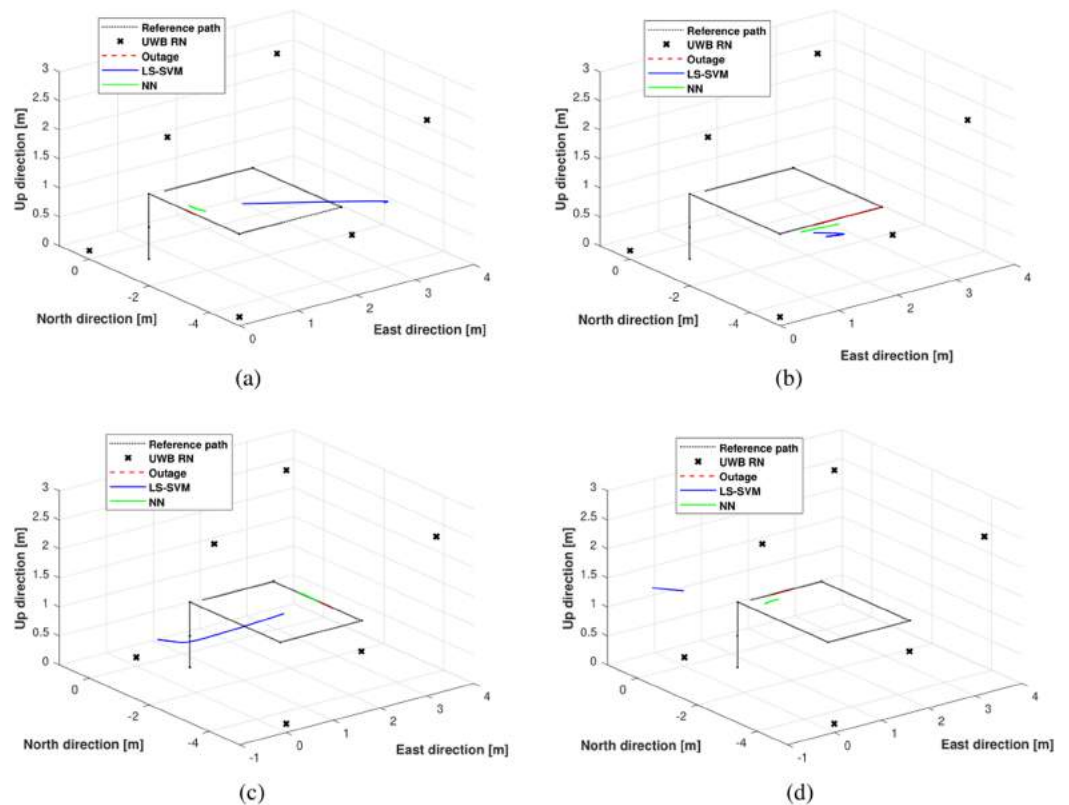


Figure 6 The trajectories estimated by the LS-SVM and the NN in outage areas: (A) outage #1, (B) outage #2, (C) outage #3, and (D) outage #4.

Full-size [DOI: 10.7717/peerj-cs.630/fig-6](https://doi.org/10.7717/peerj-cs.630/fig-6)

Experimental settings

In this section, the real test will be considered to show the validity of the proposed method. The real test is done in the No. 1 building, University of Jinan, China, the test environment is displayed in Fig. 3. The quadrotor used in this work is shown in Fig. 4. Here, we employ the quadrotor to carry UWB blind node (BN) and the inertial measurement unit (IMU). The UWB BN fixed on the target quadrotor is able to collect the distances $d_i, i \in [1, 6]$ between the target quadrotor and the UWB reference node (RN). Here, the i has the same number as the UWB RN. Then, the UWB position $\mathbf{Po}^{(U)}$ can be computed via the $d_i, i \in [1, 6]$. And the INS position $\mathbf{Po}^{(I)}$ is provided by the IMU. The difference $\delta\mathbf{Po}$ between the $\mathbf{Po}^{(I)}$ and $\mathbf{Po}^{(U)}$ is used as the measurement of the KF. In the test, the quadrotor runs following the reference path, which is shown in Fig. 5. In this work, the sample time is set to 0.02s. In order to indicate the effect of the proposed method, four UWB outage areas (#1, #2, #3, and #4) are simulated as shown in Fig. 5.

Localization errors

In this subsection, the performance of the proposed NN assisted KF will be investigated. Here, we compare the NN assisted KF's performance with the least squares support vector machine (LS-SVM) assisted KF. In this work, we employ the mean square error (MSE) at each time index, which is calculated by the follows:

$$\text{MSE}(\mathbf{Po})_t = \frac{1}{3} \left((x_t - x_t^{\text{ref}})^2 + (y_t - y_t^{\text{ref}})^2 + (z_t - z_t^{\text{ref}})^2 \right), \quad (3)$$

where $\text{MSE}(\mathbf{Po})_t$ means the MSE of the position at time index t , (x_t, y_t, z_t) is the estimated position in x , y , and z directions at the time index t , $(x_t^{\text{ref}}, y_t^{\text{ref}}, z_t^{\text{ref}})$ is the reference position in x , y , and z directions at the time index t .

Figure 6 shows the trajectories estimated by the LS-SVM and the NN in outage areas #1, #2, #3, and #4. From the figures, one can see easily that in the outage areas #1, #2, #3, and #4, when UWB measurements are unavailable, the NN can still make decisions that are close to the reference path, while the LS-SVM algorithm gives a large accumulated error.

The MSEs estimated by NN (green line) and LS-SVM (blue line) in the outage areas #1, #2, #3, and #4 are shown in Fig. 7. From the figures, one can see that the MSE of the LS-SVM algorithm has a larger accumulated error compared with the NN. The average MSEs Produced by NN and LS-SVM in the outage areas #1, #2, #3, and #4 are listed in Table 1. It can be inferred from the table that the average MSEs of the NN are smaller than the LS-SVM in the outage areas #1, #2, #3, and #4. Compared with the LS-SVM, the proposed NN reduced the localization error by about 54.34%. Thus, we can conclude that the proposed NN-based method can effectively reduce the localization error.

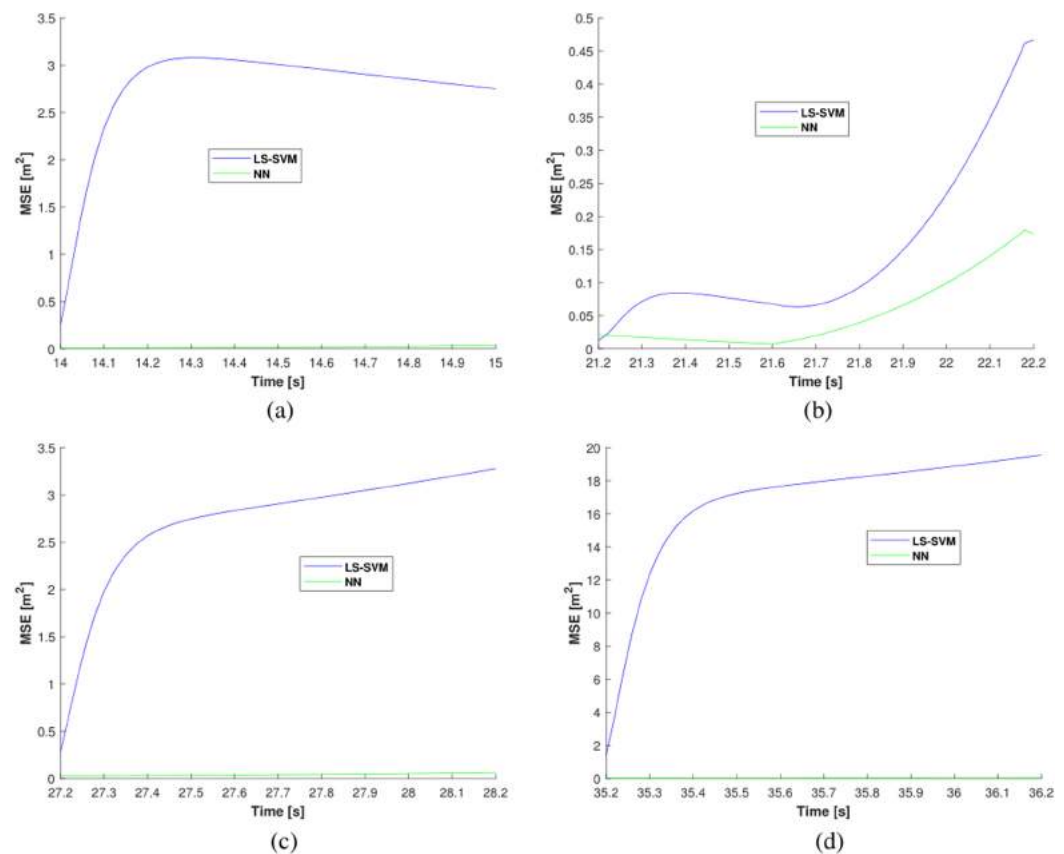


Figure 7 The MSEs estimated by the LS-SVM and the NN in outage areas: (A) outage #1, (B) outage #2, (C) outage #3, and (D) outage #4.

Full-size [DOI: 10.7717/peerj-cs.630/fig-7](https://doi.org/10.7717/peerj-cs.630/fig-7)

Table 1 Average MSEs produced by NN and LS-SVM in outages #1–#4.

Method	MSE (m^2)				Mean
	#1	#2	#3	#4	
LS-SVM	2.7445	0.1453	2.7147	16.6635	5.5670
NN	0.0190	0.0524	0.0422	0.0537	2.5418

CONCLUSION

In this work, in order to make the data fusion filter work properly under the condition that the UWB data is unavailable due to some harsh indoor environments, the NN assisted KF for fusing the UWB and the INS data seamlessly has been investigated. The contributions of this work are summarized as following:

- An NN assisted KF scheme has been designed for fusing the INS and UWB measurement.
- The model of the KF for the integrated scheme has been investigated.
- The NN assisted KF for fusing the UWB and the INS data seamlessly has been investigated. In the proposed approach, the KF provides the localization information

when the UWB data is available. Meanwhile, the KF is used to assist the NN to build the mapping between the $\hat{\mathbf{x}}_{t|t-1}$ and \mathbf{Y}_t off-line. The NN can estimate the measurement vector of the KF when the UWB data is unavailable.

- Real tests have been done to show better performance of the proposed approach.

Based on the results presented in this work, we are now working on further developments of the proposed algorithms to build the mapping with the deep learning and plan to report the results in the near future.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by National Natural Science Foundation of China No. 61803175, Shandong Provincial Key Research and Development Project No. 2019GGX104026, and Shandong Provincial Natural Science Foundation ZR2020KF027. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

National Natural Science Foundation of China: 61803175.

Shandong Provincial Key Research and Development Project: 2019GGX104026.

Shandong Provincial Natural Science Foundation: ZR2020KF027.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Shuhui Bi conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Liyao Ma analyzed the data, prepared figures and/or tables, and approved the final draft.
- Tao Shen conceived and designed the experiments, prepared figures and/or tables, and approved the final draft.
- Yuan Xu performed the computation work, authored or reviewed drafts of the paper, and approved the final draft.
- Fukun Li performed the experiments, prepared figures and/or tables, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The raw data and codes are available in the [Supplemental Files](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.630#supplemental-information>.

REFERENCES

- Benzerrouk H, Nebylov A, Salhi H. 2016. Quadrotor UAV state estimation based on high-degree cubature Kalman filter. *IFAC-PapersOnLine* 49(17):349–354 DOI 10.1016/j.ifacol.2016.09.060.
- Bonny T, Abdelsalam MB. 2019. Autonomous navigation of unmanned aerial vehicles based on android smartphone. *International Journal of Advanced Computer Science and Applications* 10(11):589–598 DOI 10.14569/issn.2156-5570.
- Camci E, Kayacan E. 2019. Learning motion primitives for planning swift maneuvers of quadrotor. *Autonomous Robots* 43(7):1733–1745 DOI 10.1007/s10514-019-09831-w.
- Chiella ACB, Machado HN, Teixeira BOS, Pereira GAS. 2019. GNSS/LiDAR-based navigation of an aerial robot in sparse forests. *Sensors* 19(19):1–22 DOI 10.3390/s19194061.
- Cui B, Wei X, Chen X, Li J, Li L. 2019. On sigma-point update of cubature Kalman filter for GNSS/INS under GNSS-challenged environment. *IEEE Transactions on Vehicular Technology* 68(9):8671–8682 DOI 10.1109/TVT.2019.2931923.
- Do T, Carrillo-Arce LC, Roumeliotis SI. 2019. High-speed autonomous quadrotor navigation through visual and inertial paths. *International Journal of Robotics Research* 38(4):486–504.
- Gośliński J, Kasiński A, Giernacki W, Owczarek P, Gardecki S. 2019. A study on coaxial quadrotor model parameter estimation: an application of the improved square root unscented Kalman filter. *Journal of Intelligent and Robotic Systems* 95(2):491–510.
- Guo S, Chang L, Li Y, Sun Y. 2020. Robust fading cubature Kalman filter and its application in initial alignment of sins. *OPTIK* 202(2):163593 DOI 10.1016/j.ijleo.2019.163593.
- Kou L, Xiang J, Li Y, Bian J. 2018. Stability and nonlinear controllability analysis of a quadrotor-like autonomous underwater vehicle considering variety of cases. *International Journal of Advanced Robotic Systems* 15(6):1–10.
- Li T, Zhang H, Gao Z, Niu X, El-Sheimy N. 2019. Tight fusion of a monocular camera, MEMS-IMU, and single-frequency multi-GNSS RTK for precise navigation in GNSS-challenged environments. *Remote Sensing* 11(6):1–24.
- Liang X, Wang Q, Hu C, Dong C. 2019. Observer-based h-infinity fault-tolerant attitude control for satellite with actuator and sensor faults. *Aerospace Science and Technology* 95(4):105424 DOI 10.1016/j.ast.2019.105424.
- Liu D, Chen X, Xu Y, Liu X, Shi C. 2019. Maximum correntropy generalized high-degree cubature Kalman filter with application to the attitude determination system of missile. *Aerospace Science and Technology* 95(6):105441 DOI 10.1016/j.ast.2019.105441.
- Liu P, Huda MN, Sun L, Yu H. 2020. A survey on underactuated robotic systems: bio-inspiration, trajectory planning and control. *Mechatronics* 72:102443 DOI 10.1016/j.mechatronics.2020.102443.
- Liu P, Yu H, Shuang C. 2019. Adaptive neural network tracking control for underactuated systems with matched and mismatched disturbances. *Nonlinear Dynamics* 98(2):1447–1464 DOI 10.1007/s11071-019-05170-8.
- Nguyen X-M, Hong SK. 2019. Robust adaptive formation control of quadcopters based on a leader-follower approach. *International Journal of Advanced Robotic Systems* 16(4):1–11.
- Qin H-D, Yu X, Zhu Z-B, Deng Z-C. 2020. An expectation-maximization based single-beacon underwater navigation method with unknown ESV. *Neurocomputing* 378:295–303.

- Schmid K, Lutz P, Tomić T, Mair E, Hirschmüller H. 2014.** Autonomous vision-based micro air vehicle for indoor and outdoor navigation. *Journal of Field Robotics* **31**(4):537–570.
- Wang R, Xiong Z, Liu J, Cao Y. 2018.** Stepwise fusion algorithm with dual correction for multi-sensor navigation. *International Journal of Advanced Robotic Systems* **15**(3):1–11.
- Xu Y, Shmaliy YS, Ahn CK, Shen T, Zhuang Y. 2020a.** Tightly-coupled integration of INS and UWB using fixed-lag extended UFIR smoothing for quadrotor localization. *IEEE Internet of Things Journal* **8**(3):1716–1727.
- Xu Y, Shmaliy YS, Chen X, Li Y, Ma W. 2020b.** Robust inertial navigation system/ultra wide band integrated indoor quadrotor localization employing adaptive interacting multiple model-unbiased finite impulse response/Kalman filter estimator. *Aerospace Science and Technology* **98**(3):UNSP 105683.
- Zahran S, Moussa AM, Sesay AB, El-Sheimy N. 2019.** A new velocity meter based on hall effect sensors for UAV indoor navigation. *IEEE Sensors Journal* **19**(8):3067–3076
DOI [10.1109/JSEN.2018.2890094](https://doi.org/10.1109/JSEN.2018.2890094).
- Zhang Y, Dong Z, Wang SH, Yu X, Górriz JM. 2020.** Advances in multimodal data fusion in neuroimaging: overview, challenges, and novel orientation. *Information Fusion* **64**:149–187.
- Zhang Y, Satapathy SC, Guttery DS, Górriz JM, Wang S. 2021.** Improved breast cancer classification through combining graph convolutional network and convolutional neural network—sciencedirect. *Information Processing and Management* **58**(2):102439
DOI [10.1016/j.ipm.2020.102439](https://doi.org/10.1016/j.ipm.2020.102439).
- Zhao S, Huang B. 2020.** Trial-and-error or avoiding a guess? Initialization of the Kalman filter. *Automatica* **121**(21):109184 DOI [10.1016/j.automatica.2020.109184](https://doi.org/10.1016/j.automatica.2020.109184).

SeedSortNet: a rapid and highly efficient lightweight CNN based on visual attention for seed sorting

Chunlei Li¹, Huanyu Li¹, Zhoufeng Liu¹, Bicao Li¹ and Yun Huang²

¹ School of Electrical and Information Engineering, Zhongyuan University of Technology, Zhengzhou, Henan, China

² Xiamen Vision+ Technology Co. Ltd, Xiamen, Fujian, China

ABSTRACT

Seed purity directly affects the quality of seed breeding and subsequent processing products. Seed sorting based on machine vision provides an effective solution to this problem. The deep learning technology, particularly convolutional neural networks (CNNs), have exhibited impressive performance in image recognition and classification, and have been proven applicable in seed sorting. However the huge computational complexity and massive storage requirements make it a great challenge to deploy them in real-time applications, especially on devices with limited resources. In this study, a rapid and highly efficient lightweight CNN based on visual attention, namely SeedSortNet, is proposed for seed sorting. First, a dual-branch lightweight feature extraction module Shield-block is elaborately designed by performing identity mapping, spatial transformation at higher dimensions and different receptive field modeling, and thus it can alleviate information loss and effectively characterize the multi-scale feature while utilizing fewer parameters and lower computational complexity. In the down-sampling layer, the traditional MaxPool is replaced as MaxBlurPool to improve the shift-invariant of the network. Also, an extremely lightweight sub-feature space attention module (SFSAM) is presented to selectively emphasize fine-grained features and suppress the interference of complex backgrounds. Experimental results show that SeedSortNet achieves the accuracy rates of 97.33% and 99.56% on the maize seed dataset and sunflower seed dataset, respectively, and outperforms the mainstream lightweight networks (MobileNetv2, ShuffleNetv2, etc.) at similar computational costs, with only 0.400M parameters (vs. 4.06M, 5.40M).

Submitted 9 April 2021
Accepted 21 June 2021
Published 5 August 2021

Corresponding author
Chunlei Li, lichunlei1979@zut.edu.cn

Academic editor
Pengcheng Liu

Additional Information and
Declarations can be found on
page 17

DOI 10.7717/peerj-cs.639

© Copyright
2021 Li et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Artificial Intelligence, Computer Vision, Embedded Computing, Real-Time and Embedded Systems

Keywords Seed sorting, Computer vision, Lightweight CNN, Attention mechanism

INTRODUCTION

Seed purity directly affects the quality of seed breeding and subsequent processing products. For example, in the process of seed harvest and storage, the impurities or hybrids may be mixed in the normal seed, which results in the economic losses to agricultural production and processing. Therefore, it is crucial to sort impurities and hybrids to ensure that the seed purity meet the market criteria. However, the traditional manual sorting methods are laborious and time-consuming, and hence less efficient. With the evolution of the

technology, there has been a tremendous development in the field of machine vision ([Rehman et al., 2019](#); [Wu et al., 2020](#)) and robot control technology ([Liu, Yu & Cang, 2019](#); [Liu et al., 2020](#); [Han et al., 2020b](#)). The automatic sorting methods ([Li et al., 2019](#)) based on the above technologies provide a promising solution.

Traditional automatic seed sorting methods adopt hand-crafted features for image characterization, such as color, shape, texture, and wavelet features or their combinations ([Liu et al., 2015](#); [HemaChitra & Suguna, 2018](#); [Li et al., 2019](#)). Then, the effective classifiers are employed to realize seed recognition such as linear discriminant analysis (LDA) ([Choudhary, Paliwal & Jayas, 2008](#)), support vector machine (SVM) ([Altuntas et al., 2018](#)), decision tree (DT) ([Kayacan, Sofu & Cetisli, 2016](#)), least square (LS) ([Mebatsion, Paliwal & Jayas, 2013](#)) and artificial neural network (ANN) ([Liu et al., 2015](#)). However, these methods are designed for a specific kind of seed and lack self-adaptivity. In the last three years, mainly due to the advances of deep learning, more concretely convolutional neural networks (CNNs), the quality of image classification ([Krizhevsky, Sutskever & Hinton, 2012](#); [Han et al., 2018](#)), object detection ([Ren et al., 2015](#); [Sun et al., 2019](#); [Bochkovskiy, Wang & Liao, 2020](#)) and semantic segmentation ([Chen et al., 2014](#)) has been progressing at a dramatic pace. Recently, some researchers also adopted deep learning technology in crop identification tasks and achieved good performance ([Ni et al., 2019](#); [Kurtulmus, 2021](#)).

The crop recognition and classification methods, especially for seed sorting, should be deployed on a fast and stable embedded system due to the requirement of higher processing speed. However, the performance of these methods often depends on a deeper, wider network structure, thus it suffers from huge computational complexity and massive storage requirements ([Han et al., 2018](#)). Therefore, the deep CNN model should be compressed and streamlined while maintaining high recognition accuracy. Recently, some lightweight and efficient CNN models have been designed, such as MobileNet ([Howard et al., 2017](#)), MobileNetv2 ([Sandler et al., 2018](#)), ShuffleNet ([Zhang et al., 2018](#)) and ShuffleNetv2 ([Ma et al., 2018](#)) for the real-time detection and recognition tasks. However, due to the lower discrimination of different types of seed, the feature extraction ability of these models is insufficient, thus leads to low recognition accuracy.

In this paper, a lightweight CNN based on visual attention for seed sorting is proposed. A dual-branch lightweight feature extraction module (i.e., Shield-block) is designed to improve the feature characterization ability while utilizing fewer parameters and lower computational complexity, and the traditional MaxPool is replaced as MaxBlurPool ([Zhang, 2019](#)) to improve the shift-invariant of the network in the down-sampling layer. In addition, an extremely lightweight sub-feature space attention module (SFSAM) is proposed as the basic unit of the built CNN model to selectively emphasize fine-grained features and suppress the interference from complex backgrounds. Overall, ours contributions are three-fold as follows:

- We designed a dual-branch lightweight feature extraction module (i.e., Shield-block) to alleviate information loss and effectively characterize the multi-scale feature while utilizing fewer parameters and lower computational complexity.

- We proposed an extremely lightweight sub-feature space attention module, which divides the feature maps into different subspaces and infers different attention maps for each subspace. To selectively emphasize fine-grained features, and suppress the interference of complex backgrounds.
- Experiments are conducted on the maize seed dataset and sunflower seed dataset, and the results show that SeedSortNet achieves higher accuracy compared with the mainstream lightweight networks (MobileNetv2, ShuffleNetv2, etc.) at the similar computational cost, even outperforms the deeper and wider networks, such as VGG ([Simonyan & Zisserman, 2014](#)), GoogleNet ([Szegedy et al., 2015](#)), and ResNet ([He et al., 2016](#)).

The remainder of the paper is organized as follows. In ‘Related Work’, we summarize some related work on seed sorting and lightweight model design. ‘Proposed Method’ introduces the technical details of the proposed method and network architecture. In ‘Experiments’, we carry out a series of comparative experiments on maize and sunflower seed datasets and the experimental results are analyzed. Finally, we conclude in ‘Discussion’.

RELATED WORK

In the following, we review the existing crop identification methods and related technologies, such as CNN model compression and lightweight model design.

Crop identification

Agricultural product assessment and recognition based on machine vision technology have been a research focus in agricultural applications, which is widely used in the detection and sorting of agricultural products such as wheat, corn, fruits, and the identification of plant diseases and insect pests.

[Liu et al. \(2015\)](#) proposed a novel soybean seed sorting based on neural network. Eight shape features, three-color features, and three texture features are extracted to characterize the soybean seed, and BP neural network is used as the classification model to recognize the different defects. Experiments are conducted on the collected image set which includes 857 images of soybean seeds with insect damage, mildew, and other defects, and the results achieve an average recognition accuracy of 97.25%. [Huang \(2012\)](#) proposed a neural network-based quality evaluation and classification method for areca nuts. The axis length, secondary axis length, axis number, area, perimeter, compactness, and the average gray level are used as the feature, and a back-propagation neural network classifier is employed to sort the quality of the areca nuts. [Aznan et al. \(2016\)](#) adopted machine vision methods to discriminate the variety of cultivated rice seed, namely M263. They firstly extracted different morphological features and then adopted a stepwise discriminant function analysis (DFA) to classify different types of rice. The classification accuracy for testing and training sets is 96% and 95.8%, respectively. [HemaChitra & Suguna \(2018\)](#) presented a novel sorting method of Indian pulse seeds based on image analysis techniques. In this method, they extracted the colors, shapes, and texture features, and adopted SVM for classification. The accuracy of their method can reach 98.9% accuracy. [Li et al. \(2019\)](#) designed a system to distinguish different damaged types of corn. An image database including normal corn

and six different damaged corns is constructed. The features such as color and shape are extracted, then the maximum likelihood classifier is leveraged to discriminate these corns. Experiment results show that the classification accuracy is above 74% for all the classes. However, these methods adopt the handcraft features designed for the specific crops and the traditional classifier for sorting, and suffer from poor adaptability and low accuracy.

Due to the excellent feature representation ability, the deep learning models represented by CNN have achieved good performance in image classification, object detection, and semantic segmentation, and have also been successfully applied in plant disease detection and crop type classification. [Sladojevic et al. \(2016\)](#) proposed a CNN-based system to identify 13 types of plant diseases out of healthy leaves. The performance of this approach exhibited a top-1 success of 96.3%. [Veeramani, Raymond & Chanda \(2018\)](#) studied the effect of the number of convolution kernels in the two layers CNN on the recognition performance of haploid and diploid corn seeds. [Veeramani, Raymond & Chanda \(2018\)](#) adopted VGG19 and GoogleNet to classify corn seed defects and analyzed the influence of the two networks with different depths on the recognition performance. [Dolata & Reiner \(2018\)](#) proposed a method for the classification of barley varieties based on CNN, which is based on two separate convolutional layers to analyze dorsal and ventral sides, respectively. The network is trained on a small sample set of 200-500 cases in 8 categories, and the classification accuracy reaches 97%. [Kurtulmuş \(2021\)](#) adopted AlexNet ([Krizhevsky, Sutskever & Hinton, 2012](#)), GoogleNet, and ResNet to identify sunflower seed varieties, and then they were also evaluated in terms of both accuracy and training time, GoogleNet obtained the highest classification accuracy (95%).

The CNN-based crop identification method can achieve the better recognition rate and has higher self-adaptivity. However, the performance of the deep learning method depends on the depth and width of the model, the researchers often boost the depth and width of the model to improve the performance of the detection and recognition system. But this strategy results in slow speed and difficult deployment in industrial applications.

Model lightweight

For the specific application of crop seed sorting, due to the extremely fast production speed, it is necessary to develop the lightweight CNN model while maintaining a higher recognition accuracy. To trade off the model size and performance for deep neural network architectures has been an active research area, the related technologies include model compression, lightweight network design, etc ([Liang et al., 2021](#)).

Model compression

Model compression aims at generating the small network models from the trained large network models while keeping the performance. The typical techniques include pruning, quantization, and knowledge distillation. Pruning technology is based on the assumption that many parameters in the deep neural network are redundant, then the weights ([Guo, Yao & Chen, 2016](#); [Aghasi et al., 2016](#); [Liu et al., 2018](#)) or filters ([Li et al., 2016](#); [Liu et al., 2017](#); [Lin et al., 2020](#)) with low correlation can be removed to make the network structure sparse. Quantization methods aim to deploy the CNN model on the terminal hardware and encode

the weights and activations using 8-bit integers (INT8) without incurring a significant loss in accuracy. Some other quantization methods even adopt INT4 or lower, such as binary quantization ([Courbariaux et al., 2016](#)) and ternary quantization ([Mellempudi et al., 2017](#)) to reduce the model size. Knowledge distillation is firstly proposed by [Bucilu, Caruana & Niculescu-Mizil \(2006\)](#) and generalized by [Hinton, Vinyals & Dean \(2015\)](#) and can generate a small student network by learning the behavior of a large teacher network. [Cho & Hariharan \(2019\)](#) empirically analyzed in detail the efficacy of knowledge distillation. However, compared with the original network, model compression is difficult to achieve better performance. The compression size is too large, which will lead to significant decrease of performance.

Lightweight network design

Lightweight network design refers to the redesign of the network structure based on the existing CNN model to reduce the parameters and the computational complexity. [Lin, Chen & Yan \(2013\)](#) proposed a Network-In-Network architecture, which used 1×1 convolution to increase network performance while maintaining a lower computational complexity. SqueezeNet ([Iandola et al., 2016](#)) is a lightweight network structure based on 1×1 convolution. The squeeze and expand module proposed by this model can effectively reduce the parameters while ensuring recognition accuracy. The recognition accuracy of the proposed method can be up to 57.55%, and it is similar with the AlexNet with the model size of $50 \times$ smaller. Google developed two efficient architectures denoted as MobileNet ([Howard et al., 2017](#)) and MobileNetV2 ([Sandler et al., 2018](#)) in 2017 and 2018, respectively. MobileNet proposed depthwise separable convolutions to reduce the computational complexity and achieved the state-of-art accuracy with low latency. Thereafter, the linear bottleneck with inverted residual structure is proposed in MobileNetV2 to construct a more efficient architecture. ShuffleNet ([Zhang et al., 2018](#)) proposed the pointwise group convolution and channel shuffle operations to improve the recognition accuracy while reducing latency. Combining the advantage of MobileNet and ShuffleNet, [Ma et al. \(2018\)](#) proposed ShuffleNetV2, which improves group convolution by channel split and used channel shuffle for the split channel as well. [Wang & Yu \(2020\)](#) proposed the Tied Block Convolution (TBC) which shares the same thinner filters over equal blocks of channels and produces multiple responses with a single filter, to design a lightweight model. GhostNet ([Han et al., 2020a](#)) applied a series of linear transformations to generate many Ghost feature maps, and it can characterize the required information from the original features at a small cost, which effectively reduces calculation and parameters. However, due to the low discrimination of crop seeds, the recognition accuracy will be significantly reduced when the existing lightweight models are directly applied to the seed sorting. Therefore, a rapid and highly efficient lightweight CNN model should be developed based on the characteristics of crop seeds while keeping the accuracy.

PROPOSED METHOD

Seed sorting based on deep learning is a promising method for seed breeding and subsequent processing products. In this paper, we proposed a rapid and efficient lightweight

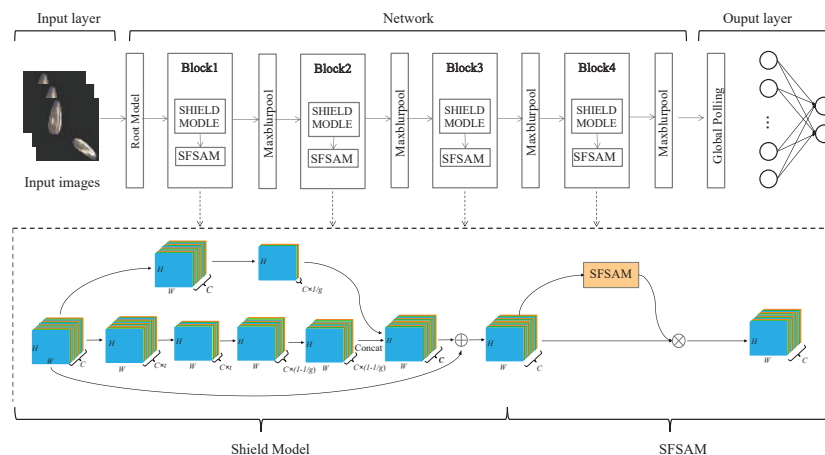


Figure 1 The flowchart of the proposed SeedSortNet model.

Full-size [DOI: 10.7717/peerjcs.639/fig-1](https://doi.org/10.7717/peerjcs.639/fig-1)

CNN model with a dual-branch network structure based on visual attention for seed sorting, denoted as SeedSortNet. It is an efficient and lightweight end-to-end recognition framework, which is mainly composed of sequential cascade layers and basic blocks, and the overall structure of the model is shown in Fig. 1. First, a dual-branch lightweight feature extraction module, namely Shield-block, is designed for effective feature extraction. Then, the traditional convolution is replaced by depthwise convolution and pointwise convolution to achieve the trade-off between classification accuracy and efficiency. Moreover, in the down-sampling layer, MaxPool is substituted as MaxBlurPool to improve the shift-invariant of the network. Finally, we propose an extremely lightweight sub-feature space attention module to selectively emphasize fine-grained features and suppress the interference of complex backgrounds. And the proposed method is specifically described as follows.

Network construction

Due to the required higher processing speed and recognition accuracy, the representative deep neural network model (eg, ResNet, VGG, GoogLeNet, etc.) cannot efficiently tackle with the seed sorting task because of the lower efficiency and insufficient feature extraction ability. To address these issues, we construct a novel lightweight and efficient network which consists of Root-model, Shield-block, and a novel down-sampling module.

A.Root-model. To effectively improve the feature representation ability while reducing calculation, a dual-branch structure, namely Root-model (Fig. 2A), is designed as the first stage of SeedSortNet. First, the sixteen 3×3 filters are utilized to extract the shallow feature information (such as texture, shape, color, etc.) of the test image. Then, in one branch, the MaxBlurPool which is a non-overlapping 2×2 window is designed for reducing the aliasing effect and improving the shift-invariant of the network. In another branch, we firstly use 3×3 filters with the stride of 2 to convolute the input features and then adopt 1×1 filters

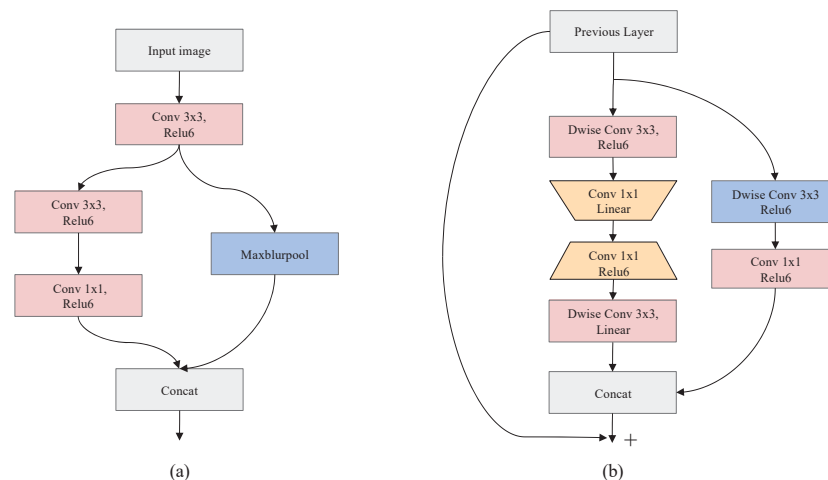


Figure 2 The schema of Root-model (A) and Shield-block (B).

Full-size [DOI: 10.7717/peerjcs.639/fig-2](https://doi.org/10.7717/peerjcs.639/fig-2)

to reduce the output dimension of the branch. Finally, the features generated by the two branches are concatenated together as the input of the next layer.

B.Shield-block. The inverted residual block (Sandler et al., 2018) which shifts the identity mapping from high-dimensional representations to low-dimensional ones (i.e., the bottlenecks), has been successfully applied in the design of lightweight networks. However, the connection of identity mapping between thin bottlenecks would inevitably lead to information loss since the residual representations are compressed (Daquan et al., 2020). In addition, this connection would also weaken the propagation capability of gradients across layers due to gradient confusion arising from the narrowed feature dimensions, and hence affect the training convergence and model performance (Sankararaman et al., 2020). To address these issues, we propose a dual-branch feature extraction module by improving the inverted residual block in this article (shown in Fig. 2B).

In the main branch, two 3×3 depthwise convolution layers are utilized to encode richer spatial information to generate a more expressive representation. Then we adopt two pointwise convolutional layers between two 3×3 depthwise convolutional layers, the first point convolution layer reduces the feature channel dimension and the latter increases its dimension, to encode the cross-channel information of the feature maps and reduce the computational complexity. Also, the linear activation function is adopted after the first pointwise convolutional layer and the last depthwise convolutional layer, which can prevent the feature values from being zeroed and hence reduce information loss.

For the other branch, a 3×3 depthwise separable convolution is designed to acquire the spatial representation of different receptive fields, and thus improve the feature representation ability. In the end, the concatenation of the two branches and its shortcut connection with the input feature are combined as the final output.

In the following, we present the detailed data processing operator of Shield-block, and it is shown in Table 1, where H , W and C represents the height, width, and channel number of the feature map, and $1/t$ represents the reduction rate of channels. Moreover, to ensure

Table 1 Data processing in the Shield-block.

Input dimension	Operator	Output dimension
$\begin{bmatrix} H \times W \times C \\ H \times W \times C \\ H \times W \times C \times \frac{1}{t} \\ H \times W \times C \times \left(1 - \frac{1}{r}\right) \end{bmatrix}$	$\begin{bmatrix} 3 \times 3 \text{ Dwise Conv, ReLU6} \\ 1 \times 1 \text{ Conv, Linear} \\ 1 \times 1 \text{ Conv, ReLU6} \\ 3 \times 3 \text{ Dwise Conv, Linear} \end{bmatrix} \quad \text{Concat} \quad \begin{bmatrix} 3 \times 3 \text{ Dwise Conv, ReLU6} \\ 1 \times 1 \text{ Conv, ReLU6} \end{bmatrix}$	$\begin{bmatrix} H \times W \times C \\ H \times W \times C \times \frac{1}{t} \\ H \times W \times C \times \left(1 - \frac{1}{r}\right) \\ H \times W \times C \times \left(1 - \frac{1}{r}\right) \end{bmatrix}$

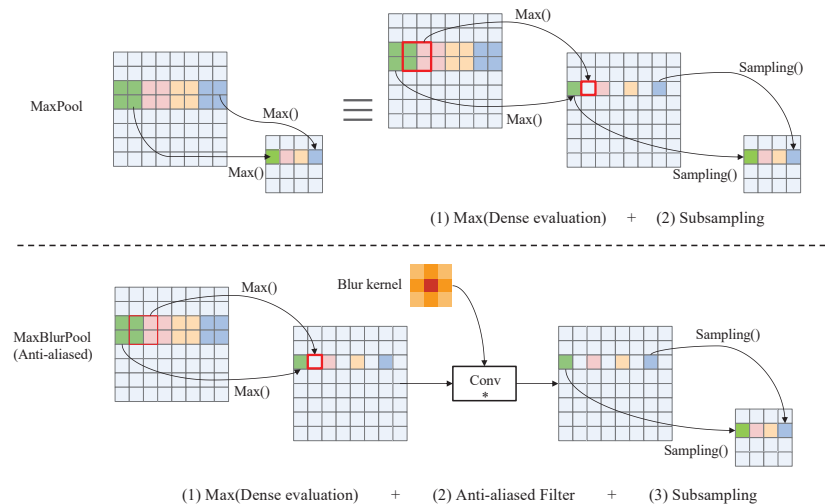


Figure 3 Operation details of MaxPool and anti-aliased MaxBlurPool.

Full-size [DOI: 10.7717/peerjcs.639/fig-3](https://doi.org/10.7717/peerjcs.639/fig-3)

the input channel dimension is consistent with the output channel dimensional, and the hyperparameter $1/r$ (referring to the proportion of the input channel of the sub-branch output channel) is adopted in this paper, here we empirically set r to 6.

C.Down-sampling. Down-sampling operator can reduce the feature dimensionality while retaining the valid information. Traditional down-sampling methods (eg. MaxPool, Strided-Convolution, AvgPool, etc.) violate the shift-equivariance and results in small shifts in the input that can drastically change the output (Azulay & Weiss, 2018). And this phenomenon will become more obvious with the increase of the network depth. To solve this problem, a fuzzy sampling method, namely MaxBlurPool, proposed in (Zhang, 2019) is adopted in our method. The specific process of MaxPool and MaxBlurPool is shown in Fig. 3. From this figure, we can see that a blur kernel is inserted between max and subsampling to remove aliased in the MaxBlurPool method, thereby improving the shift-invariant and enhancing the robustness of the CNN model.

Lightweight sub-feature space attention module (SFSAM)

Due to the low discrimination of different types of seeds, the fine-grained spatial features are crucial for seed sorting. Therefore, an extremely lightweight sub-feature space attention module is proposed to selectively emphasize fine-grained features, and suppress the

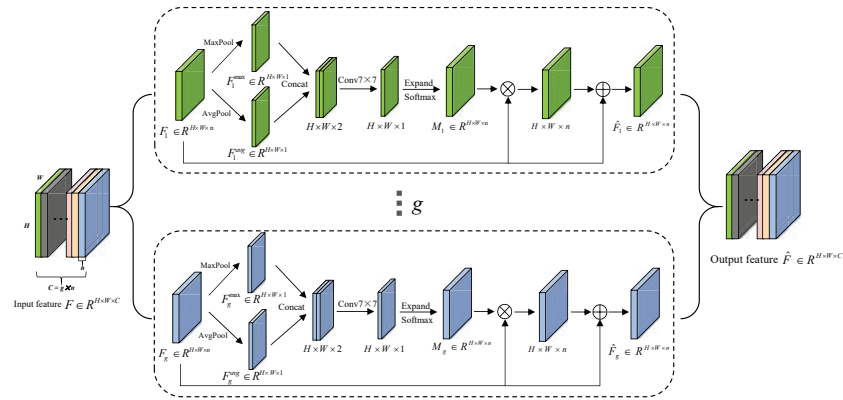


Figure 4 The schematic diagram of SFSAM structure.

Full-size [DOI: 10.7717/peerjcs.639/fig-4](https://doi.org/10.7717/peerjcs.639/fig-4)

interference of complex backgrounds. It divides the feature maps into different subspaces and infers different attention maps for each subspace, thus can generate the multi-scale feature representation, it is shown in Fig. 4. The detailed process is described as follows.

The input feature map $F \in R^{H \times W \times C}$ is firstly divided into g mutually exclusive groups $[F_1, F_2, F_3, \dots, F_i, \dots, F_g]$ (i.e., sub-feature spaces), where each sub-feature space F_i contains n intermediate feature maps. Zagoruyko & Komodakis (2016) have proved that pooling operations along the channel axis are effective in highlighting informative regions. Therefore, AvgPool and MaxPool operations are applied to g sub-feature spaces along the channel axis to generate g groups of average-pooled features $F_i^{\max} \in R^{1 \times H \times W}$ and max-pooled features $F_i^{\text{avg}} \in R^{1 \times H \times W}$. Then, these features are concatenated separately to generate g efficient feature descriptors $[F_i^{\max}, F_i^{\text{avg}}]$. Thereafter, the g group's subspace attention maps are generated using Eq.(1).

$$M_i = \text{softmax}(f^{k \times k}([\text{MaxPool}(F_i), \text{AvgPool}(F_i)])) \\ = \text{softmax}(f^{k \times k}([F_i^{\max}, F_i^{\text{avg}}])) \quad (1)$$

where $f^{k \times k}$ represents a convolution operation with a filter size of $k \times k$. In this paper, k is empirically set to 7. The attention map in each group (subspace) can capture the non-linear dependencies among the feature maps by learning to gather cross-channel information. Meantime, we employ a gating mechanism with a softmax activation to map the attention weighting tensor into $[0, 1]$.

Then, each group of feature maps gets the refined set of feature maps (\hat{F}_i) after the feature redistribution in Eq. (2).

$$\hat{F}_i = (M_i \otimes F_i) \oplus F_i \quad (2)$$

where \otimes is element-wise multiplication and \oplus is element-wise addition.

The final output \hat{F} of SFSAM is obtained by concatenating the feature maps of each group, and it is described as Eq. (3).

$$\hat{F} = \text{concat}(\hat{F}_1, \hat{F}_2, \hat{F}_3, \dots, \hat{F}_i, \dots, \hat{F}_g). \quad (3)$$

Table 2 Parameter configuration diagram of the SeedSortNet.

Stage	Input	Operator	t	C	R
1	$224 \times 224 \times 3$	Root-module	—	32	1
2	$112 \times 112 \times 32$	Shield-block	2 6	64	2
3	$112 \times 112 \times 64$	SFSAM	—	64	1
4	$112 \times 112 \times 64$	MaxBlurPool	—	64	1
5	$56 \times 56 \times 64$	Shield-block	2 6	128	4
6	$56 \times 56 \times 128$	SFSAM	—	128	1
7	$56 \times 56 \times 128$	MaxBlurPool	—	128	1
8	$28 \times 28 \times 128$	Shield-block	2 6	128	5
9	$28 \times 28 \times 128$	SFSAM	—	192	1
10	$28 \times 28 \times 192$	MaxBlurPool	—	192	1
11	$14 \times 14 \times 192$	Shield-block	2 6	256	4
12	$14 \times 14 \times 256$	SFSAM	—	256	1
13	$14 \times 14 \times 256$	MaxBlurPool	—	256	1
14	$7 \times 7 \times 256$	GlobalAvgpool	—	256	—
15	$1 \times 1 \times 256$	Dropout 2D-FC	—	2	—

Network topology

In this paper, a novel lightweight CNN model with a dual-branch network structure based on visual attention, denoted as SeedSortNet, is proposed for seed sorting with higher efficiency and recognition accuracy. The setting of the proposed SeedSortNet is outlined in Table 2. Each row denotes a sequence of building blocks, which is the repeated times of ‘R’. The reduction ratio of channels is used in each Shield-block is denoted by ‘ $1/t$ ’, and ‘C’ represents the number of channels in the output feature map.

We first use Root-module to generate 32 feature maps with the size of 112×112 . Then, it is followed by the 15 Shield-blocks, four SFSAM attention modules, and four down-sampling layers (i.e., MaxBlurPool) spatial location distributions described in Table 2. At the first Shield-block of stages 2, 5, 8, and 11, the identity mappings do not need to be set because of the increasement of the feature map depth. Besides, we set ‘ t ’ to 2 to avoid the information loss due to the low-dimensional input. Finally, the output of the fourth down-sampling layer is followed by a global average pooling layer, which can convert 2D feature maps into 1D feature vectors.

EXPERIMENTS

Experimental datasets

In this section, two datasets are selected for verifying the effectiveness of the proposed network architecture.

Maize seed dataset

The first dataset is a public haploid and diploid maize seed dataset of the maize research institute in Sakarya (Turkey), including 3000 RGB images of corn seeds (Altunta, Cömert & Kocamaz, 2019), and it includes 1230 haploid seeds images and 1770 diploid seed images. The dimensions of these images depend on the sizes of the seeds and vary between 300×289

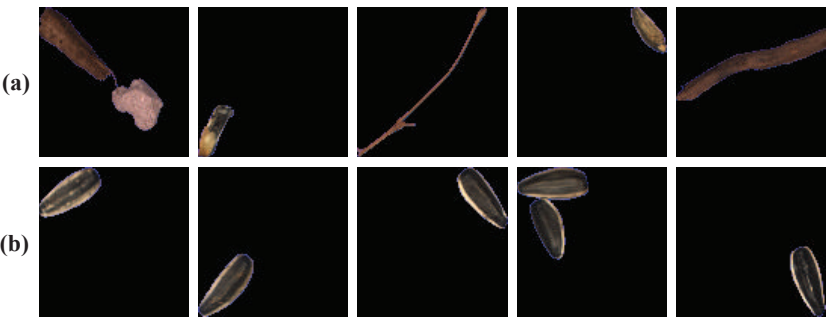


Figure 5 Representative samples of sunflower seed dataset. (A) Abnormal sunflower seeds; (B) normal sunflower seeds.

[Full-size](#) [DOI: 10.7717/peerjcs.639/fig-5](https://doi.org/10.7717/peerjcs.639/fig-5)

Table 3 Category distribution and proportion of the training set on maize seed dataset and sunflower seed dataset.

Dataset	Category distribution	The proportion of the training set
Maize	1770 (Diploid) 1230 (Haploid)	75%
Sunflower	7837 (Normal) 7997 (Abnormal)	≈75%

pixels and 610×637 pixels. In the experiment, three-quarters of the dataset are used for training, and the remaining images are used for testing, as shown in [Table 3](#).

The number of maize seed dataset is limited and may bring the overfitting for the proposed model. Therefore, the data augmentation methods, such as horizontal flip, vertical flip, and angle rotation, are adopted to augment the maize seed data set by a factor of 4. The experimental results prove that such a large dataset is enough to train a model with very strong generalization ability.

Sunflower seed dataset

To thoroughly evaluate the effectiveness of the proposed method, we constructed our sunflower seed dataset on an industrial production line for the experiments. The image acquisition device equipped with a color line scan camera is established to collect 15834 sunflower seed RGB images with the size of 100×100 pixels. And we divided them into two categories, as shown in [Fig. 5](#). The top row is the abnormal seed images composed of leaves, stones, defective seeds, etc. The bottom row is the normal sunflower seed images. It is worth noting that when the picture contains several seeds and impurities or hybrids, we will classify them as abnormal to ensure a low false alarm. In our experiment, about three-quarters of the dataset are randomly selected as the training set, and the remaining images are used for testing, as shown in [Table 3](#).

Implementation details and evaluation metric

Implementation details

All experiments were performed on a 64-bit Linux-based operation system, Ubuntu 18.04. The software is mainly based on the deep learning architecture of Pytorch and python development environment Spyder. The hardware is based on an Intel(R) Xeon(R)

Table 4 Calculation formulas and explanations of binary class metrics.

Measure	Formulation	Evaluation Focus
Accuracy (Acc)	$\frac{tp+tn}{tp+fp+tn+fn}$	The overall accuracy of a model.
Precision (p)	$\frac{tp}{tp+fp}$	The ratio of correctly classified positive samples to estimated total positive sample.
Recall (r)	$\frac{tp}{tp+fn}$	The proportion of positive values classified as true.
F1-score	$\frac{2 \cdot p \cdot r}{p+r} = \frac{2 \cdot \frac{tp}{tp+fp} \cdot \frac{tp}{tp+fn}}{\frac{tp}{tp+fp} + \frac{tp}{tp+fn}}$	The harmonic mean between precision and recall.

CPU E5-2650 v4 @2.20 GHz and two NVIDIA Quadro M5000 GPUs, with CUDA10.2 accelerating calculation.

And we train the network by mini-batch SGD, with an initial learning rate of 0.001 and a reducing factor of 0.1 after 30 epochs. The momentum parameter is set to 0.9 and the weight decay parameter is 0.0001. The number of iterations in training is 100, and the batch size is set to 16 and 64 on the maize and sunflower datasets, respectively. Besides, the input image size is resized to 224×224 -pixel by the CenterCrop function, and the parameter g is set to 4 by analyzing experimental results.

Evaluation metric

To quantitatively evaluate the effectiveness of the proposed method, four metrics, such as true positive (tp), true negative (tn), false positive (fp), and false negative (fn) are adopted in our method. tp is the true positive and represents correctly recognized haploid maize seeds or the normal sunflower seed. tn is the true negative and represents correctly recognized diploid maize seeds or the abnormal sunflower seed. fp is the false positive and represents the falsely recognized haploid maize seeds or the normal sunflower seed. fn is the false negative and represents falsely recognized diploid maize seeds or the abnormal sunflower seed. Based on these metrics, four evaluation metrics, accuracy (Acc), precision (p), recall (r) and F1-score, are calculated as Table 4.

It should be noted that the F1-score metric can better interpret the true performance when the number of samples is not balanced. Receiver operating characteristic (ROC) curves are also a useful tool for measuring a model performance without considering class distribution or error costs. Also, the number of parameters and required float points operations (denoted as FLOPs) are also employed to evaluate the model size and computational complexity, which are widely-used protocols.

Result analysis

Results on maize seed dataset

To assess the performance of our network (i.e., SeedSortNet) in the maize dataset. Six representative CNN models (i.e., AlexNet, VGG, ResNet, GoogleNet, DenseNet (Huang et al., 2017), and Resnext (Xie et al., 2017) are selected to conduct comparative experiments. The experimental results are shown in Table 5.

From the results in Table 5, we can observe that the adopted network can achieve good classification accuracy, and reach more than 90% under the same experimental environment. SeedSortNet has the best performance, with accuracy, precision, recall,

Table 5 Performance comparison of different network on maize seed dataset.

Model	Parameters	FLOPs	Acc	Precision	Recall	F1-score
AlexNet	57.01M	711.46M	93.33	93.39	92.80	93.07
VGG11	128.77M	7.63G	94.67	94.54	94.43	94.48
VGG13	128.96M	11.33G	94.50	94.51	94.10	94.29
ResNet18	11.18M	1.82G	95.33	95.18	95.18	95.18
ResNet50	23.51M	4.12G	96.00	95.73	96.05	95.88
DenseNet121	6.96M	2.88G	95.83	95.58	95.85	95.71
GoogleNet	5.60M	1.51G	96.67	96.50	96.62	96.56
ResNext101	86.75M	16.48G	96.00	95.77	95.99	95.88
SeedSortNet	0.40M	512.06M	97.33	97.30	97.18	97.24

Table 6 Performance comparison of maize seed dataset in lightweight CNNs.

Model	Parameters	FLOPs	Acc	F1-score
MobileNetv1	3.22M	587.94M	94.00	93.81
MobileNetv2 1.4×	4.06M	566.33M	96.00	95.89
ShuffleNetv1 2×(g=3)	3.53M	537.48M	96.67	96.55
ShuffleNetv2 2×	5.35M	591.79M	96.00	95.90
GhostNet 2×	12.96M	529.89M	96.50	96.41
SeedSortNet	0.40M	512.06M	97.33	97.24
MobileNetv1 0.75×	1.83M	339.80M	91.83	91.65
MobileNetv2	2.23M	318.96M	95.83	95.71
ShuffleNetv1 1.5×(g=3)	2.00M	301.90M	96.17	96.05
ShuffleNetv2 1.5×	2.48M	302.65M	95.50	95.38
GhostNet 1.5×	7.79M	310.76M	95.83	95.72
SeedSortNet 0.75×	0.23M	338.64M	97.00	96.90

and F1-score of 97.33%, 97.30%, 97.18%, and 97.24%, respectively, with a relatively low computational complexity and model size. These results verify the effectiveness of the proposed method.

Meanwhile, we also compared the performance of mainstream lightweight CNN models (eg, MobileNetv1, MobileNetv2, ShuffleNetv1, ShuffleNetv2, GhostNet) under different calculation benchmarks. The experimental results on the maize dataset in terms of computational complexity, model parameters, classification accuracy, and F1-score are shown in Table 6. The models are typically grouped into two levels of computational complexity for embedded device applications, i.e., ~300MFLOPs and 500 ~600MFLOPs. From the results, we can see that the larger FLOPs lead to higher accuracy in these lightweight networks. SeedSortNet outperforms other competitors consistently in classification accuracy and F1-score at various computational complexity levels. Furthermore, the number of parameters has also greatly decreased for the proposed method.

In order to further demonstrate the effectiveness of the proposed method, ROC curve is adopted to measure the model performance. Figures 6A–6C shows the ROC curves and

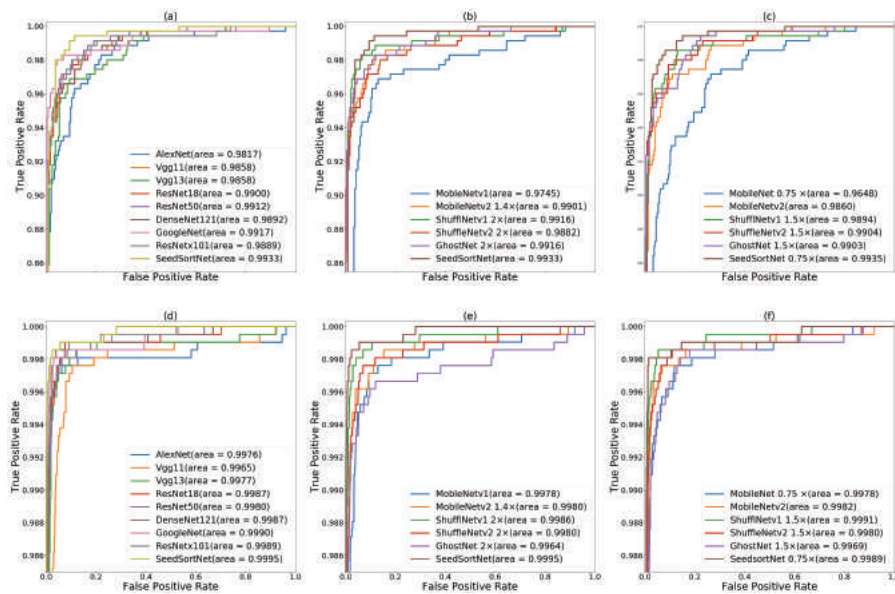


Figure 6 ROC curves of the CNN models on maize seed dataset (A, B, C) and sunflower seed dataset (D, E, F). (A & D) ROC curves of SeedSortNet and Six representative CNN models (i.e., AlexNet, VGG, ResNet, GoogleNet, DenseNet, and Resnext), (B & E) ROC curves of the lightweight network (500 ~600MFLOPs), (C & F) ROC curves of the lightweight network (~300MFLOPs).

Full-size [DOI: 10.7717/peerjcs.639/fig-6](https://doi.org/10.7717/peerjcs.639/fig-6)

the calculated area under curve (AUC) scores for using the proposed method and other network models (i.e., the above comparison network) on the maize seed dataset. From AUC scores, it is observed that the method achieve the best result of 99.33% compared with other models, which is superior to the above representative representative CNN models and lightweight networks.

Results on sunflower seed dataset

Table 7 demonstrates the model size, computational complexity, accuracy, recall, specificity and F1-score of different network models on the sunflower seed dataset. From the table, we can see that the proposed method has the highest accuracy, precision, recall, and F1-score, while it has the lower FLOPs and parameters.

Similar to the maize seed sorting, we also conducted comparative experiments with the mainstream lightweight CNN models (eg, MobileNetv1, MobileNetv2, ShuffleNetv1, ShuffleNetv2, GhostNet) under different calculation benchmarks on the sunflower seed dataset. From **Table 8**, we can find that the classification accuracy and F1-score of the SeedSortNet are higher than other network models under a similar calculation cost. Meantime, we find that the test dataset is relatively balanced, thus its F1-score and accuracy are almost the same. Therefore, the proposed SeedSortNet is more suitable for deployment on edge devices and has the ideal sorting accuracy.

In **Figs. 6D–6F**, we find that the AUC score of the proposed method is closer to 1.0 (i.e., 0.9995) on the sunflower seed dataset compared with other CNN models, which demonstrates SeedSortNet has a good ability to prevent misclassification.

Table 7 Performance comparison of different network on sunflower seed dataset.

Model	Parameters	FLOPs	Acc	Precision	Recall	F1-score
AlexNet	57.01M	711.46M	99.00	99.01	98.99	99.00
VGG11	128.77M	7.63G	97.78	97.78	97.80	97.78
VGG13	128.96M	11.33G	98.44	98.43	98.45	98.44
ResNet18	11.18M	1.82G	99.05	99.04	99.05	99.05
ResNet50	23.51M	4.12G	99.22	99.22	99.22	99.22
DenseNet121	6.96M	2.88G	98.90	98.89	98.91	98.90
GoogleNet	5.60M	1.51G	99.34	99.34	99.34	99.34
ResNext101	86.75M	16.48G	99.22	99.22	99.22	99.22
SeedSortNet	0.40M	512.06M	99.56	99.56	99.56	99.56

Table 8 Performance comparison of sunflower seed dataset in lightweight CNNs.

Model	Parameters	FLOPs	Acc	F1-score
MobileNetv1	3.22M	587.94M	98.36	98.36
MobileNetv2 1.4×	4.06M	566.33M	98.83	98.83
ShuffleNetv1 2×(g=3)	3.53M	537.48M	99.19	99.19
ShuffleNetv2 2×	5.35M	591.79M	99.00	99.00
GhostNet 2×	12.96M	529.89M	98.80	98.80
SeedSortNet	0.40M	512.06M	99.56	99.56
MobileNetv1 0.75×	1.83M	339.80M	98.32	98.31
MobileNetv2	2.23M	318.96M	98.90	98.90
ShuffleNetv1 1.5×(g=3)	2.00M	301.90M	99.12	99.12
ShuffleNetv2 1.5×	2.48M	302.65M	98.73	98.73
GhostNet 1.5×	7.79M	310.76M	98.44	98.44
SeedSortNet 0.75×	0.23M	338.64M	99.34	99.34

Ablation study

The ablation study is carried on SeedSortNet and the network without SFSAM attention mechanism. The experimental results in Table 9 show that F1-score of 96.33% and 99.37% are obtained without SFSAM on the maize and sunflower seed datasets, respectively, which proves that Root-model and Shield block have better information extraction abilities. Meanwhile, SeedSortNet can get 97.33% and 99.56% F1-score, respectively. These demonstrate that SFSAM can selectively emphasize information features and suppress the interference of complex backgrounds, thereby improving the performance. At the same time, it can also be observed from Table 9 that SFSAM does not introduce too many parameters and calculations.

Effects of g selection in SFSAM

As described in the SFSAM section, the feature maps are divided into g groups and generate g attention maps. Each attention map can capture cross-channel information from the feature maps in its respective group. When $g = 1$, the cross channel information for the whole feature volume is captured by a single attention map, which is not sufficient

Table 9 F1-score of SeedSortNet and SeedSortNet (without SFSAM) on maize seed dataset and sunflower seed dataset.

Model	Parameters (M)	FLOPs	F1-score(maize)	F1-score (sunflower)
SeedSortNet(without SFSAM)	0.399M	505.26M	96.22	99.39
SortSeedNet	0.400M	512.06M	97.24	99.56

Table 10 F1-score of SeedSortNet (with fewer parameters/FLOPs; $g = 1, 4, 8, 16$) on maize seed dataset and sunflower seed dataset.

Model	Parameters	FLOPs	F1-score(maize)	F1-score(sunflower)
SortSeedNet ($g=1$)	0.399M	506.56M	96.39	99.44
SortSeedNet ($g=4$)	0.400M	512.06M	97.24	99.56
SortSeedNet ($g=8$)	0.402M	518.85M	96.90	99.51
SortSeedNet ($g=16$)	0.405M	532.45M	96.57	99.46

to capture the complex relationships in the entire feature space and will result in lower predictive performance. When $1 < g < C$, the better exchange of cross-channel information can be obtained. Therefore, we conduct experiments on the different parameters assigned by g (such as $g = 1, 4, 8, 16$), and the results in Table 10 confirm the correctness of the above analysis. It can also be observed that the maize and sunflower seed datasets have achieved higher performance gains, and the FLOPs and parameters increase with the increase g . Based on the experimental results, we adopt $g = 4$ to conduct the above series of comparative experiments which provides a reasonable trade-off between preserving good performance and improving computational efficiency.

DISCUSSION

In this paper, we present a rapid and highly efficient lightweight CNN for seed sorting (i.e., SeedSortNet). We first design a novel dual-branch lightweight feature extraction module (i.e., Shield block) for building efficient neural network architectures. In the down-sampling layer, MaxBlurPool is employed instead of frequently-used MaxPool to improve the shift-invariant of the network. Then we proposed a lightweight sub-feature space attention module (SFSAM), which improves the representational power of the model by learning different attention feature maps. A wide range of experiments show the effectiveness of SeedSortNet, which achieves state-of-the-art identification performance on maize seed and sunflower seed datasets while utilizing fewer parameters and lower computational complexity. In future research, the number of seed varieties and images will be further increased to test the performance of these models, and we hope that these methods can be applied in the seed market.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by NSFC (U1804157, No. 61772576, No. 62072489), the Henan Science and Technology Innovation Team (CXTD2017091), IRTSTHN (21IRTSTHN013), the ZhongYuan Science and Technology Innovation Leading Talent Program (14200510013), and the Program for Interdisciplinary Direction Team in Zhongyuan University of Technology. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

NSFC: U1804157, No. 61772576, No. 62072489.

Henan science and technology innovation team: CXTD2017091.

IRTSTHN: 21IRTSTHN013.

ZhongYuan Science and Technology Innovation Leading Talent Program: 14200510013.

Program for Interdisciplinary Direction Team in Zhongyuan University of Technology.

Competing Interests

Yun Huang is employed by Xiamen Vision+ Technology Co. Ltd.

Author Contributions

- Chunlei Li and Huanyu Li conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Zhoufeng Liu conceived and designed the experiments, performed the experiments, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Bicao Li conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Yun Huang conceived and designed the experiments, performed the computation work, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The maize seed dataset comes from Altuntaş et al. (2019): <https://doi.org/10.1016/j.compag.2019.104874> and is available at: <http://www.rovile.org/datasets/haploid-and-diploid-maize-seeds-dataset/>.

The seedsortnet code and sunflower seed dataset are available at GitHub: <https://github.com/Huanyu2019/Seedsortnet>.

REFERENCES

- Aghasi A, Abdi A, Nguyen N, Romberg J. 2016. Net-trim: convex pruning of deep neural networks with performance guarantee. ArXiv preprint. [arXiv:1611.05162](https://arxiv.org/abs/1611.05162).

- Altuntas Y, Kocamaz AF, Cengiz R, Esmeray M. 2018.** Classification of haploid and diploid maize seeds by using image processing techniques and support vector machines. In: *2018 26th signal processing and communications applications conference (SIU)*.
- Altunta Y, Cömert Z, Kocamaz AF. 2019.** Identification of haploid and diploid maize seeds using convolutional neural networks and a transfer learning approach. *Computers & Electronics in Agriculture* **163**(2019):1–11.
- Aznan A, Rukunudin I, Shakaff A, Ruslan R, Zakaria A, Saad F. 2016.** The use of machine vision technique to classify cultivated rice seed variety and weedy rice seed variants for the seed industry. *International Food Research Journal* **23**:S31–S325.
- Azulay A, Weiss Y. 2018.** Why do deep convolutional networks generalize so poorly to small image transformations? ArXiv preprint. [arXiv:1805.12177](https://arxiv.org/abs/1805.12177).
- Bochkovskiy A, Wang C-Y, Liao H-YM. 2020.** Yolov4: optimal speed and accuracy of object detection. ArXiv preprint. [arXiv:2004.10934](https://arxiv.org/abs/2004.10934).
- Bucilu C, Caruana R, Niculescu-Mizil A. 2006.** Model compression. In: *Proceedings of the 12th ACM SIGKDD international conference on knowledge discovery and data mining*. New York: ACM, 535–541.
- Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL. 2014.** Semantic image segmentation with deep convolutional nets and fully connected crfs. ArXiv preprint. [arXiv:1412.7062](https://arxiv.org/abs/1412.7062).
- Cho JH, Hariharan B. 2019.** On the efficacy of knowledge distillation. In: *Proceedings of the IEEE/CVF international conference on computer vision*. 4794–4802.
- Choudhary R, Paliwal J, Jayas DS. 2008.** Classification of cereal grains using wavelet, morphological, colour, and textural features of non-touching kernel images. *Biosystems Engineering* **99**(3):330–337 DOI [10.1016/j.biosystemseng.2007.11.013](https://doi.org/10.1016/j.biosystemseng.2007.11.013).
- Courbariaux M, Hubara I, Soudry D, El-Yaniv R, Bengio Y. 2016.** Binarized neural networks: training deep neural networks with weights and activations constrained to+ 1 or-1. ArXiv preprint. [arXiv:1602.02830](https://arxiv.org/abs/1602.02830).
- Daquan Z, Hou Q, Chen Y, Feng J, Yan S. 2020.** Rethinking bottleneck structure for efficient mobile network design. In: Vedaldi A, Bischof H, Brox T, Frahm JM, eds. *Computer Vision—ECCV 2020. ECCV 2020. Lecture Notes in Computer Science*, vol. 12348. Cham: Springer DOI [10.1007/978-3-030-58580-8_40](https://doi.org/10.1007/978-3-030-58580-8_40).
- Dolata P, Reiner J. 2018.** Barley variety recognition with viewpoint-aware double-stream convolutional neural networks. In: *2018 federated conference on computer science and information systems (FedCSIS)*. 101–105.
- Guo Y, Yao A, Chen Y. 2016.** Dynamic network surgery for efficient dnns. ArXiv preprint. [arXiv:1608.04493](https://arxiv.org/abs/1608.04493).
- Han K, Guo J, Zhang C, Zhu M. 2018.** Attribute-aware attention model for fine-grained representation learning. In: *Proceedings of the 26th ACM international conference on multimedia, MM '18*. New York, NY, USA: Association for Computing Machinery, 2040–2048.

- Han K, Wang Y, Tian Q, Guo J, Xu C. 2020a.** Ghostnet: more features from cheap operations. In: *2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*. Piscataway: IEEE.
- Han S-Y, Zhou J, Chen Y-H, Zhang Y-F, Tang G-Y, Wang L. 2020b.** Active fault-tolerant control for discrete vehicle active suspension via reduced-order observer. In: *IEEE transactions on systems, man, and cybernetics: systems* DOI 10.1109/TSMC.2020.2964607.
- He K, Zhang X, Ren S, Sun J. 2016.** Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Piscataway: IEEE, 770–778.
- HemaChitra HS, Suguna S. 2018.** Optimized feature extraction and classification technique for indian pulse seed recognition. *International Journal of Computer Engineering & Applications* 12:421–427.
- Hinton G, Vinyals O, Dean J. 2015.** Distilling the knowledge in a neural network. ArXiv preprint. [arXiv:1503.02531](https://arxiv.org/abs/1503.02531).
- Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H. 2017.** Mobilenets: efficient convolutional neural networks for mobile vision applications. ArXiv preprint. [arXiv:1704.04861](https://arxiv.org/abs/1704.04861).
- Huang G, Liu Z, Laurens V, Weinberger KQ. 2017.** Densely connected convolutional networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Piscataway: IEEE, 4700–4708.
- Huang K-Y. 2012.** Detection and classification of areca nuts with machine vision. *Computers & Mathematics with Applications* 64(5):739–746 Advanced Technologies in Computer, Consumer and Control DOI 10.1016/j.camwa.2011.11.041.
- Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K. 2016.** Squeezenet: alexnet-level accuracy with 50x fewer parameters and 0.5 mb model size. ArXiv preprint. [arXiv:1602.07360](https://arxiv.org/abs/1602.07360).
- Kayacan MC, Sofu MM, Cetisli BO. 2016.** Design of an automatic apple sorting system using machine vision. *Computers and Electronics in Agriculture* 127:395–405.
- Krizhevsky A, Sutskever I, Hinton GE. 2012.** Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems* 25:1097–1105.
- Kurtulmuş F. 2021.** Identification of sunflower seeds with deep convolutional neural networks. *Food Measure* 15:1024–1033 DOI 10.1007/s11694-020-00707-7.
- Li H, Kadav A, Durdanovic I, Samet H, Graf HP. 2016.** Pruning filters for efficient convnets. ArXiv preprint. [arXiv:1608.08710](https://arxiv.org/abs/1608.08710).
- Li X, Dai B, Sun H, Li W. 2019.** Corn classification system based on computer vision. *Symmetry* 11(4):591 DOI 10.3390/sym11040591.
- Liang F, Tian Z, Dong M, Cheng S, Sun L, Li H, Chen Y, Zhang G. 2021.** Efficient neural network using pointwise convolution kernels with linear phase constraint. *Neurocomputing* 423:572–579 DOI 10.1016/j.neucom.2020.10.067.
- Lin M, Chen Q, Yan S. 2013.** Network in network. ArXiv preprint. [arXiv:1312.4400](https://arxiv.org/abs/1312.4400).

- Lin M, Ji R, Wang Y, Zhang Y, Zhang B, Tian Y, Shao L. 2020.** Hrank: filter pruning using high-rank feature map. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. Piscataway: IEEE, 1529–1538.
- Liu D, Ning X, Li Z, Yang D, Li H, Gao L. 2015.** Discriminating and elimination of damaged soybean seeds based on image characteristics. *Journal of Stored Products Research* **60**:67–74 DOI [10.1016/j.jspr.2014.10.001](https://doi.org/10.1016/j.jspr.2014.10.001).
- Liu P, Huda MN, Sun L, Yu H. 2020.** A survey on underactuated robotic systems: bio-inspiration, trajectory planning and control. *Mechatronics* **72**:102443 DOI [10.1016/j.mechatronics.2020.102443](https://doi.org/10.1016/j.mechatronics.2020.102443).
- Liu P, Yu H, Cang S. 2019.** Adaptive neural network tracking control for underactuated systems with matched and mismatched disturbances. *Nonlinear Dynamics* **98**(2):1447–1464 DOI [10.1007/s11071-019-05170-8](https://doi.org/10.1007/s11071-019-05170-8).
- Liu Z, Li J, Shen Z, Huang G, Yan S, Zhang C. 2017.** Learning efficient convolutional networks through network slimming. In: *Proceedings of the IEEE international conference on computer vision*. 2736–2744.
- Liu Z, Xu J, Peng X, Xiong R. 2018.** Frequency-domain dynamic pruning for convolutional neural networks. In: *Proceedings of the 32nd international conference on neural information processing systems*. 1051–1061.
- Ma N, Zhang X, Zheng H-T, Sun J. 2018.** Shufflenet v2: practical guidelines for efficient cnn architecture design. In: *Proceedings of the European conference on computer vision (ECCV)*. 116–131.
- Mebatsion HK, Paliwal J, Jayas DS. 2013.** Automatic classification of non-touching cereal grains in digital images using limited morphological and color features. *Computers & Electronics in Agriculture* **90**(90):99–105 DOI [10.1016/j.compag.2012.09.007](https://doi.org/10.1016/j.compag.2012.09.007).
- Mellempudi N, Kundu A, Mudigere D, Das D, Kaul B, Dubey P. 2017.** Ternary neural networks with fine-grained quantization. ArXiv preprint. [arXiv:1705.01462](https://arxiv.org/abs/1705.01462).
- Ni C, Wang D, Vinson R, Holmes M, Tao Y. 2019.** Automatic inspection machine for maize kernels based on deep convolutional neural networks. *Biosystems Engineering* **178**:131–144 DOI [10.1016/j.biosystemseng.2018.11.010](https://doi.org/10.1016/j.biosystemseng.2018.11.010).
- Rehman TU, Mahmud MS, Chang YK, Jin J, Shin J. 2019.** Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. *Computers and Electronics in Agriculture* **156**:585–605 DOI [10.1016/j.compag.2018.12.006](https://doi.org/10.1016/j.compag.2018.12.006).
- Ren S, He K, Girshick R, Sun J. 2015.** Faster r-cnn: towards real-time object detection with region proposal networks. ArXiv preprint. [arXiv:1506.01497](https://arxiv.org/abs/1506.01497).
- Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C. 2018.** Mobilenetv2: inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Piscataway: IEEE, 4510–4520.
- Sankararaman KA, De S, Xu Z, Huang WR, Goldstein T. 2020.** The impact of neural network overparameterization on gradient confusion and stochastic gradient descent. In: *International conference on machine learning*. PMLR, 8469–8479.
- Simonyan K, Zisserman A. 2014.** Very deep convolutional networks for large-scale image recognition. ArXiv preprint. [arXiv:1409.1556](https://arxiv.org/abs/1409.1556).

- Sladojevic S, Arsenovic M, Anderla A, Culibrk D, Stefanovic D. 2016.** Deep neural networks based recognition of plant diseases by leaf image classification. *Computational Intelligence and Neuroscience* **2016**:3289801 DOI [10.1155/2016/3289801](https://doi.org/10.1155/2016/3289801).
- Sun L, Zhao C, Yan Z, Liu P, Duckett T, Stolkin R. 2019.** A novel weakly-supervised approach for rgb-d-based nuclear waste object detection. *IEEE Sensors Journal* **19**(9):3487–3500.
- Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. 2015.** Going deeper with convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Piscataway: IEEE, 1–9.
- Veeramani B, Raymond JW, Chanda P. 2018.** Deepsort: deep convolutional networks for sorting haploid maize seeds. *BMC Bioinformatics* **19**(9):1–9 DOI [10.1186/s12859-017-2006-0](https://doi.org/10.1186/s12859-017-2006-0).
- Wang X, Yu SX. 2020.** Tied block convolution: leaner and better cnns with shared thinner filters. ArXiv preprint. [arXiv:2009.12021](https://arxiv.org/abs/2009.12021).
- Wu S, Wang Z, Shen B, Wang J-H, Dongdong L. 2020.** Human–computer interaction based on machine vision of a smart assembly workbench. *Assembly Automation*.
- Xie S, Girshick R, Dollár P, Tu Z, He K. 2017.** Aggregated residual transformations for deep neural networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Piscataway: IEEE, 1492–1500.
- Zagoruyko S, Komodakis N. 2016.** Paying more attention to attention: Improving the performance of convolutional neural networks via attention transfer. ArXiv preprint. [arXiv:1612.03928](https://arxiv.org/abs/1612.03928).
- Zhang R. 2019.** Making convolutional networks shift-invariant again. In: *International conference on machine learning*. PMLR, 7324–7334.
- Zhang X, Zhou X, Lin M, Sun J. 2018.** Shufflenet: an extremely efficient convolutional neural network for mobile devices. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Piscataway: IEEE, 6848–6856.

Musculoskeletal modeling and humanoid control of robots based on human gait data

Jun Yu¹, Shuaishuai Zhang², Aihui Wang², Wei Li² and Lulu Song²

¹ Zhongyuan-Petersburg Aviation College, Zhongyuan University of Technology, Zhengzhou, China

² School of Electric and Information Engineering, Zhongyuan University of Technology, Zhengzhou, China

ABSTRACT

The emergence of exoskeleton rehabilitation training has brought good news to patients with limb dysfunction. Rehabilitation robots are used to assist patients with limb rehabilitation training and play an essential role in promoting the patient's sports function with limb disease restoring to daily life. In order to improve the rehabilitation treatment, various studies based on human dynamics and motion mechanisms are still being conducted to create more effective rehabilitation training. In this paper, considering the human biological musculoskeletal dynamics model, a humanoid control of robots based on human gait data collected from normal human gait movements with OpenSim is investigated. First, the establishment of the musculoskeletal model in OpenSim, inverse kinematics, and inverse dynamics are introduced. Second, accurate human-like motion analysis on the three-dimensional motion data obtained in these processes is discussed. Finally, a classic PD control method combined with the characteristics of the human motion mechanism is proposed. The method takes the angle values calculated by the inverse kinematics of the musculoskeletal model as a benchmark, then uses MATLAB to verify the simulation of the lower extremity exoskeleton robot. The simulation results show that the flexibility and followability of the method improves the safety and effectiveness of the lower limb rehabilitation exoskeleton robot for rehabilitation training. The value of this paper is also to provide theoretical and data support for the anthropomorphic control of the rehabilitation exoskeleton robot in the future.

Submitted 30 April 2021

Accepted 12 July 2021

Published 9 August 2021

Corresponding author

Jun Yu, 6523@zut.edu.cn

Academic editor

Pengcheng Liu

Additional Information and
Declarations can be found on
page 15

DOI 10.7717/peerj-cs.657

© Copyright
2021 Yu et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Human-Computer Interaction, Robotics

Keywords Human-gait, Musculoskeletal model, OpenSim, Humanoid control, Lower limb rehabilitation exoskeleton, Motion capture

INTRODUCTION

The main causes of limb motor dysfunction in patients include stroke and limb injury (*Sousa et al., 2011*). The sequelae of this injurious seriously affect the quality of life for patients and their families. Worldwide, there are a large number of patients with limb dysfunction caused by various accidents (*Bai et al., 2019*). Meanwhile, the current global population is facing a very serious aging situation (*Fuster, 2017*). The number of patients with limb dysfunction is further increasing with the aging society. The lower limbs play a crucial role in our lives. The lower limb training helps to expand the scope of daily activities of the recovered person and reduce the risk of falling. Therefore, in the rehabilitation treatment of the limbs, the rehabilitation treatment of lower limbs is especially essential. Studies have

shown that patients with limb motor dysfunction will recover their normal walking ability after a certain amount of scientific training at an appropriate time after injury ([Mekki et al., 2018](#); [Chen et al., 2016](#); [LI & JIANG, 2010](#)). However, the rehabilitation training work for patients with limb dysfunction in the later period is heavy, the existing medical staff cannot complete this huge task. The emergence of rehabilitation exoskeleton provides an effective solution to solve these social problems ([Bernhardt et al., 2017](#); [Coleman et al., 2017](#)), fills the gap in the number of medical staff, and brings hope to the majority of patients with limb dysfunction. At present, the exoskeleton robot for lower limb rehabilitation training mostly uses force/position control or trajectory planning method ([Shi et al., 2019](#); [Bernhardt et al., 2005](#)). Although the existing exoskeleton training process is mostly mechanized rigid training, the complexity of the human movement process determines that the exoskeleton is difficult to track the human movement track in the training process. Therefore, the traditional exoskeleton rehabilitation training cannot effectively meet the needs of patients with lower limb motor dysfunction.

In order to solve this problem, current researchers have proposed a musculoskeletal model of human-based on the characteristics of movement mechanisms to improve the effectiveness of rehabilitation training. The human musculoskeletal system is a complex non-linear, multi-redundant system, which is difficult for non-human physiology researchers, then most lower limb rehabilitation exoskeleton robot researchers rarely analyze the real human gait and human musculoskeletal model. The open-source software OpenSim developed by Stanford University brings a feasible solution for non-human physiology researchers. [Seth et al. \(2018\)](#) have jointly developed OpenSim, which can create a musculoskeletal model and then predict new motions through the model and perform motion analysis. OpenSim is the software based on computational modeling and simulation of biomechanical systems ([Seth et al., 2018](#)). Based on OpenSim, [Guo et al. \(2020\)](#) studied the biomechanical characteristics of human lower limbs at different speeds and different weights, performed gait simulation, and proposed joint torque and muscle activation during walking ([Saul et al., 2015](#)). Space circulation characteristics and biomechanical characteristics are the main content of gait analysis research. Researchers employed OpenSim to perform musculoskeletal modeling and analyze the joint kinematics and muscle force characteristics of gait ([Wang et al., 2018](#)). [Cardona & Cena \(2019\)](#) studied and analyzed the biomechanics of the lower limbs of the human body, and estimated the kinematics and dynamics parameters of healthy gait and pathological gait. [Zhou et al. \(2020\)](#) combined the human musculoskeletal model and exoskeleton modeling control, then conducted simulation research on exoskeleton design and control methods with humans in the loop ([Branson et al., 2010](#)). [Humphreys \(2019\)](#) uses OpenSim to test in an environment lacking measurement test data and microgravity to generate predictive kinematics. It is of great significance to study the coupling and synergy between the exoskeleton and humans. [Dembia et al., \(2017\)](#) employed OpenSim to simulate auxiliary equipment and reduce the metabolic cost of weight-bearing walking through simulation; this research will provide help for experimenters in the manufacture of exoskeleton devices. In the field of human body mechanism analysis and research, OpenSim has been widely used. However, most of these studies and applications start from the software itself to

simulate and analyze motion, the results of OpenSim simulation analysis are rarely used for extended applications in the field of rehabilitation exoskeleton.

Therefore, this paper expands the results of OpenSim simulation analysis and applies them to lower extremity rehabilitation exoskeleton robots. Starting from the human body motion mechanism, the human body kinematics analysis is carried out, and a PD control strategy based on real gait and musculoskeletal model is proposed. The schematic diagram of the principle is shown in Fig. 1. In this picture, A represents gait data acquisition, B represents data preprocessing, C represents modeling and analysis, D represents the controller, E represents the robot. First, this paper uses the NOKOV motion capture system and force measurement platform to collect normal human gait data, and preprocess the collected data. Then, the processed gait data was imported into OpenSim, and the musculoskeletal model of the experimental object was established for human kinematics and dynamic analysis, moreover obtained the mechanical characteristics of human motion. Finally, the human motion mechanical characteristics are proposed to control the torque of the lower limb exoskeleton robot based on the PD controller, and the error-free tracking is achieved by adjusting the controller parameters. This method improves the flexibility of the exoskeleton robot movement and meets the anthropomorphic design requirements of rehabilitation training.

The Ethics Committee of the School of Electronic Information, Zhongyuan University of Technology (approval batch number: ZUTSEI202008-001), approved this research protocol, and all participating patients signed an informed consent form.

Analysis of the mechanism of human lower limb movement

Human gait data collection

At present, the motion capture system is divided into five categories according to the principle: mechanical motion capture system ([Wu et al., 2005](#)), acoustic motion capture system, electromagnetic motion capture system ([Guo et al., 2011](#)), inertial motion capture system ([Kim & Nussbaum, 2013](#)) and optical motion capture system ([Kurihara et al., 2002](#); [Guerra Filho, 2005](#); [Kirk, O'Brien & Forsyth, 2005](#)). Among them, the optical motion capture system is divided into two categories: motion capture system based on computer vision (optical non-calibration) and optical motion capture system based on mark point (optical calibration).

This paper selected Nokov 3D infrared passive optical motion capture system with high accuracy and good effect after comparing several existing motion capture systems and combining with the needs for current research topic. In the scene set up by this system, the infrared camera is used to fully cover the experimental scene, infrared light is emitted by the infrared camera array in the process of data collection, and the position information of the reflective Marker points are captured in the experimental scene. In the process of collecting gait data, the experimental subject walks in the experimental scene with affixed Marker points. In order to meet the needs of the research, the experimental platform is equipped with a three-dimensional force measuring platform, which can synchronously collect the three-dimensional ground reaction force during the movement of the experimental object.

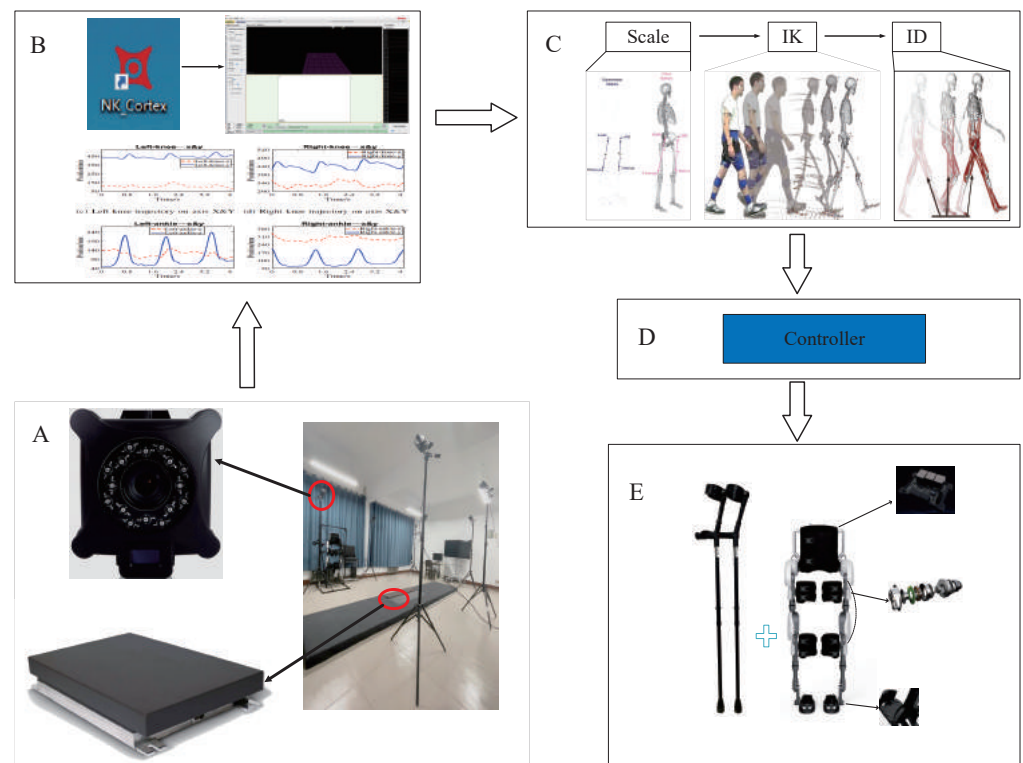


Figure 1 The schematic diagram of the principle.

[Full-size](#) DOI: [10.7717/peerjcs.657/fig-1](https://doi.org/10.7717/peerjcs.657/fig-1)

Before collecting experimental data, the deployment of the experimental platform is also critical. The deployment of the camera position has a fatal impact on the experimental data (Kurihara *et al.*, 2002). In the experiment, the influences of different camera arrangement methods on experimental data were tested. It was found that the best data is obtained by using the approximate circular camera arrangement. This arrangement allows each camera to maximize its utilization value. In the experiment, the cameras scene is arranged around the force measuring platform in an oval shape. The calibration origin is positioned as far as possible in the center of each camera's field of view by adjusting the orientation of the cameras. The experimental collection scene is shown in Fig. 2.

Gait data processing

After the data collection is completed, it is preprocessed to ensure the completeness and accuracy of each frame of data. For missing data points, we had appropriate discarding or interpolation methods for processing. For severely missing data, the entire group will be deleted without applying.

The force was collected by three-dimensional measuring platform that is the force between foot and ground when the experimenter walks. During the process of gait collection, the force in the vertical direction is the most important force, it reflects the phenomenon of overweight and weightlessness during the gait cycle. The three-dimensional force as shown in Fig. 3. It can be found that the force between left foot and right foot with

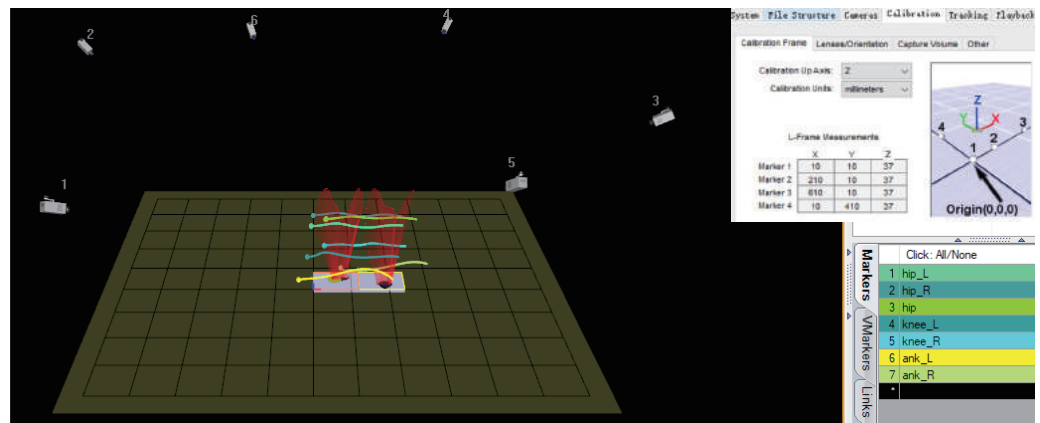


Figure 2 The human gait data collection scene.

[Full-size](#) DOI: 10.7717/peerjcs.657/fig-2

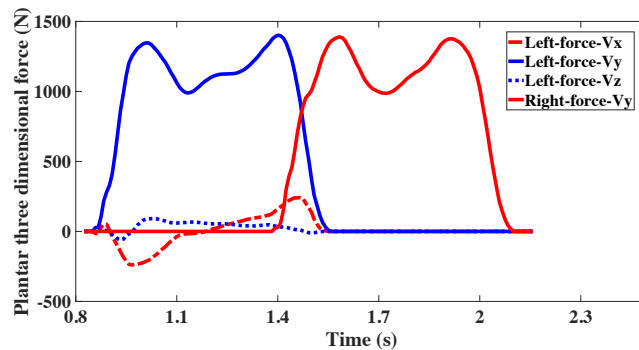


Figure 3 Component of plantar force on coronal, sagittal, and vertical axis.

[Full-size](#) DOI: 10.7717/peerjcs.657/fig-3

the ground is basically symmetrical on same axis. The data of the force platform is zero before the foot contacts it. Next, a small fluctuation is formed in the coronal and sagittal axes at first, then increases to a maximum and gradually decreases to zero. On the vertical axis, it increases rapidly, forming a bimodal curve similar to M, and then rapidly returns to zero too.

The acceleration of the left and right joint has certain symmetry and periodicity, as shown in Fig. 4, it can be seen from the figure that the acceleration from the left leg joint and the right leg joint can basically coincide with each other in the case of 0.5 s of translation. During the walking process of the subject, the acceleration of hip joint points were maintained from 1 m/s^2 to 5 m/s^2 . The acceleration of knee joint points were maintained from 0 m/s^2 to 13 m/s^2 . The acceleration of ankle joint points were maintained from 0 m/s^2 to 25 m/s^2 . It can be easily observed that the acceleration at the ankle joint points is greater than the acceleration at the knee joint points, and the acceleration at the knee joint points is greater than acceleration at the hip joint points. During walking, consistent with experience, the further away from the torso, the acceleration of the nodes greater. Here, a small idea is proposed: based on the acceleration of the joints, a body

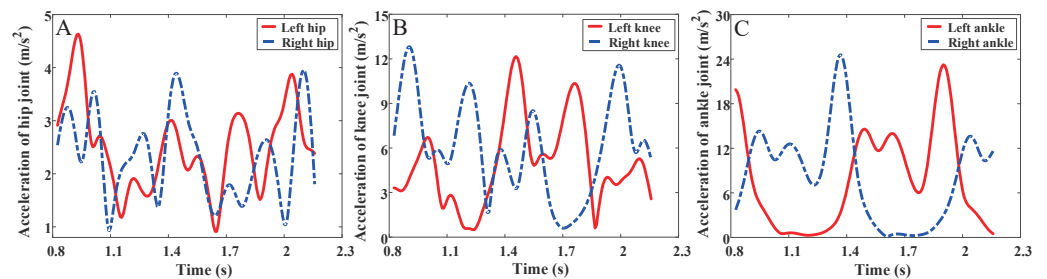


Figure 4 Combined acceleration of joints. (A) The acceleration of hip joint. (B) The acceleration of knee joint. (C) The acceleration of ankle joint.

Full-size DOI: 10.7717/peerjcs.657/fig-4

movement comfort function is designed to evaluate the patients' comfort in the process of lower limb exoskeleton rehabilitation training. This will be a research direction in the next stage of this subject.

Musculoskeletal modeling

In order to build musculoskeletal models and obtain relatively accurate biomechanical information, several common musculoskeletal modeling and computation software on the market, such as LIFEMOD (Huynh et al., 2015), OpenSim (Seth et al., 2011), Anybody (Damsgaard et al., 2006), SIMM software, were compared in the research process. The comparison results are shown in Table 1. It was found that OpenSim meets the needs of this study by comparison. It is an open sources free software developed by Stanford University. OpenSim calculates the motion process based on biomechanical knowledge and combining forward kinematics and inverse kinematics. OpenSim can be applied to human musculoskeletal model development, motion simulation, motion analysis, muscle strength calculation, normal and pathological gait analysis, etc.

The first reason for using OpenSim modeling: rigid exoskeleton rehabilitation robot is a typical multi-input and multi-output complex mechanical system with nonlinear, strong coupling and other uncertain factors. There is a great inaccuracy when modeling the exoskeleton using a traditional linkage model. These inaccuracies are mainly reflected in the following aspects:

1. Mass moment;
2. Inertial matrix;
3. Changes in stiffness and damping (in the process of human–computer interaction);
4. Static friction force of the robot.

The second reason for using OpenSim modeling: compared with the exoskeleton robot, the musculoskeletal system of the human body is a more complex system with multiple redundancy, nonlinear and strong coupling. The most basic way of human movement is to pull the bones around the joints through muscle contraction to achieve the purpose of limb movement. Compared with the traditional connecting rod modeling, the musculoskeletal system modeling is more in line with the movement and texture characteristics of human body, and can better simulate some movements of human body, which is closer to the actual movement characteristics of human body. Using musculoskeletal model for simulation

Table 1 Musculoskeletal modeling software comparison.

Software	Main Features
LifeMOD	<ul style="list-style-type: none"> -Commercial: Yes. -Import Simulink from another: CAD, CATIA, Pro/E, SolidWorks, Unigraphics. -Inverse Kinematics utility. -Inverse Dynamics utility. -Allows simulations with implants.
OpenSim	<ul style="list-style-type: none"> -Commercial: NO (free). -Simulink Export: No native. -Muscle-driven forward dynamic (from data recorded). -Inverse Dynamics utility. -Inverse kinematic simulation. -Allows simulations with implants.
Anybody	<ul style="list-style-type: none"> -Commercial: Yes. -Simulink Export: No native. -Friction forces modeling. -Inverse Dynamics utility. -Allows simulations with implants.
SIMM	<ul style="list-style-type: none"> -Commercial: Yes. -Simulink Export: No native. -Real-time viewing. -Bone deformation modeling. -Inverse kinematics utility.

will get more accurate and reliable simulation results. Considering comprehensively, this paper chooses the musculoskeletal modeling method which is closer to the human body for data analysis and processing.

The model selected in this paper is based on the Gait 2354 model, which is from the OpenSim open-source community. This is a three-dimensional model with 23-degrees of freedom of the human musculoskeletal system. The model embodies the achievements of many predecessors. First, the original model is created by Thelen et al. The model uses Delp et al.'s definition of lower extremity joints ([Delp et al., 1990](#)), Anderson and Pandy's low back joints and anthropometry ([Anderson & Pandy, 1999](#); [Anderson & Pandy, 2001](#)), and Yamaguchi and Zajac's plane knee model ([Yamaguchi & Zajac, 1989a](#)). The Gait2392 model features 92 muscle actuators to represent 76 muscles in the lower extremities and torso. For the Gait2354 model, the number of muscles was reduced by Anderson to improve simulation speed for demonstrations and educational purposes. Seth removed the patella to avoid kinematic constraints; insertions of the quadriceps are handled with moving points in the tibia frame.

Musculoskeletal model scaling

In this paper, the open-source musculoskeletal model was scaled to obtain the exclusive model equivalent to the experimental object. In order to ensure the accuracy of model

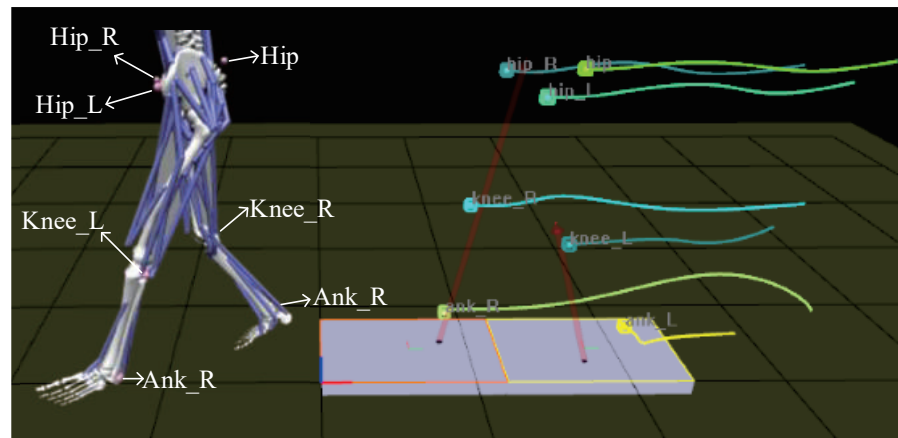


Figure 5 Markers.

Full-size  DOI: [10.7717/peerjcs.657/fig-5](https://doi.org/10.7717/peerjcs.657/fig-5)

scaling, the model was scaled several times. Finally, the accuracy of the model with a scaling error of less than one-thousandth is achieved.

Model scaling allows the open sources model to match our experimental subjects as closely as possible. In the scaling process, static data collected are mainly used in this paper (the experimental data collected while the experimental object is standing still). Before scaling the model, Marker points were added at the appropriate position on the model in accordance with the experimental object. As shown in Fig. 5. Meanwhile, these Marker points were connected to the model bones. In order to ensure the scaled model more accurate, the collected action data are processed in this paper. Through calculation, the left and right width of the pelvis and the length of the left thigh, left calf, right thigh and right calf were obtained (as shown in Table 2), where the mass and length are calculated through Zatsiorsky (Vaughan, Andrews & Hay, 1982) and Harless study (Drillis, Contini & Bluestein, 1964). Finally, the length of these body segments in the model was built.

All the above body segment lengths were obtained from the processing of experimental data, and the measured body segment lengths were averaged. The comparison shows that the data is relatively accurate. In this scaling process, we preserve mass distribution during scale, and change the scale weight of the makers to get a more accurate model.

Inverse Kinematics (IK)

Forward kinematics calculate the final position of the model by giving the initial position, velocity and acceleration of the model. The IK are opposite to the forward dynamics. IK figures out the motion process of the model based on the given position, including the change of physical quantity such as velocity and acceleration in this process.

The IK uses the motion capture data collected in the experiment (walk.trc). And the internal algorithm of the software was used for biological simulation, and the inverse solution was used to calculate the joint angle, pelvis tilt, et al. Among the lower limb movements of the human body, the hip joint movement is most complicated. The hip joint has three degrees of freedom. Therefore, the leg will perform three axial movements on the

Table 2 Subject's physical information.

Body segment	Thigh(L)	Calf(L)	Thigh(R)	Calf(R)
Origin model segment mass (Kg)	9.3014	3.7075	9.3014	3.7075
Scaled model segment mass (Kg)	8.0441	3.2063	8.0441	3.2063
Body mass by Harless study (Kg)	7.67	2.925	7.67	2.925
Segment length by experiment (mm)	458.17	395.12	456.36	394.22
Length by Harless study (mm)	438.48	394.98	438.48	394.98

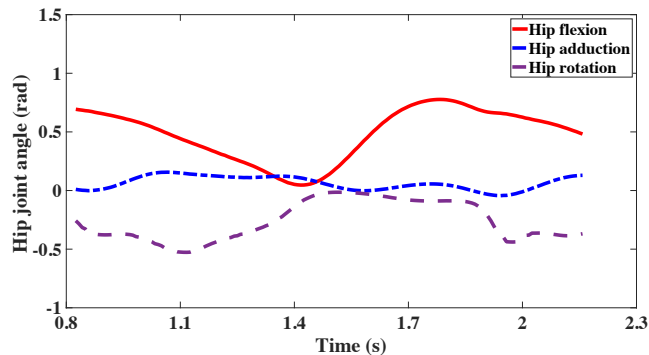


Figure 6 The adduction, flexion and rotation of hip joint.

Full-size DOI: [10.7717/peerjcs.657/fig-6](https://doi.org/10.7717/peerjcs.657/fig-6)

hip joint, including flexion and extension of the hip joint on the sagittal plane, adduction and abduction on the coronal plane, and internal rotation and external rotation direction of the thigh. During the gait cycle, the hip joint angles change as shown in Fig. 6. The hip joint of flexion is the movement in the sagittal plane, and its range of change is stable at $0 \sim 0.8$ rad. This movement change is the leading and effective movement of the hip joint during the gait, and the movement in this direction will drive the body to move forward. The hip joint of adduction changes smoothly in the gait, with only a slight fluctuation. The hip joint of rotation includes internal rotation and external rotation of the thigh during the gait. It can be seen that the data in this part has strong characteristics, and the range of variation is stable at $-0.6 \sim 0$ rad. The hip joint of adduction and the hip joint of rotation have more personal characteristics related to personal habits and leg health conditions. It is also a key factor in judging whether the gait is abnormal. Through IK, the collected motion capture data can be matched with the calibration data of the experimental object, and the motion simulation process of the model can be optimized. The IK tool calculates universal coordinate values for each time step (frame) of the movement. Then, the model is positioned in the “best match” pose of the experimental marker time step. In other words, mark points in the collected motion process are matched with the motion capture data, and the weighted square error of the mark points and motion capture data is minimized. The law of weighted least squares problem during IK solved by the function:

$$\min_q \left[\sum_{i \in \text{markers}} w_i \|x_i^{\text{exp}} - x_i(q)\|^2 + \sum_{j \in \text{unprescribed-coords}} \omega_j (q_j^{\text{exp}} - q_j)^2 \right] \quad (1)$$

Where, q is the vector of the generalized coordinates solved, X_i^{exp} is the experimental location of mark i , $X_i(q)$ is the position of the corresponding mark points in the model (depending on the coordinate value), q_j^{exp} is the experimental value of coordinate j , and set their experimental values to the specified coordinates.

The comparison between the knee angle analyzed by Cortex data acquisition software and the knee angle analyzed by OpenSim musculoskeletal simulation software is shown in Fig. 7. The data in the figures represents the gait information of 1.5 cycles. The overall trend of knee joint angle obtained by two methods are similar and can be seen from the figures. However, there are still some differences. It can be seen that the variation range of knee joint angle obtained by using musculoskeletal simulation software OpenSim is larger and the variation trend is more stable. The reason of this phenomenon maybe is the Cortex software get the angles just by simple calculating with the collected position data. OpenSim combines the characteristics of musculoskeletal model in the calculation of joint angles, so the joint angles obtained by OpenSim are better than Cortex.

Inverse dynamics (ID)

The ID problem refers to: given the position q , velocity \dot{q} and acceleration \ddot{q} of each joint of the robot at a certain moment, calculate the driving force (include: states or motion) imposed on each joint at this time. The ID can be solved by the Newton-Euler equation or the Lagrange equation.

Dynamics is the study of motion, the forces and torques that cause motion. The purpose of ID is to estimate the forces and torques required to produce a particular motion, and the results also used to predict how muscles contribute to motion. In order to calculate the forces and moments, the system's equations of motion need to be solved iteratively. The motion equations are derived from the motion description and the mass property of the model. In the solution process, the IK is employed to calculate the joint angle and the ground reaction force during the experiment. And combined with the dynamic equilibrium conditions and boundary conditions, the net forces and moments at each joint are obtained.

When solving the ID problem, the data of motion and force measuring platform are needed to ensure that the number of equations of motion more than unknowns (degrees of freedom), which turns the problem into a statically indeterminate problem. The error of experimental motion data and the inaccuracy of the model will lead to Newton's second law $F=m \cdot a$ invalid. In order to solve this dynamic discontinuity problem, residual forces and torques are introduced into a specific section of the model. The following equation is established, which relates the ground reaction moment to the residual moment. Where, \vec{F}_{exp} is the ground reaction moment and $\vec{F}_{\text{residual}}$ is the residual moment.

$$\vec{F}_{\text{exp}} + \vec{F}_{\text{residual}} = m \cdot \vec{a} \quad (2)$$

Inverse_Dynamics.sto: is generated by inverse kinetic operations, including time series, net joint moments of each bone joint, etc.

The net joints acting torque of the hip joints in three motion states are shown in Fig. 8. The net joint acting moment of hip joint is same to the angle of the hip joint,

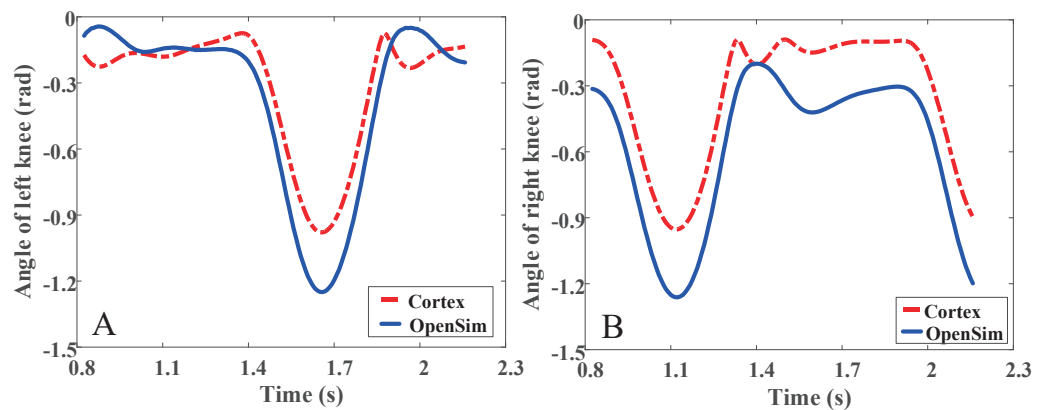


Figure 7 The angle of knee with Cortex & OpenSim. (A) The angle of left knee. (B) The angle of right knee.

Full-size [DOI: 10.7717/peerjcs.657/fig-7](https://doi.org/10.7717/peerjcs.657/fig-7)

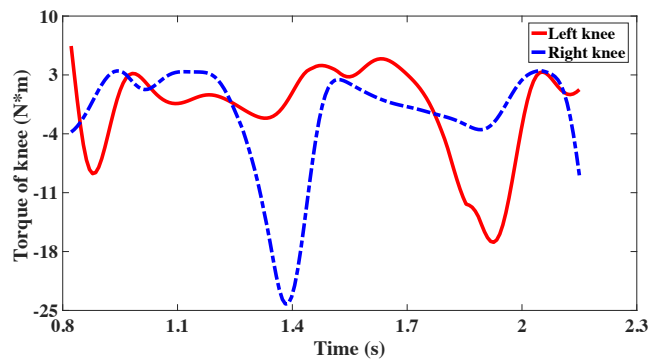


Figure 8 The torque of knee joint from OpenSim.

Full-size [DOI: 10.7717/peerjcs.657/fig-8](https://doi.org/10.7717/peerjcs.657/fig-8)

including adduction, flexion and rotation. Torque drives the body segments, therefore corresponds to the joint angles of the hip joint. The net joint torque of the knee joint is the joint torque connecting the thigh and the calf, which is significant in the research of exoskeleton rehabilitation robots of lower limbs.

CONTROL SYSTEM DESIGN

PD controller

The PD controller is one of the most widely used and effective methods in the field of robotics. As we all know, a PD controller is sufficient to stabilize any kind of rigid manipulator near the reference position. In particular, even when the inertia and friction parameters of the robot are unknown, it can be guaranteed to be asymptotically stable. Under the premise of ignoring friction and other disturbances, *Kelly (1993)* developed a PD controller with adaptive desired gravity compensation and demonstrated the closed-loop global asymptotically stability that obviates LaSalle's theorem. In order for the robot to walk like a human, the PD controller is employed to accurately control the robot's

posture and gait, and the stability of the robot is maintained through the feedback system. [Putri & Machbub \(2018\)](#) designed a PD controller with Center of Mass (COM) as system feedback, and verified the effectiveness of PD controller on the Bioloid GP under uneven environments with many obstacles. [Hu, Wang & Wu, \(2021\)](#) proposed a robust adaptive PD-like control based on healthy human gait data. [Zhou et al. \(2021\)](#) proposed a trajectory deformation algorithm and chose a PD position controller to ensure the trajectory tracking effect. [Ali et al., \(2018\)](#) used a fuzzy PID-based position control method in the design of the upper limb rehabilitation robot system to solved the robot's precise position/force control under the imprecise model. [Han, Wang & Tian \(2020\)](#) used intelligent PD controllers for the motion control of the lower extremity rehabilitation exoskeleton. This method uses a linear state observer to compensate for the control input, solves the inaccuracy of the exoskeleton robot model and the interaction between the human and the exoskeleton. In reality, the robot is a multi-degree-of-freedom, mutually coupled nonlinear system, the performance of robot system depends highly on the availability of high quality differential signal based on the non-continuous measured position signal.

Dynamics

Fourier is a rigid body robot, in the absence of friction and other disturbances, the dynamics of a serial n-link rigid robot can be written as:

$$M(q)\ddot{q} + D(q, \dot{q})\dot{q} + G(q) = \tau + \tau_{OpenSim}. \quad (3)$$

We simplify Fourier to a 2-link rigid robot, where, q is the 2×1 vector of joint angle, τ is the 2×1 vector of joint torques, $\tau_{OpenSim}$ is the 2×1 vector of joint torques from the OpenSim software, $M(q)$ is the 2×2 symmetric positive definite manipulator inertia matrix, $D(q, \dot{q})$ is the 2×1 coriolis force and centrifugal force matrix, $G(q)$ is the 2×1 gravity matrix. In this paper, $\tau_{OpenSim}$ as a reference torque input to the controller, to realize the anthropomorphic torque output of the exoskeleton robot.

The PD control law is given as flowing:

$$\tau = K_d \dot{e} + K_p e - \tau_{OpenSim} \quad (4)$$

where, K_d and K_p are the 2×2 symmetric positive definite matrices, $e = q_d - q$, q_d is the desired joint angle. The $\tau_{OpenSim}$ is a bounded matrix.

$$T_{\min} \leq \|\tau_{OpenSim}\| \leq T_{\max}.$$

Eqs. (3) and (4) imply.

$$M(q)(\ddot{q}_d - \ddot{q}) + D(q, \dot{q})(\dot{q}_d - \dot{q}) + K_d \dot{e} + K_p e = 0. \quad (5)$$

Then the $M(q)\ddot{e} + D(q, \dot{q})\dot{e} + K_p e = -K_d \dot{e}$ is obtained. Considering the candidate Lyapunov function,

$$V = \frac{1}{2} \dot{e}^T M(q) \dot{e} + \frac{1}{2} e^T K_p e \quad (6)$$

Eq. (6) is position definite, and the time derivative of the function,

$$\dot{V} = e^T M \ddot{e} + \frac{1}{2} e^T \dot{M}(q) \dot{e} + e^T K_p e. \quad (7)$$

There is an oblique symmetry feature: $\dot{M}(q) - 2D(q, \dot{q}) = 0$. Substituting the condition into Eq. (5), get

$$\dot{V} = e^T (M \ddot{e} + D \dot{e} + K_p e) = -e^T K_d \dot{e} \leq 0. \quad (8)$$

Then, the stability of design control system can be guaranteed.

Results

The simulation is carried out in MATLAB-Simulink, and the results are shown in Figs. 9 and 10. Figures 9A, 9C and 10A are the joints angle tracking and tracking error of OpenSim combined with PD controller, respectively. Figures 9B, 9D and 10B are the joints angle tracking and tracking error of the PD controller, respectively. Comparing Figs. 9A and 9B, it can be found that the PD controller combined with OpenSim has a better tracking effect, after the tracking, error-free tracking can be achieved, and the controller that only uses the PD control can clearly found that there is still an error in the peak position of the gait angle in the later stage of the tracking. In the method proposed in this paper, the joint torque from OpenSim plays a good role in compensating for the control of the controller. Comparing Figs. 10A and 10B, we can be found that the PD controller combined with OpenSim has a better tracking effect, the tracking error is smoother with only small fluctuations, and the stability is higher. The patient can enter the rehabilitation training comfortably, and achieve the safety and comfort of the rehabilitation training.

The parameter of PD controller, such as: K_p , K_d . $K_p = \text{diag}(150, 150)$, $K_d = \text{diag}(120, 120)$. The performance indexes of the two controllers designed are shown in Table 3.

The RMSE, ISE, and ITSE indicated that the two controllers are almost the same, suggesting that they have faster response speed and smaller oscillation. In terms of the properties of IAE and ITAE, the controller with OpenSim output torque is slightly better than the PD controller's transient response and the transient response oscillation is smaller.

CONCLUSIONS

The collected real human gait data combines with the human musculoskeletal model in this paper, then obtains human motion characteristics by inverse kinematics analysis. These motion characteristics were employed to design the controller and verify it in Simulink. The simulation results show that this method is more flexible and anthropomorphic in the exoskeleton robot control.

Research innovation points summary shown as follows:

(1) The human biological musculoskeletal dynamics model was identified using OpenSim and the real human gait of the experimental data source. Combining the human lower limb movement characteristics extracted from the gait data with the musculoskeletal model established by OpenSim, the musculoskeletal model is based on the physiological

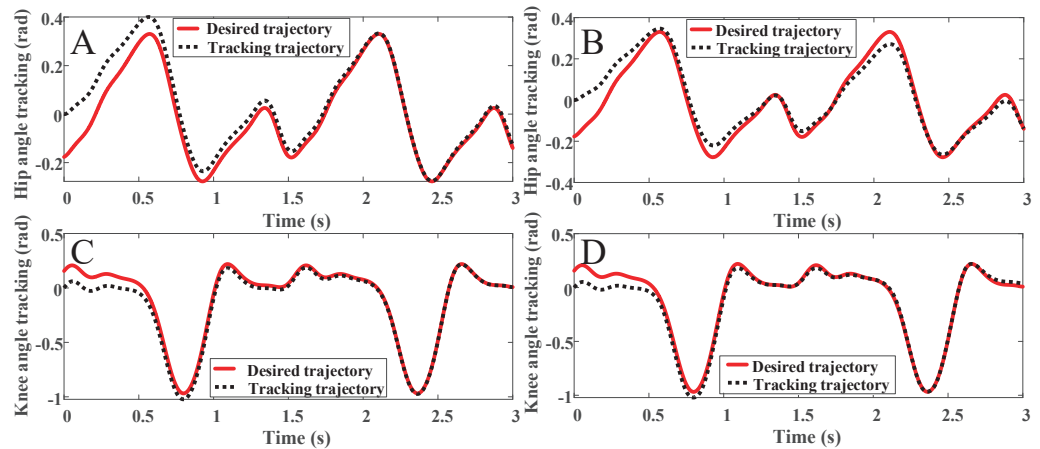


Figure 9 (A) The hip angle track with OpenSim. (B) The hip angle track with link rod model. (C) The knee angle track with OpenSim. (D) The knee angle track with link rod model.

[Full-size](#) [DOI: 10.7717/peerjcs.657/fig-9](https://doi.org/10.7717/peerjcs.657/fig-9)

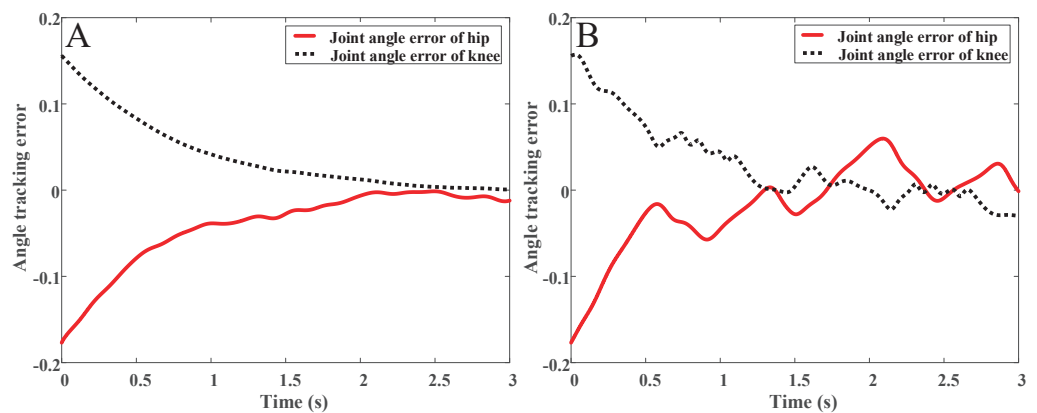


Figure 10 (A) The angle track error with OpenSim. (B) The angle track error with link rod model.

[Full-size](#) [DOI: 10.7717/peerjcs.657/fig-10](https://doi.org/10.7717/peerjcs.657/fig-10)

Table 3 The performance indexes of the two controllers.

ERROR	OpenSim with PD controller		PD controller only	
	Hip error	Knee error	Hip error	Knee error
RMSE	1.7×10^{-4}	1.6×10^{-4}	1.4×10^{-4}	1.5×10^{-4}
ISE	0.010	0.009	0.008	0.009
ITSE	0.003	0.003	0.004	0.003
IAE	0.1219	0.1164	0.1105	0.1100
ITAE	0.0845	0.0790	0.1076	0.0813

characteristics of the human body (muscle and tendon characteristics, skeletal and tendon connection structure, nervous system, etc.). This method will obtain more accurate gait kinematics and gait dynamics data. Compared with the connecting rod model, this model

has a better human-like design, so the precision and accuracy of the model are better than ever.

(2) The humanoid control of robots based on human gait data of normal human gait movements was discussed. The PD controller design and the simulation results based on experimental were analyzed. Experimental results have shown that the control strategy based on OpenSim and PD controller is more in line with the characteristics of human kinematics and physiology. Compared with only using the PD controller, this method has better trajectory tracking effect, faster adjustment time, and more comfortable patient rehabilitation experience.

ACKNOWLEDGEMENTS

Of course, we also sincerely thank the subjects who provided gait data for this paper. They assisted us in completing the work of gait data collection from October 28th to 30th, 2020. With their help, we completed this paper.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the Henan Province Science and Technology R&D projects (No. 182102410056, No. 202102210097, No. 202102210135), the National Natural Science Foundation (No. 62073297, No. U1813201), and the Science and Technology Planning Project of China National Textile and Apparel Council (No. 2019060). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

Henan Province Science and Technology R&D projects: 182102410056, 202102210097, 202102210135.

National Natural Science Foundation: 62073297, U1813201.

Science and Technology Planning Project of China National Textile and Apparel Council: 2019060.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Jun Yu and Aihui Wang conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Shuaishuai Zhang performed the experiments, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Wei Li performed the experiments, prepared figures and/or tables, and approved the final draft.

- Lulu Song analyzed the data, prepared figures and/or tables, and approved the final draft.

Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

The Zhongyuan University of Technology granted Ethical approval to carry out the study within its facilities (ZUTSEI202008-001).

Data Availability

The following information was supplied regarding data availability:

The codes and the raw data are available in the [Supplemental Files](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.657#supplemental-information>.

REFERENCES

- Ali A, Ahmed SF, Kadir KA, Joyo MK, Yarooq RS. 2018.** Fuzzy PID controller for upper limb rehabilitation robotic system. In: *2018 IEEE international conference on innovative research and development (ICIRD)*. Piscataway: IEEE, 1–5.
- Anderson FC, Pandy MG. 1999.** A dynamic optimization solution for vertical jumping in three dimensions. *Computer Methods in Biomechanics and Biomedical Engineering* 2(3):201–231 DOI 10.1080/10255849908907988.
- Anderson FC, Pandy MG. 2001.** Dynamic optimization of human walking. *Journal of Biomechanical Engineering* 123(5):381–390 DOI 10.1115/1.1392310.
- Bai X, Guo Z, He L, Ren L, McClure MA, Mu Q. 2019.** Different therapeutic effects of transcranial direct current stimulation on upper and lower limb recovery of stroke patients with motor dysfunction: a meta-analysis. *Neural Plasticity* 2019.
- Bernhardt J, Godecke E, Johnson L, Langhorne P. 2017.** Early rehabilitation after stroke. *Current Opinion in Neurology* 30(1):48–54 DOI 10.1097/WCO.0000000000000404.
- Bernhardt M, Frey M, Colombo G, Riener R. 2005.** Hybrid force-position control yields cooperative behaviour of the rehabilitation robot LOKOMAT. In: *9th international conference on rehabilitation robotics, ICORR 2005*, 536–539.
- Branson S, Wah C, Schroff F, Babenko B, Welinder P, Perona P, Belongie S. 2010.** Visual recognition with humans in the loop. In: Daniilidis K, Maragos P, Paragios N, eds. *Computer Vision – ECCV 2010. ECCV 2010. Lecture Notes in Computer Science, vol 6314*. Berlin, Heidelberg: Springer, 438–451 DOI 10.1007/978-3-642-15561-1_32.
- Cardona M, Cena CEG. 2019.** Biomechanical analysis of the lower limb: a full-body musculoskeletal model for muscle-driven simulation. *IEEE Access* 7:92709–92723 DOI 10.1109/ACCESS.2019.2927515.
- Chen B, Ma H, Qin LY, Gao F, Chan KM, Law SW, Qin L, Liao WH. 2016.** Recent developments and challenges of lower extremity exoskeletons. *Journal of Orthopaedic Translation* 5:26–37 DOI 10.1016/j.jot.2015.09.007.

- Coleman ER, Moudgal R, Lang K, Hyacinth HI, Awosika OO, Kissela BM, Feng W. 2017. Early rehabilitation after stroke: a narrative review. *Current Atherosclerosis Reports* 19(12):1–12 DOI 10.1007/s11883-017-0636-3.
- Damsgaard M, Rasmussen J, Christensen ST, Surma E, De Zee M. 2006. Analysis of musculoskeletal systems in the anybody modeling system. *Simulation Modelling Practice and Theory* 14(8):1100–1111 DOI 10.1016/j.simpat.2006.09.001.
- Delp SL, Loan JP, Hoy MG, Zajac FE, Topp EL, Rosen JM. 1990. An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. *IEEE Transactions on Biomedical Engineering* 37(8):757–767 DOI 10.1109/10.102791.
- Dembia CL, Silder A, Uchida TK, Hicks JL, Delp SL. 2017. Simulating ideal assistive devices to reduce the metabolic cost of walking with heavy loads. *PLOS ONE* 12(7):e0180320 DOI 10.1371/journal.pone.0180320.
- Drillis R, Contini R, Bluestein M. 1964. Body segment parameters. *Artificial Limbs* 8:44–66.
- Fuster V. 2017. *Changing demographics: a new approach to global health care due to the aging population*. Washington, D.C.: American College of Cardiology Foundation.
- Guerra Filho G. 2005. Optical motion capture: theory and implementation. *RITA* 12(2):61–90.
- Guo C, He YM, Sun ZY, Wang ZH. 2020. Research of biomechanical characteristics of human lower limb walking using the OpenSim technology. *Mechanical Science and Technology for Aerospace Engineering*. Epub ahead of print 2020 2 November DOI 10.13433/j.cnki.1003-8728.20200225.
- Guo LY, Yang CC, Yang CH, Hou YY, Chang JJ, Wu WL. 2011. The feasibility of using electromagnetic motion capture system to measure primary and coupled movements of cervical spine. *Journal of Medical and Biological Engineering* 31(4):245–254 DOI 10.5405/jmbe.721.
- Han S, Wang H, Tian Y. 2020. A linear discrete-time extended state observer-based intelligent PD controller for a 12 DOFs lower limb exoskeleton LLE-RePA. *Mechanical Systems and Signal Processing* 138:106547–106561 DOI 10.1016/j.ymssp.2019.106547.
- Hu N, Wang A, Wu Y. 2021. Robust adaptive PD-like control of lower limb rehabilitation robot based on human movement data. *PeerJ Computer Science* 7:e394 DOI 10.7717/peerj-cs.394.
- Humphreys B. 2019. Predictive kinematic analysis methodologies for use with open-sim. Available at <https://ntrs.nasa.gov/api/citations/20190031761/downloads/20190031761.pdf?attachment=true>.
- Huynh K, Gibson I, Jagdish B, Lu W. 2015. Development and validation of a discretised multi-body spine model in LifeMOD for biodynamic behaviour simulation. *Computer Methods in Biomechanics and Biomedical Engineering* 18(2):175–184 DOI 10.1080/10255842.2013.786049.
- Kelly R. 1993. Comments on AdaptivePD controller for robot manipulators. *IEEE Transactions on Robotics And Automation* 9:117–119 DOI 10.1109/70.210803.
- Kim S, Nussbaum MA. 2013. Performance evaluation of a wearable inertial motion capture system for capturing physical exposures during manual material handling tasks. *Ergonomics* 56(2):314–326 DOI 10.1080/00140139.2012.742932.

- Kirk AG, O'Brien JF, Forsyth DA. 2005.** Skeletal parameter estimation from optical motion capture data. In: *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*. Piscataway: IEEE, 782–788 DOI [10.1109/CVPR.2005.326](https://doi.org/10.1109/CVPR.2005.326).
- Kurihara K, Hoshino S, Yamane K, Nakamura Y. 2002.** Optical motion capture system with pan-tilt camera tracking and real time data processing. In: *Proceedings 2002 IEEE international conference on robotics and automation (Cat. No. 02CH37292)*, 1241–1248 DOI [10.1109/ROBOT.2002.1014713](https://doi.org/10.1109/ROBOT.2002.1014713).
- LI H, Jiang YF. 2010.** Research progress on early rehabilitation nursing for stroke patients. *Chinese Journal of Nursing* 2(3):93–95.
- Mekki M, Delgado AD, Fry A, Putrino D, Huang V. 2018.** Robotic rehabilitation and spinal cord injury: a narrative review. *Neurotherapeutics* 15(3):604–617 DOI [10.1007/s13311-018-0642-3](https://doi.org/10.1007/s13311-018-0642-3).
- Putri DIH, Machbub C. 2018.** Gait controllers on humanoid robot using kalman filter and PD controller. In: *2018 15th international conference on control, automation, robotics and vision*, 36–41.
- Saul KR, Hu X, Goehler CM, Vidt ME, Daly M, Velisar A, Murray WM. 2015.** Benchmarking of dynamic simulation predictions in two software platforms using an upper limb musculoskeletal model. *Computer Methods in Biomechanics and Biomedical Engineering* 18(13):1445–1458 DOI [10.1080/10255842.2014.916698](https://doi.org/10.1080/10255842.2014.916698).
- Seth A, Hicks JL, Uchida TK, Habib A, Dembia CL, Dunne JJ, Ong CF, DeMers MS, Rajagopal A, Millard M, Hamner SR, Arnold EM, Yong JR, Lakshmikanth SK, Sherman MA, Ku JP, Delp SL. 2018.** OpenSim: simulating musculoskeletal dynamics and neuromuscular control to study human and animal movement. *PLOS Computational Biology* 14(7):e1006223 DOI [10.1371/journal.pcbi.1006223](https://doi.org/10.1371/journal.pcbi.1006223).
- Seth A, Sherman M, Reinbolt JA, Delp SL. 2011.** OpenSim: a musculoskeletal modeling and simulation framework for in silico investigations and exchange. *Procedia Iutam* 2:212–232 DOI [10.1016/j.piutam.2011.04.021](https://doi.org/10.1016/j.piutam.2011.04.021).
- Shi D, Zhang W, Zhang W, Ding X. 2019.** A review on lower limb rehabilitation exoskeleton robots. *Chinese Journal of Mechanical Engineering* 32(1):1–11 DOI [10.1186/s10033-018-0313-7](https://doi.org/10.1186/s10033-018-0313-7).
- Sousa CO, Barela JA, Prado Medeiros CL, Salvini TF, Barela AM. 2011.** Gait training with partial body weight support during overground walking for individuals with chronic stroke: a pilot study. *Journal of Neuroengineering and Rehabilitation* 8(1):1–8 DOI [10.1186/1743-0003-8-1](https://doi.org/10.1186/1743-0003-8-1).
- Vaughan C, Andrews J, Hay J. 1982.** Selection of body segment parameters by optimization methods. *Journal of Biomechanical Engineering* 104(7):38–44 DOI [10.1115/1.3138301](https://doi.org/10.1115/1.3138301).
- Wang Y, Li X, Huang P, Li G., Fang P. 2018.** An analysis of biomechanical characteristics of gait based on the musculoskeletal model. In: *2018 IEEE international conference on Cyborg and Bionic Systems (CBS)*. Piscataway: IEEE, 151–154.
- Wu T, McGinley J, Duffy VG, Liu L. 2005.** Application and validation of a mechanical motion capture-based industrial ergonomics assessment system. Technical report, SAE technical paper.

- Yamaguchi GT, Zajac FE. 1989.** A planar model of the knee joint to characterize the knee extensor mechanism. *Journal of Biomechanics* **22**(1):1–10
DOI [10.1016/0021-9290\(89\)90179-6](https://doi.org/10.1016/0021-9290(89)90179-6).
- Zhou X. 2020.** Predictive human-in-the-loop simulations for assistive exoskeletons. In: *2020 international design engineering technical conferences and computers and information in engineering conference*. DOI [10.1115/DETC2020-22668](https://doi.org/10.1115/DETC2020-22668).
- Zhou J, Li Z, Li X, Wang X, Song R. 2021.** Human-robot cooperation control based on trajectory deformation algorithm for a lower limb rehabilitation robot. *IEEE/ASME Transactions on Mechatronics* DOI [10.1109/TMECH.2021.3053562](https://doi.org/10.1109/TMECH.2021.3053562).

Service humanoid robotics: a novel interactive system based on bionic-companionship framework

Jiaji Yang¹, Esyin Chew¹ and Pengcheng Liu²

¹ Cardiff School of Technologies, Cardiff Metropolitan University, Cardiff, Cardiff, United Kingdom

² The Department of Computer Science, University of York, York, United Kingdom

ABSTRACT

At present, industrial robotics focuses more on motion control and vision, whereas humanoid service robotics (HSRs) are increasingly being investigated and researched in the field of speech interaction. The problem and quality of human-robot interaction (HRI) has become a widely debated topic in academia. Especially when HSRs are applied in the hospitality industry, some researchers believe that the current HRI model is not well adapted to the complex social environment. HSRs generally lack the ability to accurately recognize human intentions and understand social scenarios. This study proposes a novel interactive framework suitable for HSRs. The proposed framework is grounded on the novel integration of Trevarthen's (2001) companionship theory and neural image captioning (NIC) generation algorithm. By integrating image-to-natural interactivity generation and communicating with the environment to better interact with the stakeholder, thereby changing from interaction to a bionic-companionship. Compared to previous research a novel interactive system is developed based on the bionic-companionship framework. The humanoid service robot was integrated with the system to conduct preliminary tests. The results show that the interactive system based on the bionic-companionship framework can help the service humanoid robot to effectively respond to changes in the interactive environment, for example give different responses to the same character in different scenes.

Submitted 19 April 2021
Accepted 22 July 2021
Published 13 August 2021

Corresponding author
Jiaji Yang, JYang@cardiffmet.ac.uk

Academic editor
Qichun Zhang

Additional Information and
Declarations can be found on
page 17

DOI 10.7717/peerj-cs.674

© Copyright
2021 Yang et al.

Distributed under
Creative Commons CC-BY 4.0

Subjects Human-Computer Interaction, Artificial Intelligence, Robotics

Keywords Humanoid robotics, Human-robot interaction, Social robotics

INTRODUCTION

Humanoid service robots (HSRs) have seen a sharp rise in adoption recently and are seen as one of the major technologies that will drive the service industries in the next decade (Harris, Kimson & Schwedel, 2018). An increasing number of researchers are committed to investigating HSRs to help humans complete repetitive or high-risk service and interactive tasks such as serving patients with infectious diseases, delivering meals and so on. Delivery robots, concierge robots, and chat robots have been increasingly used by travel and hospitality companies (Ivanov, 2019). Although the contribution of these achievements mainly comes from the rapid development of robotics engineering, Ivanov et al. (2019) indicated that future research focus will gradually shift from robotics

OPEN ACCESS

engineering to human-robot interaction (HRI), thus opening up interdisciplinary research directions for researchers.

In the early days, [Fong, Thorpe & Baur \(2003\)](#) proposed that in order to make robots perform better, the robot needs to be able to use human skills (perception, cognition, etc.) and benefit from human advice and expertise. This means that robots that rely solely on self-determination have limitations in performing tasks. The authors further propose that the collaborative work between humans and robots will be able to break this constraint, and research on human-robot interaction has begun to emerge. [Fong, Thorpe & Baur \(2003\)](#) believe that to build a collaborative control system and complete human-robot interaction, four key problems must be solved. (1) The robot must be able to detect limitations (what can be done and what humans can do), determine whether to seek help, and identify when it needs to be resolved. (2) The robot must be self-reliant and secure. (3) The system must support dialog. That is, robots and humans need to be able to communicate with each other effectively. However, dialog is restricted at present. Through collaborative control, dialog should be two-way and require a richer vocabulary. (4) The system must be adaptive. Although most of the current humanoid service robots already support dialog and can complete simple interactive tasks, as propounded in the research, such dialog in the present time remains limited and “inhuman.” In the process of interacting with robots, humans always determine the state of the robot (the position of the robot or the action the robot is doing) through vision, and then communicate with the robot through a dialog system. However, HSR cannot perform this yet as they do not seem to fully satisfy the two-way nature of dialog. Therefore, this research responds to the current gap and attempts to differ from the current HRI research. This research attempts to introduce deep learning into the existing dialog system of HSR, thus advancing the field.

With the continuous development of humanoid robots, more and more humanoid robots are used in the service industry, especially the hospitality industry. Human-Robot Interaction (HRI) has become a hot potato by more and more researchers ([Yang & Chew, 2020](#)). However, with the deepening of research, some researchers found that when humans interact with humanoid service robots (HSRs), humans hope that HSRs should have the ability and interest to interact with the dynamic thoughts and enthusiasm of the partner's relationship, and can recognize the environment, blended with what others think is meaningful and the emotions to express sympathy ([Yang & Chew, 2020](#)). This coincides with [Trevathan's \(2001\)](#) companionship theory, so the concept of human robot companion (HRC) was proposed this research. The earlier concept of the robot companion is mentioned by [Dautenhahn et al. \(2005\)](#): HSRs need to have a high degree of awareness and sensitivity to social environment. Through the review of the above literatures, it is proposed to establish an interactive and companion framework for HSRs using deep learning and neural image caption generation, thus advance the current field of HSRs to tackle with bionic-interactive tasks of the service industry and further evolve from conventional HRI to Human and Robot Companion (HRC) (See [Table 1](#)).

This study proposes that the introduction of visual data into the current HRI model of HSRs enables HSRs to have a high level of sensitivity to the social dynamic environment while interacting with humans, thereby enhancing the current HRI model to HRC. With

Table 1 Scenario based comparison for HRI and HRC.

Scenario:	HRI	HRC
<p><i>Scenario 1: Hospitality:</i></p> <p>The enhancement from HRI to HRC:</p> <p>(1) Compared with HRI, robots in HRC can recognize the environment (luggage) and changes in customers' appearance (red shirt), which is in line with the proposed concept of companion should become dynamic in the theory of companionship.</p> <p>(2) More enthusiastic and bionic interaction capabilities (automatically detect whether they are regular customers, and greet enthusiastically).</p> <p>(3) The robot in HRC remember the customers' past orders and provide meal recommendation for well-being.</p>	<p>When you enter a hotel, you see a reception area dominated by robots. When you approach the reception area, the HSR will say "Welcome to Hotel XYZ, please follow the instructions to check-in on my display screen". After completing the check-in, the robot will tell you the room number and issue you a room card, you go to your room, change a suit and prepare to go downstairs to eat. When you go back to the reception, the robot says 'welcome, please follow the instructions to place an order on my display'. You choose a few dishes that look good on the screen of the robot, but when the food comes up you don't seem to be satisfied with the taste...</p>	<p>When you enter a hotel, you can a reception area thoughtfully served by robots. The robots also see you and wave to you, 'Welcome Jack, you have a nice luggage, I can help you to check-in. What else can I do for you?' After completing the check-in, the robot will tell you the room number and issues you a room card. You go to your room and change to a red shirt to go downstairs to eat. When you go back to the reception, the robot says, "Welcome Jack, you wear nice red shirt, what can I do for you?" You choose a few dishes that look good on the robot's screen, but the robot tells you that 'According to your past order and diet preferences, these meals may not be suitable for you. Feel free to change it to a less cholesterol dishes with special house promotion and I recommend you to take this quality wine as a treat to have a healthy eating while enjoying your stay with us.'</p>
<p><i>Scenario 2: Health care:</i></p> <p>(1) The enhancement from HRI to HRC: Robots have dynamic thinking and real-time neural image captioning ability: able to deal with emergencies and a quick decision making from what it sees the environment in real-time.</p> <p>(2) Robots has been improved from conventional smart Q&A and interactions to new concept of bionic companionship.</p>	<p>You bought a robot at home to monitor your health. The robot obtains some of your health indicators (such as temperature, blood pressure, etc.) through some external devices. When there is a problem with your indicators, the robot can give you corresponding suggestions or help you contact a doctor. A total of 1 day you suddenly fainted at home for some reasons, but because you did not aim at the detection device connected to the robot, the robot did not find your condition. Fortunately, your neighbor found you fainted at home. . .</p>	<p>You bought a home care robot to monitor your health. The robot obtains some of your health indicators (such as pulses, blood pressure, etc.) through some external devices. When there is a problem with your indicators, the robot can give you corresponding suggestions or help you contact a doctor. A total of 1 day you suddenly fainted at home for some reason. The robot discovered your real-time condition through the deep learning vision system and contacted the your family member or hospital in time, subject to what the robot sees, e.g., fainted human with lots of blood or motion (call hospital for emergency); fainted human with conscious and free speech (call family members).</p>

the continuous development of deep learning, some researchers have recently realized the transformation of static pictures or videos from conventional camera input into text descriptions (*Li et al., 2020; Hu et al., 2020; Luo, Hsu & Ye, 2019*). This deep learning algorithm model is called neural image capturing (NIC). This research attempts to adapt and integrate NIC into HSRs and propose a novel framework (bionic-companionship framework) to enhance the traditional HRI experience. This framework aims to improve the current HRI interaction mode in the field of HSRs to a higher level of HRC (*Yang & Chew, 2021*). The bionics in this research refers to the humanoid service robot imitating all the tastes of life, trying to adapt to the seven emotions of ancient human nature (joy, anger, sadness, fear, love, disgust, liking) and six biological wills (life, death, eyes, ears, mouth, nose) (*Chew et al., 2021*). The system proposed in this study combines visual intelligence and Speech Intelligence, and imitates human behavior in social activities, which is in line with the

concept of robot bionics proposed by researchers such as *Chew et al. (2021)*. Therefore, this study believes that the proposed system is a bionic system.

RELATED WORKS

With the continuous development of HRI research, industrial robots have been able to interact with humans accurately and self-adaptively. Some advanced control systems (*Zhang, Hu & Gow, 2020*) and algorithms (*Tang, Zhang & Hu, 2020*) have been proposed as Industrial robots provide reliable support for completing interactive tasks in an industrial environment. However, as HSRs began to enter the service industry, some research cases began to discover that there are still problems with the interaction of HSRs in the social environment. *Caleb-Solly et al. (2018)* believed that users can also help robots when robots help users; meanwhile, users can give feedback to optimize the system. The feedback reflects not only the optimization of the robot system but also the satisfaction of customers. *Chung & Cakmak (2018)* study indicated that hotels in the hospitality industry want to collect customer feedback in real-time to immediately disseminate positive feedback and respond to unsatisfactory customers while they are still on the scene. Guests want to inform their experience without affecting their privacy. Stakeholders in the hospitality industry hope that intelligent robots can interact more with users. Besides, *Rodriguez-Lizundia et al. (2015)* concluded that the optimal distance between users and robots is 69.58 cm. Specifically, interaction with a certain greeting mode can attract users to maintain a longer interaction time; robots with the active search are more attractive to participants. The interaction time is longer than that of passively searching robots, suggesting that robots should be designed to keep at a certain distance from humans and consider adding the ability to allow robots to actively identify customers and attract them.

Research suggests that the current interactive system used by HSRs lacks the ability to process and adapt to dynamic social environments. The dynamic social environment here refers to the same human behavior and language often expressing different meanings in different social situations, such as In different situations, the handshake may require two completely different interactive messages to respond. Therefore, this research proposes the concept of HRC to develop a new interactive mode to solve the current problems faced by HRI in the hospitality industry. For a more detailed comparison of HRI and HRC, please refer to the video in the appendix link (<https://youtu.be/fZmV4MKeYtQ>).

Review of neural image captioning

The challenge of generating natural language descriptions from visual data has been extensively researched in the field of computer vision. However, early research has mainly focused on generating natural language descriptions from video-type visual data (*Gerber & Nagel, 1996; Mitchell et al., 2012*). These systems convert complex visual data into natural languages using rule-based systems. However, because the rules are artificially designed, these systems are sufficiently robust, bionic, and have been shown to be beneficial in limited applications such as traffic scenarios (*Vinyals et al., 2015*). In the past decade, various researchers, inspired by the successful use of sequence-to-sequence

training with neural networks for machine translation, proposed a method for generating image descriptions based on recurrent neural networks (RNNs) (Cho et al., 2014; Sutskever, Vinyals & Le, 2014). In fact, this method of replacing the encoder in the encoder-decoder framework in machine translation with image features transforms the original complex task of generating image data caption into a simple process of “translating” the image into a sentence (Cho et al., 2014). Furthermore, Donahue et al. (2014) used long short-term memory (LSTM) for end-to-end large-scale visual learning processes. In addition to images, Donahue et al. (2014) also applied LSTM to videos, allowing their models to generate video descriptions. Vinyals et al. (2015) and Kiros, Salakhutdinov & Zemel (2014) initially proposed the structure of a currently popular neural image generation algorithm based on the combination of a convolutional neural network (CNN) image recognition model and a natural language processing (NLP) structured model. Moreover, the neural image captioning algorithm based on the attention mechanism has also attracted extensive attention in the field of computer vision. Denil et al. (2012) proposed a real-time target tracking and attention recognition model driven by sight data. Tang, Srivastava & Salakhutdinov (2014) proposed an attention-generation model based on deep learning. From the perspective of visual neuroscience, the model requires object-centric data collection for model generation. Subsequently, Mnih et al. (2014) proposed a new recurrent neural network model, which can adaptively select specific areas or locations to extract information from images or videos and process the selected area at high resolution. As the algorithm has increasingly mature, the application of the algorithm in related fields has also been breaking through recently, such as the caption generation of car images (Chen, He & Fan, 2017), the description generation of facial expressions (Kuznetsova et al., 2014), and educational NAO robots driven by image caption generation for video Q&A games for children’s education (Kim et al., 2015). Recent research on image caption generation also shows that the accuracy and reliability of the technology have increased (Ding et al., 2019). In addition, reinforcement learning to automatically correct image caption generation networks have also been proposed (Fidler, 2017). These deep learning-based studies have undoubtedly laid a foundation for the possible NIC integration with HSRs as proposed in this study. The novel integration led to the possibility for humanoid robots to interact with humans while recognizing the social environment in real time, thereby improving the interactive service quality of the HSRs.

Neural image caption generation algorithm ‘crash into’ robot

An increasing number of studies have been conducted on HRI combined with image caption generation algorithm. Kim et al. (2015) used the structure of a convolutional neural network (CNN) combined with RNN + deep concept hierarchies (DCH) to design and develop an educational intelligent humanoid robot system for play video games with children. In this study, CNN was used to extract and pre-process cartoons with educational features, and RNN and DCH were used to convert the collected video features into Q&A about cartoons. During the game, after watching the same cartoon, the child and the robot ask and answer questions based on the content of the cartoon. The research results

show that such a system can interact effectively with children. However, for HRIs, such simple and limited-structured Q&A conditions cannot satisfy all the interaction scenarios required. [Cascianelli et al. \(2018\)](#) used a gull-gated recurrent unit (GRU) encoder-decoder architecture to develop a human-robot interface that provides interactive services for service robots. This research solves a problem called natural language video description (NLVD). The authors also compared the performance when using LSTM and GRU with two different algorithms to solve these problems. They demonstrated that the GRU algorithm runs faster and consumes less memory. This type of model may be more suitable for HSRs. Although the research model is competitive on public datasets, the experimental results on the designed datasets show that the model suffers from significant overfitting. This proves that in the actual model training process, a specific training dataset for HSR interaction should be established, and other methods (such as transfer learning) should be considered to improve the generalization ability of the model for interactive tasks. [Luo et al. \(2019\)](#) created a description template to add various image features collected by the robot, such as face recognition and expression, to the generated description. Compared with the previous models, their interaction is slightly more natural and closer to the human description. However, [Luo et al. \(2019\)](#) use the model to provide limited services to industry managers, hard to generalize, and not for developing an entire HRI framework.

Like the research on robot vision language, research on robot vision action is in its infancy. [Yamada et al. \(2016\)](#) used RNNs to enable robots to learn commands online from humans and respond with corresponding behaviors. This research furthermore provides a reference and direction for humanoid robots to use deep learning to obtain online learning capabilities for human commands. Inspired by the above study, the rationale and hypothesis proposed in the present research are that the description generated by the neural image captions can drive HSRs to perform appropriate behaviors, and HSRs can even obtain online learning capabilities of interacting with surrounding people through studying and analyzing social environments. [Tremblay et al. \(2018\)](#) and [Nguyen et al. \(2018\)](#) believe that non-experts often lack the rationality of task descriptions when issuing instructions to robots. They use deep learning to allow robots to automatically generate human-readable instructions' descriptions according to the surrounding social environment. In addition, [Nguyen et al. \(2018\)](#) also used visual data to make humanoid robots imitate and learn human actions under corresponding commands so that the robot can learn how to complete the corresponding tasks only through visual data; however, social robots cannot complete precise control of movements when they imitate movements of visual data.

CONTRIBUTION TO THE KNOWLEDGE: THE BIONIC-COMPANIONSHIP FRAMEWORK WITH NIC FOR HSRs

The contribution of the present study is the novel investigation and design of the bionic-companionship framework for HSRs, adapting and integrating neural image caption generation algorithms and bionic humanoid robots, to be validated in a lab-controlled environment and real-life exploration. The new HRC framework is anticipated to enhance HRI to reach a new state, making it possible for HSRs to become bionic companions of humans.

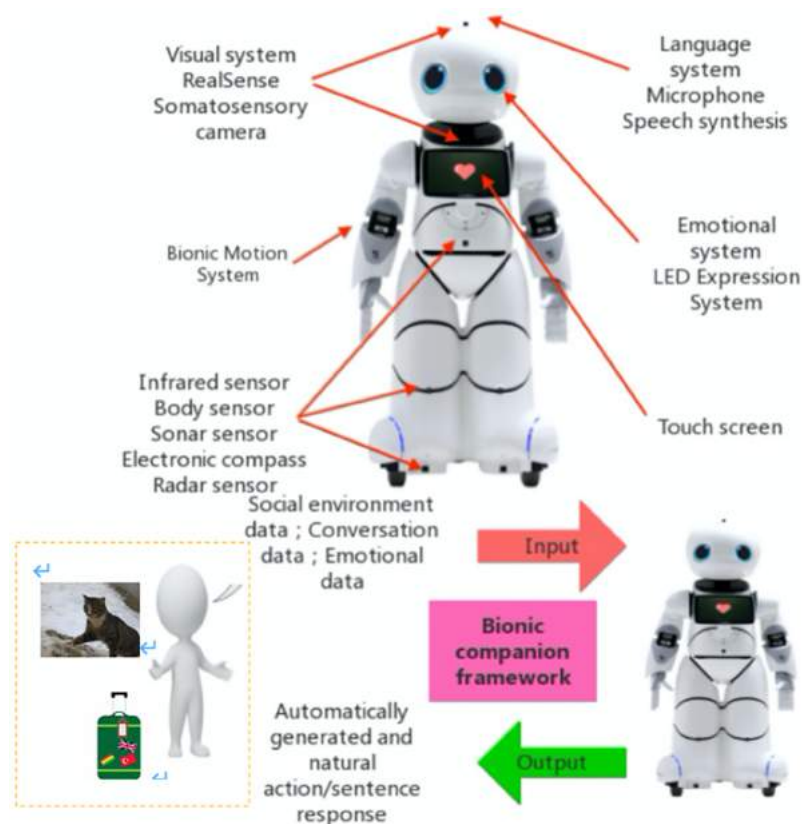


Figure 1 HSR capabilities with the proposed high-level of HRC conceptual model.

Full-size [DOI: 10.7717/peerj-cs.674/fig-1](https://doi.org/10.7717/peerj-cs.674/fig-1)

This study proposes adapting and integrating deep learning techniques to one of the world's most advanced HSRs so that robots can autonomously and in a timely fashion convert pictures or data information captured by robotic visions and sensors into texts or sentences in order to respond and communicate more naturally with humans. The conceptual model of the proposed system consists of various modules, as shown in Figs. 1, 2. The contributions of this research are summarized as follows:

1. In order to solve the current problems of HSRs in the hospitality industry, a new interactive concept-HRC is proposed.
2. A novel bionic interaction framework is designed based on the proposed HRC.
3. A system that can be used on HSRs is developed based on the bionic interaction framework, and the system has been tested and verified. The preliminary results prove that the system can enable HSRs to handle dynamic social environments.

Humanoid service robot used in research

The design and investigation of this HRC framework involves using the Canbot U05E humanoid robot (see Fig. 1 for the high-level design, Figs. 2–5 for further details) (CANBOT, 2020). The robot's 22-degree-of-freedom motion joints enable it to perform a

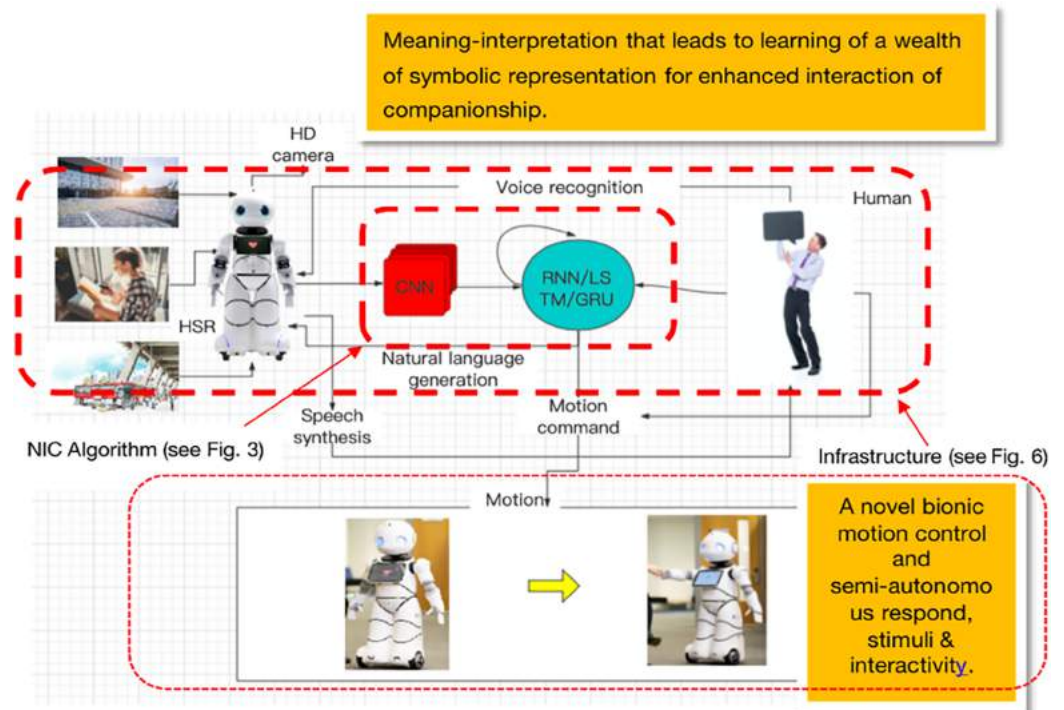


Figure 2 Bionic-companionship framework design.

Full-size [DOI: 10.7717/peerj-cs.674/fig-2](https://doi.org/10.7717/peerj-cs.674/fig-2)

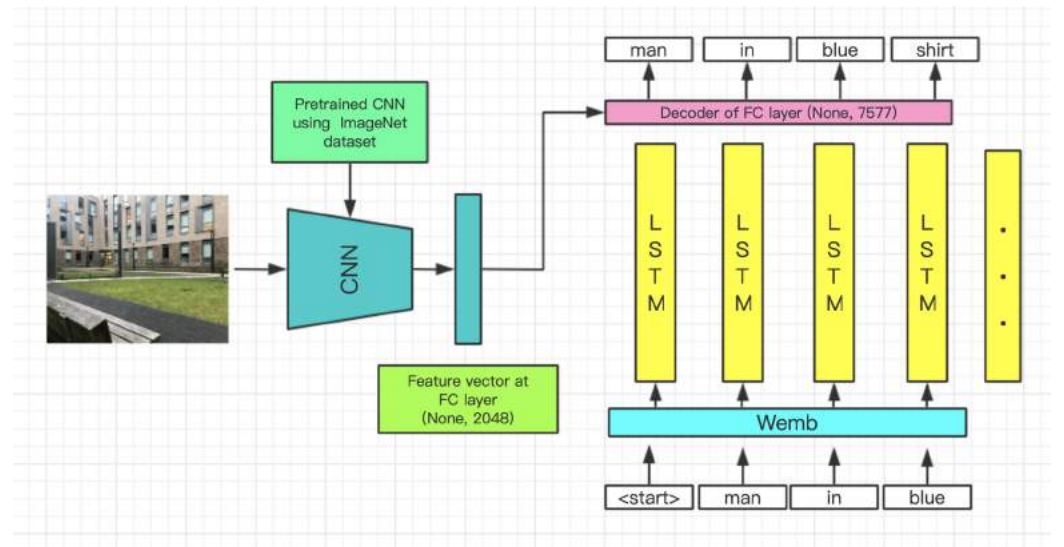


Figure 3 Neural image captioning model structure for HSR.

Full-size [DOI: 10.7717/peerj-cs.674/fig-3](https://doi.org/10.7717/peerj-cs.674/fig-3)

variety of simulated movements, such as raising the head, turning the head, raising the arm, shaking the crank, shaking hands, leaning back, walking, and turning, and based on the proposed framework, it can acquire natural human behaviors and, as a result, efficiently interact with humans. In addition, Canbot U05E's advanced vision system and sensors can collect more complete environmental data for the proposed design and make



0. A black dog is running after a white dog in the snow .
1. Black dog chasing brown dog through snow
2. Two dogs chase each other across the snowy ground .
3. Two dogs play together in the snow .
4. Two dogs running through a low lying body of water .



0. A man and a dog are sitting on a park bench with several people walking down the path that runs alongside it .
1. A man and a dog on a park bench in the foreground , with a group of walking people in the distance .
2. A man sits on a bench in the park with a dog , while others walk nearby .
3. People sit on a park bench while others jog on a path .
4. The man sits with his dog on a bench in a park .

Figure 4 Samples Flickr 8k (*Rashtchian et al., 2010*) training data set.

Full-size DOI: [10.7717/peerj-cs.674/fig-4](https://doi.org/10.7717/peerj-cs.674/fig-4)

the novel framework more robust. The robot is designed to imitate the human's seven senses, providing strong support for the concept and implementation of the bionic partner designed in this study.

Bionic-companionship framework

In this study, we review the previous works on this topic and research gaps in the literature and describe a novel humanoid service robot and human interaction framework with neural image subtitles as its core (details are shown in Fig. 2). The framework uses the structure of the NIC algorithm to better realize the interaction of HSRs from HRI to the direction of bionic-companionship. According to the initial descriptions of robot companions, as in the studies by *Turkle (2006)* and (*Kim et al., 2015*), the proposed framework should provide HSRs with more natural interactions and a more sensitive understanding of the environment, and hence, the design of the framework is divided into two subsystems (see the dotted red).

Image/video description generation system

These subsystems are the core modules of the entire interactive framework. HSRs collect visual data of the surrounding environment through equipped visual sensors (such as HD or 3D cameras) and sensors (such as tactile and radar). The type of visual data collected depends on the complexity of the interactive task to be completed by HSRs. It is generally considered that more complex interactive tasks require the use of continuous images or real-time videos. The system uses the latest neural image generation algorithm structure and CNN to perform feature extraction on the pictures and video data of the surrounding social environment, and converts the data into feature vector sequences that

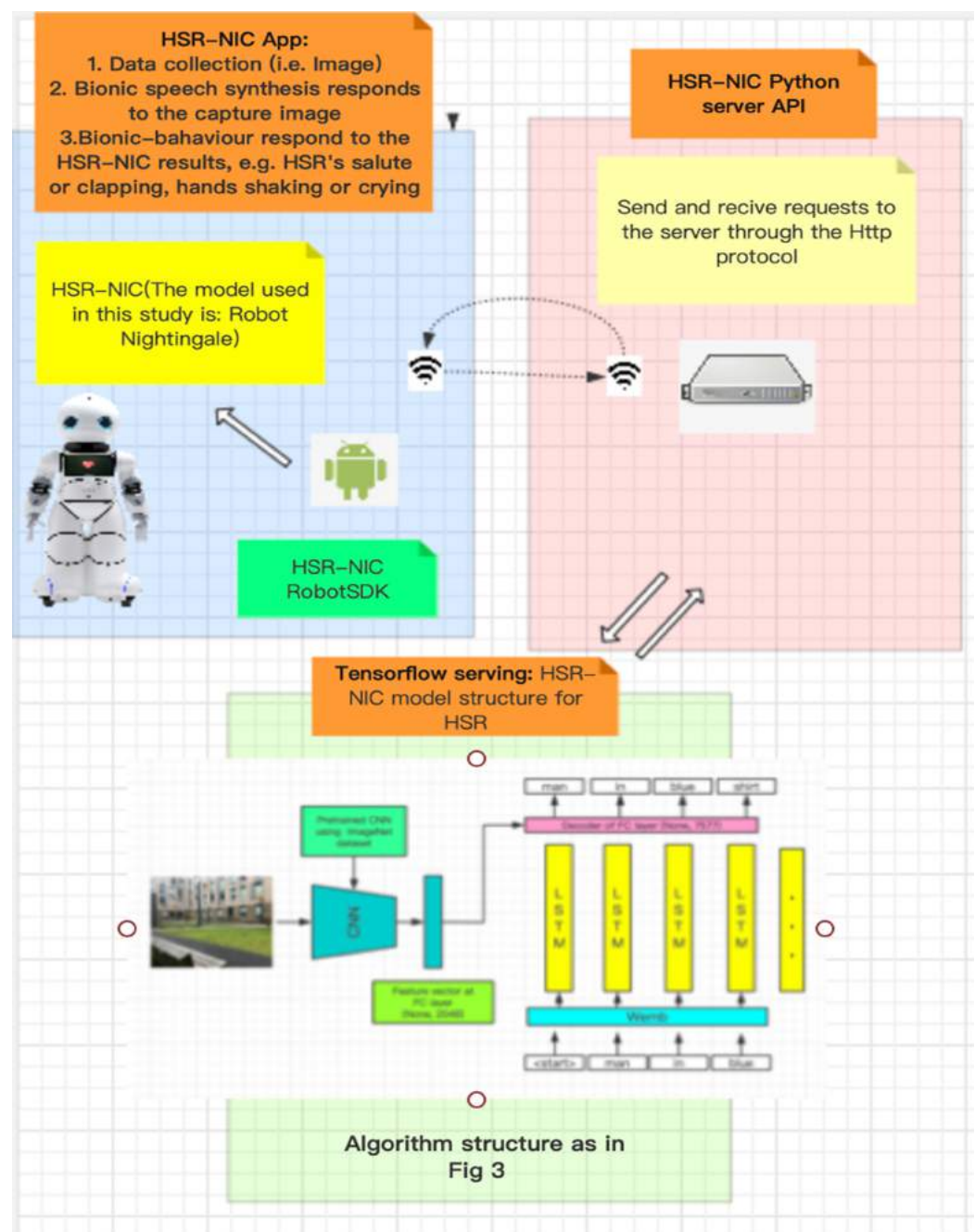


Figure 5 The infrastructure of the humanoid service robot generating neural image captions as part of the bionic companionship framework. [Full-size DOI: 10.7717/peerj-cs.674/fig-5](https://doi.org/10.7717/peerj-cs.674/fig-5)

can be used by RNN. Finally, the RNN completes the process of generating an interactive description from the visual data. HSRs use a speech synthesis system that converts these descriptions into voices to communicate with humans. This process is different from the past mode of using HSRs human sensing sensors and setting fixed interactive feedback; the innovation of this system is that HSRs can automatically and naturally generate interactive feedback. This means that the change in the scene during the interaction will

cause a continuous change in the interaction feedback, and this change is not preset by humans. In addition, in further conversation interactions, human voice response and social environment data will be coordinated by HSRs and produce continuous conversation interaction behavior.

Command-robot behavior system

For HSRs, simple conversation interactions are insufficient. HSRs should generate corresponding motions based on visual and human behavior data. For example, when humans wave to a robot, the robot should also actively respond. The hypothesis of this study is to classify or cluster description text generated from visual data and use these classified description texts to control the motions of HSRs in response to complex interactive tasks. For example, when the description generated by neural image captions is “Hello”, then HSRs will automatically determine whether ‘Hello’ matches a category that requires interactive motion and performs corresponding motions such as waving.

PILOT TESTING, PRELIMINARY RESULTS, AND DISCUSSION

In the present study, we designed and integrated a classic NIC model on the HSR and performed a preliminary evaluation.

Introduction to HSR-NIC model

The structure of the HSR-NIC algorithm used in this study was adapted and enhanced from the model structure proposed by [Mao et al. \(2014\)](#) who used a classic encoder-decoder structure. In this study, the encoder uses the Xception pre-trained CNN to convert the input image into a feature vector. The word sequence is then input into the LSTM after a layer of word embedding layer, and finally, an add operation is performed on the word features output by the LSTM and the image features extracted by the trained CNN. These are then input into a decoder composed of a single-layer fully connected layer, which generates the probability distribution of the next word using a softmax layer. The LSTM introduced by the model can solve the long-term dependency problem in the traditional RNN, thereby improving the accuracy of the model. The dense representation of word embedding can reduce the amount of calculations involved in the model; it also enables the model to capture similar relationships between words. In addition, the model used in this study also introduces a dropout layer with a probability of 50% to increase the robustness of the model. The teacher forcing mechanism was used during model training to accelerate the model training process. The optimizer used in the research is Adam, which has the advantages of making the model converge more quickly and automatically adjusting the learning rate with learning. The variables of the model are updated by minimizing the cross-entropy loss between the probability distribution of the predicted result and the probability distribution of the true result and back-propagation. The model structure diagram as follow ([Fig. 3](#)).

Model forward propagation process

The training process of the image captioning task can be described as follows: For a picture in the training set, its corresponding description is a sequence that represents the words in the sentence. For model θ , given input image I from the HSR's vision, the probability of the model generating sequence is expressed as

$$P(S|I; \theta) = \prod_{t=0}^N P(S_t|S_0, S_1, \dots, S_{t-1}, I; \theta) \quad (1)$$

The logarithm of the likelihood function is used to obtain the log-likelihood function:

$$\log P(S|I; \theta) = \sum_{t=0}^N \log P(S_t|S_0, S_1, \dots, S_{t-1}, I; \theta) \quad (2)$$

The training objective of the model is to maximize the sum of the log-likelihoods of all training samples:

$$\theta^* = \arg\max_{\theta} \sum_{(I,S)} \log P(S|I; \theta) \quad (3)$$

where (I, S) is the training sample. This method of maximum likelihood estimation is equivalent to empirical risk minimization using the log-loss function. Therefore, in the forward propagation process of this research model, the image feature vector I_v is extracted from the image using the CNN, and a two-dimensional vector of shape (batch size, 2048) is the output.

$$I_v = \text{CNN}_{\theta_c}(I) \quad (4)$$

The extracted image features need to be encoded by a fully connected layer into the context feature vector C that can be matched with word features. The word feature vector is the output O_t of the LSTM over the time step. The input word of LSTM passes through a word-embedding layer to generate a dense vector representation $W(s)$.

$$C = W_{\theta}(I_v), O_t = \text{LSTM}_{\theta}(W(s)) \quad (5)$$

Finally, word feature O_t and context feature C are together input into a decoder composed of a single fully connected layer after the softmax calculation generates the probability distribution of the next word $P(S_i|I; \theta)$.

$$P(S_i|I; \theta) = \text{softmax}(W_{\theta}(C + O_t)) \quad (6)$$

The loss function is expressed as

$$L = \sum_{t=1}^T y^{(t)} \log p^{(t)} + (1 - y^t) \log(1 - p^t) \quad (7)$$

Training dataset

For the present study, we use Flickr 8k ([Rashtchian et al., 2010](#)) as the training dataset. This is a new benchmark collection for sentence-based image descriptions and searches. It consists of 8,000 images. Each image was paired with five different captions. These

captions provide content descriptions of the objects and events in the picture. The images do not contain any well-known people or locations but depict random scenes and situations. Examples of datasets are shown in Fig. 4. The Flickr 8k dataset not only contains images of animals and objects, but also of some social scenes. These data can help robots to better understand natural, day-to-day scenes.

The process of humanoid service robot generating image captions

To explore the feasibility of the bionic-companionship framework, preliminary tests were conducted on a real humanoid service robot (Canbot U05E). The process of generating image captions by a humanoid service robot is divided into four steps, as shown in Fig. 5.

Step 1. The HSR-NIC API is responsible for controlling the robot to call the high-definition camera to collect surrounding environment information (the data collection in this study is focused on HSR capture images). The collected data will be sent to the local host service program through the HTTP protocol and wait for a response from the HSR.

Step 2. The HSR-NIC localhost server program receives the data, and the requests perform preliminary processing and cleaning of the data (image) and send the data (image) to the HSR-NIC model server program to wait for the calculation result (the generated caption description).

Step 3. The HSR-NIC model server program analyzes the image data according to the training parameters saved before, generates the descriptive caption, and returns it to the local server.

Step 4. The HSR-NIC local server program sends the caption description to the robot application through the HTTP protocol, and the robot application controls the robot to respond according to the caption description, such as speech synthesis and motion control.

Preliminary test results and limitation

In this study, we conducted a preliminary test on a humanoid service robot integrated with the NIC algorithm. The results of the preliminary test were found to be promising.

With the discuss of the last chapter, the research will integrate the NIC into the HSRs to make the HSRs take advantage of the change of the surrounding environment interact with the human better. Therefore, the system proposed by this research will combine qualitative analysis and quantitative analysis to initially validate the performance of the system.

This study introduces the cross-entropy loss curve of the last 50 epochs of the model as the evaluation metric for quantitative analysis. As shown in the Fig. 6, the model finally converges to the minimum loss value of 2.65 in the training set and 2.71 in the validation set, which proves that the model has no over-fitting and under-fitting, and has generalization ability. Since the loss value is calculated from the sum of the difference between the probability value of each predicted word in the predicted description and the true value, the loss value will be affected by the sentence length of the predicted description. In related work, researchers (Li et al., 2020; Hu et al., 2020) used some more reliable evaluation methods to evaluate the performance of the model, including the

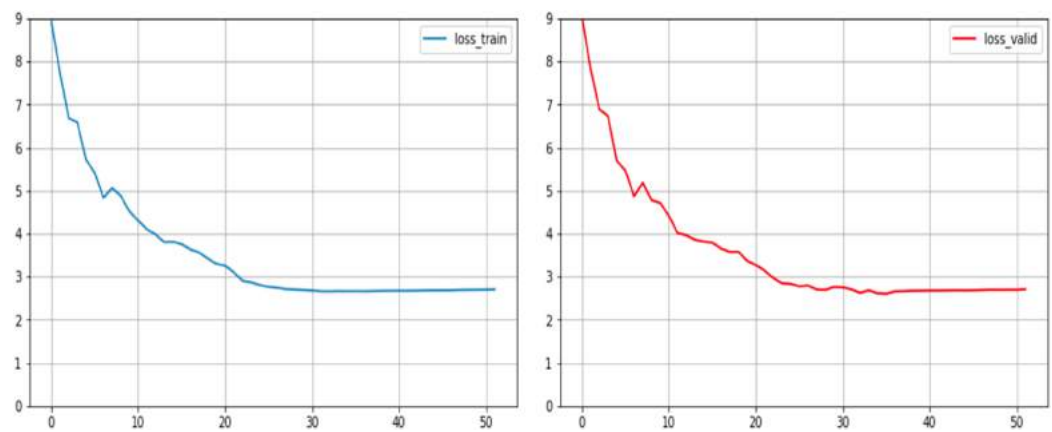


Figure 6 Loss curve of NIC model on training set and validation set.

Full-size DOI: 10.7717/peerj-cs.674/fig-6

BLUE4 (Papineni et al., 2002) and CIDEr (Vedantam, Lawrence Zitnick & Parikh, 2015). These evaluation metrics are usually used in the field of machine translation instead of manual evaluation. Since the tasks handled by the NIC model can be regarded as translated from images/scenes into English, the evaluation metrics can also be applied to the evaluation of NIC. This study will use qualitative analysis to replace quantitative analysis of metrics such as BLUE4 and CIDEr, so as to further evaluate the preliminary performance of HSR after the integrated NIC model.

As shown in Figs. 7 and 8, the researcher conducted two sets of tests in three different scenarios with HSR. In the first set of tests, the researcher wore a hat and changed scenarios. In the second set of tests, the researcher did not wear a hat, and the scene switching method was the same as in the first set. It can be seen from the experimental results that the humanoid robot can complete the perception of scene switching through this algorithm and generate a rough description of the scene. In the first set of tests, most of the content described was accurate. The robot equipped with the NIC algorithm can effectively identify 'man', 'black shirt', and 'sitting on a bench'. However, in the second group of tests, there were many errors in the recognition results. This could be attributed to the researcher's long hair. Interestingly, researchers with long hair are easily identified as women or children. This indicates that the accuracy of the NIC algorithm still has room for improvement.

In addition, in order to test the performance of the system in a dynamic environment. The researcher conducted the test in a real environment (as Fig. 8). The researcher selected six real environments as the test data and let the robot generate interactive information. Among the six real interactive environments, there are three scenes that can be more accurately recognized by the robot and produce corresponding descriptions. The description information can correspond to the test environment, and the corresponding part of the description has been highlighted with the same color in the Fig. 8. Some of the objects, facilities, and human movements in these scenes can be accurately predicted, such as sidewalk, traffic, bench, building, building, etc. However, in the other three environments, the robot did not give an accurate description. The researchers believe that

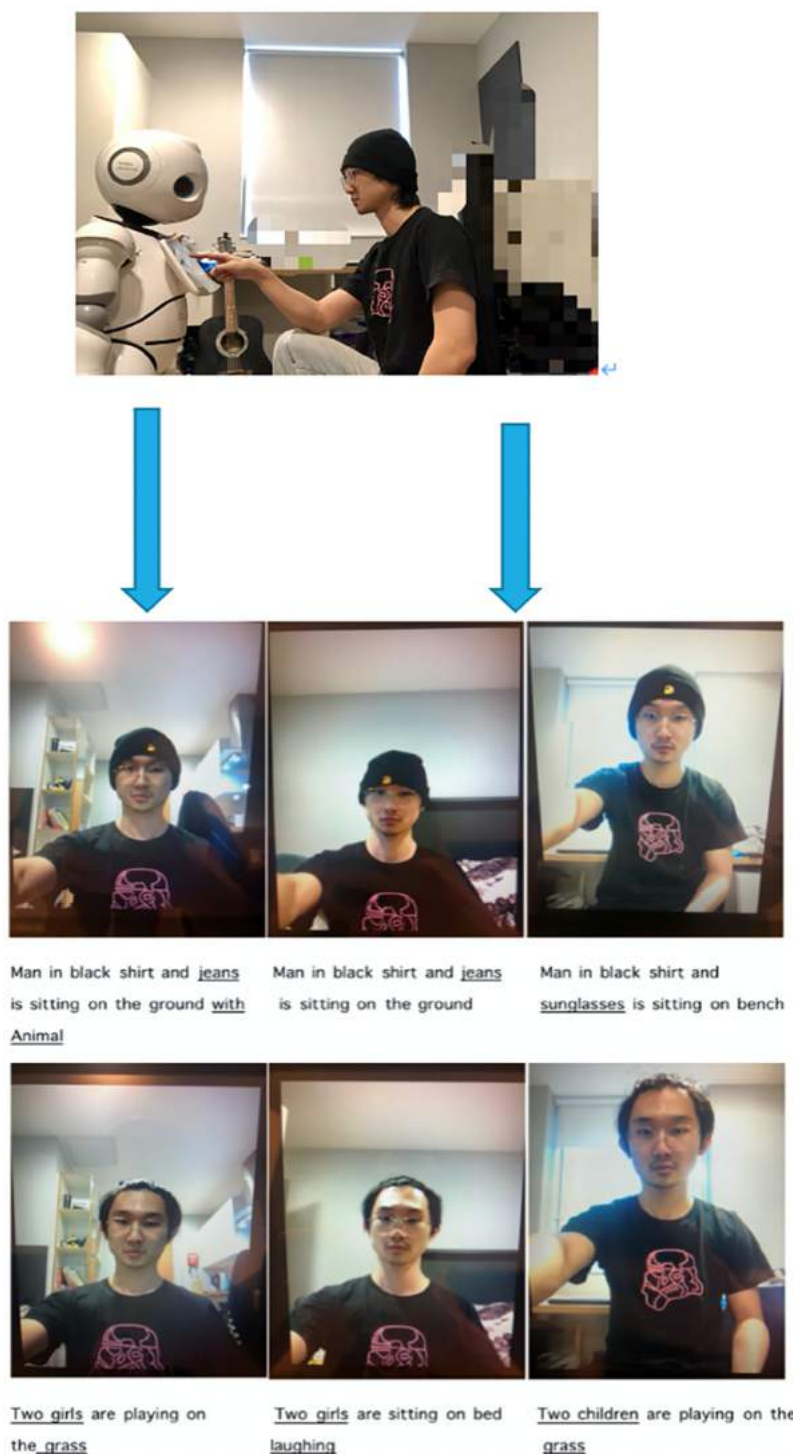
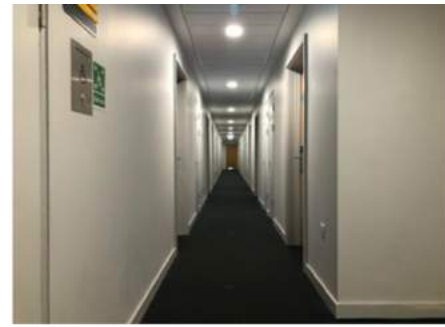


Figure 7 A series of preliminary testing results captured from Canbot U05E and bionic-companionship preliminary framework. [Full-size !\[\]\(fcc3264021d438d9732560e78099f674_img.jpg\) DOI: 10.7717/peerj-cs.674/fig-7](https://doi.org/10.7717/peerj-cs.674/fig-7)

this may be due to the fact that the training set does not contain objects in these three environments, causing the model to fail to learn how to express the ‘unfamiliar environment’.



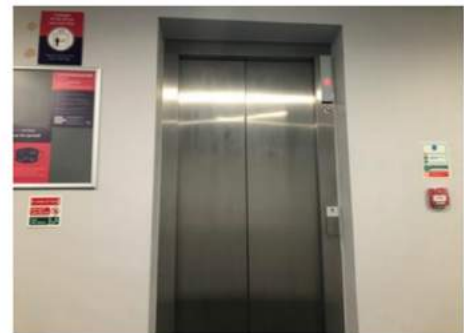
man in helmet is sitting on the sidewalk with traffic



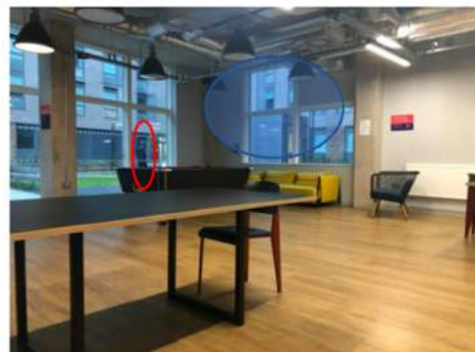
man is sitting on bench with his feet in the air



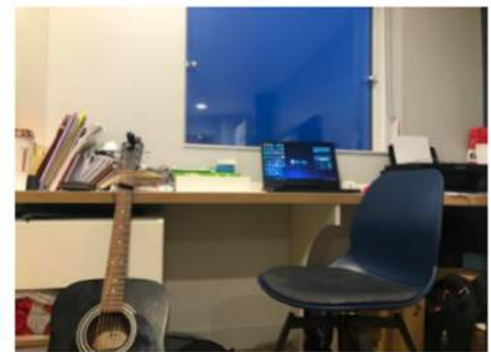
man in blue shirt is sitting on bench with his feet up in front of building



man sitting on the ground with his feet up in the air



man in black shirt is standing in front of large building



man is standing in front of crowd

Figure 8 Real social environment test examples.

Full-size  DOI: 10.7717/peerj-cs.674/fig-8

In general, as per the results of the two experimental sets, it was proven that the robot equipped with the NIC algorithm can capture the changes in the surrounding environment and generate different feedbacks according to the changes. The results also demonstrate the feasibility of the proposed bionic-companionship framework. Although there is still a gap between the prediction results of the algorithm and the real communication scene, the researcher believes that special data collection for some specific interaction scenarios and model training for these specific data can be effective in addressing this gap. Future

research directions will mainly focus on improving the accuracy of algorithms and achieving more human-like interactions. (The detailed process is shown in the HSR-NIC demo video.) In addition, the researcher believes that the scene understanding of static images is the basis for dealing with dynamic environments. Some researches have mentioned that the introduction of related algorithms of object detection into NIC can identify and generate descriptions of scenes in dynamic environments. This is also the current research limitation of this research and the research challenges that will be faced in the future.

CONCLUSIONS

This study presents a review of neural image generation algorithms and application cases in the field of robotics, and proposes a novel humanoid service robot and human interaction framework based on the bionic-companionship theory. The subsystems of the bionic-companionship framework are designed and introduced in detail. Preliminary tests also initially proved that the framework could increase the sensitivity of HSRs to changes in the surrounding environment. The proposed framework will contribute to further development from HRI to HRC. Future work will focus on implementing each of the subsystems in the framework and applying the framework to HSRs to verify its performance.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The authors received no funding for this work.

Competing Interests

Esyin Chew & Pengcheng Liu are Academic Editors for PeerJ.

Author Contributions

- Jiaji Yang conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Esyin Chew conceived and designed the experiments, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Pengcheng Liu conceived and designed the experiments, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The research code is available in the [Supplemental File](#). The training dataset for the system is Flickr 8k and is available at Kaggle: <https://www.kaggle.com/adityajn105/flickr8k>.

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.674#supplemental-information>.

REFERENCES

- CANBOT. 2020. Intelligent Robot AI Makers Laboratory. Available at <https://www.canbotrobots.com/html/yy->.
- Caleb-Solly P, Dogramadzi S, Huijnen CA, Heuvel HVD. 2018. Exploiting ability for human adaptation to facilitate improved human-robot interaction and acceptance. *The Information Society* 34(3):153–165 DOI 10.1080/01972243.2018.1444255.
- Cascianelli S, Costante G, Ciarfuglia TA, Valigi P, Fravolini ML. 2018. Full-GRU natural language video description for service robotics applications. *IEEE Robotics and Automation Letters* 3(2):841–848 DOI 10.1109/LRA.2018.2793345.
- Chen L, He Y, Fan L. 2017. Let the robot tell: describe car image with natural language via LSTM. *Pattern Recognition Letters* 98(8):75–82 DOI 10.1016/j.patrec.2017.09.007.
- Chew E, Lee PL, Hu S, Yang J. 2021. Investigating the first robotic nurses: humanoid robot nightingale and partners for COVID-19 preventive design. In: *The Seventh Edition of the International Workshop on New Trends in Medical and Service Robots Conference*. Basel, Switzerland: University Hospital Basel.
- Cho K, van Merriënboer B, Gulcehre C, Bougares F, Schwenk H. 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: *EMNLP*.
- Chung MJY, Cakmak M. 2018. How was your stay?: Exploring the use of robots for gathering customer feedback in the hospitality industry. In: *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. Piscataway: IEEE, 947–954.
- Dautenhahn K, Woods S, Kaouri C, Walters ML, Koay KL, Werry I. 2005. What is a robot companion-friend, assistant or butler? In: *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Piscataway: IEEE, 1192–1197.
- Denil M, Bazzani L, Larochelle H, de Freitas N. 2012. Learning where to attend with deep architectures for image tracking. *Neural computation* 24(8):2151–2184 DOI 10.1162/NECO_a_00312.
- Ding S, Qu S, Xi Y, Sangaiah AK, Wan S. 2019. Image caption generation with high-level image features. *Pattern Recognition Letters* 123(1):89–95 DOI 10.1016/j.patrec.2019.03.021.
- Donahue J, Hendricks LA, Guadarrama S, Rohrbach M, Venugopalan S, Saenko K, Darrell T. 2014. Long-term recurrent convolutional networks for visual recognition and description. arXiv:1411.4389v2. Available at <https://arxiv.org/abs/1411.4389>.
- Fidler S. 2017. Teaching machines to describe images with natural language feedback. In: *Advances in Neural Information Processing Systems*. 5068–5078.
- Fong T, Thorpe C, Baur C, C. Collaboration. 2003. Dialogue, human-robot interaction. In: *Robotics Research*. Berlin, Heidelberg: Springer, 255–266.
- Gerber R, Nagel N-H. 1996. Knowledge representation for the generation of quantified natural language descriptions of vehicle traffic in image sequences. In: *Proceedings of 3rd IEEE International Conference on Image Processing*. 2:805–808.
- Harris K, Kimson A, Schwedel A. 2018. Why the automation boom could be followed by a bust, Harvard business review (March 13). Available at <https://hbr.org/2018/03/why-the-automation-boom-could-be-followed-by-a-bust>.

- Hu X, Yin X, Lin K, Wang L, Zhang L, Gao J, Liu Z. 2020.** Vivo: surpassing human performance in novel object captioning with visual vocabulary pre-training. arXiv preprint arXiv:2009.13682. Available at <https://arxiv.org/abs/2009.13682>.
- Ivanov S. 2019.** Ultimate transformation: how will automation technologies disrupt the travel, tourism and hospitality industries? *Zeitschrift für Tourismuswissenschaft* **11**(1):25–43 DOI [10.1515/tw-2019-0003](https://doi.org/10.1515/tw-2019-0003).
- Ivanov S, Gretzel U, Berezina K, Sigala M, Webster C. 2019.** Progress on robotics in hospitality and tourism: a review of the literature. *Journal of Hospitality and Tourism Technology* **9**074(3):75 DOI [10.1108/JHTT-08-2018-0087](https://doi.org/10.1108/JHTT-08-2018-0087).
- Kiros R, Salakhutdinov R, Zemel R. 2014.** Multimodal neural language models. In: *International Conference on Machine Learning*. 595–603.
- Kim KM, Nan CJ, Ha JW, Heo YJ, Zhang BT. 2015.** Pororobot: a deep learning robot that plays video Q&A games. In: *AAAI Fall Symposium Series*.
- Kuznetsova P, Ordonez V, Berg TL, Choi Y. 2014.** TREETALK: composition and compression of trees for image descriptions. *Transactions of the Association for Computational Linguistics* **2**(1):351–362 DOI [10.1162/tac1_a_00188](https://doi.org/10.1162/tac1_a_00188).
- Luo RC, Hsu YT, Ye HJ. 2019.** Multi-modal human-aware image caption system for intelligent service robotics applications. In: *IEEE 28th International Symposium on Industrial Electronics (ISIE)*. Piscataway: IEEE, 1180–1185.
- Luo RC, Hsu Y, Wen Y, Ye H. 2019.** Visual image caption generation for service robotics and industrial applications,. In: *2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS)*, Taipei, Taiwan. 827–832.
- Li X, Yin X, Li C, Zhang P, Hu X, Zhang L, Wang L, Hu H, Dong L, Wei F, Choi Y, Gao J. 2020.** Oscar: object-semantics. aligned pre-training for vision-language tasks. In: *European Conference on Computer Vision*. Cham: Springer, 121–137.
- Mao J, Xu W, Yang Y, Wang J, Huang Z, Yuille A. 2014.** Deep captioning with multimodal recurrent neural networks (m-rnn). arXiv preprint. Available at <https://arxiv.org/abs/1412.6632>.
- Mitchell M, Han X, Dodge J, Mensch A, Goyal A, Berg A, Yamaguchi K, Berg T, Stratos K, Daumé H III. 2012.** Midge: generating image descriptions from computer vision detections. In: *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, 747–756.
- Mnih V, Hees N, Graves A, Kavukcuoglu K. 2014.** Recurrent models of visual attention. In: *NIPS*.
- Nguyen A, Kanoulas D, Muratore L, Caldwell DG, Tsagarakis NG. 2018.** Translating videos to commands for robotic manipulation with deep recurrent neural networks. In: *IEEE International Conference on Robotics and Automation (ICRA)*. Piscataway: IEEE, 1–9.
- Papineni K, Roukos S, Ward T, Zhu WJ. 2002.** Bleu: a method for automatic evaluation of machine translation. In: *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*. 311–318.
- Rodriguez-Lizundia E, Marcos S, Zalama E, Gómez-García-Bermejo J, Gordaliza A. 2015.** A bellboy robot: study of the effects of robot behaviour on user engagement and comfort. *International Journal of Human-Computer Studies* **82**(1):83–95 DOI [10.1016/j.ijhcs.2015.06.001](https://doi.org/10.1016/j.ijhcs.2015.06.001).
- Rashtchian C, Young P, Hodosh M, Hockenmaier J. 2010.** Collecting image annotations using amazon’s mechanical turk. In: *Proceedings of the NAACL HLT, 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*. 139–147.

- Sutskever I, Vinyals O, Le QV. 2014. Sequence to sequence learning with neural networks. In: *NIPS*. 3104–3112.
- Tang X, Zhang Q, Hu L. 2020. An EKF-based performance enhancement scheme for stochastic nonlinear systems by dynamic set-point adjustment. *IEEE Access* 8:62261–62272 DOI 10.1109/ACCESS.2020.2984744.
- Tang Y, Srivastava N, Salakhutdinov RR. 2014. Learning generative models with vi-sual attention. In: *NIPS*. 1808–1816.
- Tremblay J, To T, Molchanov A, Tyree S, Kautz J, Birchfield S. 2018. Synthetically trained neural networks for learning human-readable plans from real-world demonstrations. In: *IEEE International Conference on Robotics and Automation (ICRA)*. Piscataway: IEEE, 1–5.
- Trevarthen C. 2001. Intrinsic motives for companionship in understanding: their origin, development, and significance for infant mental health. *Infant Mental Health Journal: Official Publication of the World Association for Infant Mental Health* 22(1–2):95–131 DOI 10.1002/(ISSN)1097-0355.
- Turkle S. 2006. A nascent robotics culture: new complicities for companionship. American Association for Artificial Intelligence Technical Report Series AAAI.
- Vedantam R, Lawrence Zitnick C, Parikh DC. 2015. Consensus-based image description evaluation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Piscataway: IEEE, 4566–4575.
- Vinyals O, Toshev A, Bengio S, Erhan D. 2015. Show and tell: a neural image caption generator. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE, 3156–3164.
- Yang J, Chew E. 2020. A systematic review for service humanoid robotics model in hospitality, springer international journal for social robotics. *International Journal of Social Robotics* 22(2):253 DOI 10.1007/s12369-020-00724-y.
- Yang J, Chew E. 2021. The Design Model for Robotic Waitress. *International Journal of Social Robotics* 1–11 DOI 10.1007/s12369-021-00745-1.
- Yamada T, Murata S, Arie H, Ogata T. 2016. Dynamical integration of language and behavior in a recurrent neural network for human-robot interaction. *Frontiers in Neurorobotics* 10:5 DOI 10.3389/fnbot.2016.00005.
- Zhang QC, Hu L, Gow J. 2020. Output feedback stabilization for mimo semi-linear stochastic systems with transient optimisation. *International Journal of Automation and Computing* 17(1):83–95 DOI 10.1007/s11633-019-1193-8.

The classification of skateboarding tricks via transfer learning pipelines

Muhammad Amirul Abdullah¹, Muhammad Ar Rahim Ibrahim¹,
Muhammad Nur Aiman Shapiee¹, Muhammad Aizzat Zakaria¹,
Mohd Azraai Mohd Razman¹, Rabiul Muazu Musa², Noor Azuan Abu
Osman³ and Anwar P.P. Abdul Majeed^{1,4}

¹ Innovative Manufacturing, Mechatronics and Sports Laboratory, Faculty of Manufacturing and Mechatronics Engineering Technology, Universiti Malaysia Pahang, Pekan, Pahang, Malaysia

² Centre for Fundamental and Liberal Education, Universiti Malaysia Terengganu, Kuala Nerus, Terengganu, Malaysia

³ Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Kuala Lumpur, Malaysia

⁴ Centre for Software Development & Integrated Computing, Universiti Malaysia Pahang, Pekan, Malaysia

ABSTRACT

This study aims at classifying flat ground tricks, namely Ollie, Kickflip, Shove-it, Nollie and Frontside 180, through the identification of significant input image transformation on different transfer learning models with optimized Support Vector Machine (SVM) classifier. A total of six amateur skateboarders (20 ± 7 years of age with at least 5.0 years of experience) executed five tricks for each type of trick repeatedly on a customized ORY skateboard (IMU sensor fused) on a cemented ground. From the IMU data, a total of six raw signals extracted. A total of two input image type, namely raw data (RAW) and Continuous Wavelet Transform (CWT), as well as six transfer learning models from three different families along with grid-searched optimized SVM, were investigated towards its efficacy in classifying the skateboarding tricks. It was shown from the study that RAW and CWT input images on MobileNet, MobileNetV2 and ResNet101 transfer learning models demonstrated the best test accuracy at 100% on the test dataset. Nonetheless, by evaluating the computational time amongst the best models, it was established that the CWT-MobileNet-Optimized SVM pipeline was found to be the best. It could be concluded that the proposed method is able to facilitate the judges as well as coaches in identifying skateboarding tricks execution.

Subjects Artificial Intelligence, Data Science, Embedded Computing, Emerging Technologies

Keywords Classification, Support vector machine, Skateboarding, Machine learning, Transfer learning

INTRODUCTION

A skateboard is a short, narrow board with two small wheels attached to the bottom of either end. Skateboarders ride on this apparatus to perform tricks, including jumps (ollies), flips and mid-air spins. It is worth noting that this sport shall make its Olympic debut in the now delayed Tokyo 2020 Olympic Games. In general, in skateboarding competitions, the judging is done manually and subjectively through the observation of selected professional judges. However, it is worth mentioning at this juncture that the

Submitted 1 April 2021
Accepted 28 July 2021
Published 18 August 2021

Corresponding author
Anwar P.P. Abdul Majeed,
amajeed@ump.edu.my

Academic editor
Pengcheng Liu

Additional Information and
Declarations can be found on
page 15

DOI 10.7717/peerj-cs.680

© Copyright
2021 Abdullah et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Head Judge for Skatepark of Tampa & Board has pointed out the myriad difficulties in providing a judgement in a skateboarding event ([Pappalardo, 2014](#)). Amongst the notable factors reported were the style, speed, difficulty, consistency, trick selection and originality. Such obstacles are also faced by the coaches in providing comprehensive feedback to further improve the performance of the athletes ([Stein et al., 2018](#)).

Owing to the advancement of technology, the employment of machine learning and, to a certain extent, deep learning has received due attention in human and sports activity recognition. For instance, [Chen & Xue \(2015\)](#) employed a deep learning model on data captured *via* an accelerometer for human activity recognition (HAR). The authors extracted the acceleration data through an android phone with a sampling frequency of 100 Hz from the built-in tri-axial accelerometer. Eight activities were investigated, *i.e.*, falling, running, jumping, walking, step walking, walking quickly, walking downstairs and upstairs from 68 males and 32 females. A total of 31,688 labelled samples utilized, where 27,395 samples used for training, and the remaining 4,293 were used for testing. The authors fed the raw signal transformed images to Convolutional Neural Network with a three convolutional layer and three pooling layer architecture. Moreover, conventional feature extraction methods *via* Fast Fourier Transform, as well as Discrete Continuous Transform apart from the original time-domain signals that were paired with Support Vector Machine and Deep Belief Network models, were also investigated. It was shown from the study that the proposed CNN model could achieve a classification accuracy (CA) of 93.8% of accuracy and was better than that of the other models evaluated.

[Akula, Shah & Ghosh \(2018\)](#) investigated the use of multi-stage CNN on infrared images for HAR. The action data were collected from 18 females and 34 males between the age range of 19 to 28 years old. The FLIR E60 thermal infrared camera was utilized to capture a total of 5,278 image samples. The images consist of four main categories of actions, namely falling, sitting, walking, and standing. Additional subclasses for falling and sitting was also included, in which for falling includes fallen on the ground and fallen on the desk, whilst for sitting were sitting on a chair with and without a desk. The 5-fold cross-validation was employed. The images were split into training, validation, and testing phase with 28,844, 1,255, and 1,179 samples image, respectively. The proposed deep learning model could achieve a CA of 87.44% against the Histogram of Oriented Gradients (HOG)-SVM pipeline, which attained a CA of 85.9%.

[Lee, Yoon and Cho \(2017\)](#) evaluated the efficacy of CNN against a conventional machine learning model, *i.e.*, Random Forest (RF), in the classification of HAR from data gathered *via* a tri-axial accelerometer. Five subjects participated in the study where three activities, namely staying still, running, and walking was recorded *via* Nexus 6P Huawei smartphones. The raw x, y and z signals were transformed into a single magnitude vector data (1D) with two-size feature vectors of 10 and 20 s denoted as Feature10 and Feature20, respectively. The RF model was evaluated *via* MATLAB whilst the CNN model *via* TensorFlow. It was shown from the study that the proposed 1D CNN achieved a CA of 91.32% for Feature10 and 92.71% for Feature20, respectively outperforming the

conventional RF model, which achieved a CA of 85.72% and 89.10% for Feature10 and Feature20, respectively.

Conversely, [Rangasamy et al. \(2020\)](#) proposed the employment of the Transfer Learning paradigm for hockey activity recognition. The authors employed a pre-trained CNN, specifically VGG16, to extract features from four main hockey activities, namely free hit, goal, penalty corner and long corner, respectively. The dataset collected from International Hockey Federation (FIH) YouTube videos of the 2018 Hockey World Cup with a resolution of $1,280 \times 720$. A total of 400 frames been used and resized to 224×224 pixels. Different hyperparameters, namely the number of epochs with a different number of batch training of 100, 200 and 300, were fine-tuned at the fully connected layer utilizing a 10-fold cross-validation technique. The preliminary result showed that the model with 300 epochs achieved the highest CA of 98% then followed by 200 and 100 epochs with CA of 95% and 90%, respectively.

In relation to skateboarding, [Groh, Kautz & Schuldhaus \(2015\)](#) proposed the employment of machine learning in classifying six different skateboarding tricks using four machine learning classifiers, namely Naïve Bayes (NB), Partial Decision Tree (PART), Support Vector Machine with radial kernel basis kernel (RB-SVM) and k-Nearest Neighbor (kNN). Seven experienced male skateboarders between the age of 21 to 29 participated in the study. The data was gathered *via* an Inertial Measurement Unit (IMU) that was placed behind the front truck of the board. It was shown from the study that RB-SVM, as well as NB models, could achieve a CA of 97.8%. In an extended study, [Groh et al. \(2017\)](#) then enhanced the proposal by classifying thirteen classes for eleven skateboarding tricks, one class for bails and one class for other detected events with no trick. In this enhancement, the authors evaluated five classifiers which are NB, Random Forest (RF), Linear Support Vector Machine (LSVM), RB-SVM and kNN. It was shown from the study that the RB-SVM model was the best model amongst the models evaluated with a CA of 89.1%.

In a much earlier study, [Anlauff et al. \(2010\)](#) evaluated the efficacy of Linear Discriminant Analysis (LDA) in classifying three classes of two fundamental skateboarding tricks, *i.e.*, Ollie and Ollie180 and one for no trick event. One skateboarder participated in the study, in which the skateboarder executed the tricks repeatedly for 20 times. A 10-fold cross-validation technique was employed on the training dataset, and it was shown from the study that an average CA of 89.33% was reported. Conversely, in a recent investigation, [Corrêa et al. \(2017\)](#) develop an Artificial Neural Network (ANN) model in classifying five skateboarding tricks. Interestingly, the authors artificially generated the dataset based on the acceleration data reported in [Groh, Kautz & Schuldhaus \(2015\)](#). A single hidden layer architecture was employed with 28 hidden neurons with a tan-sigmoid activation function trained with the Scaled Conjugate Gradient learning algorithm on a dataset that is split with an 80:20 ratio for training and validation. The study evaluated the model on data attained from the Z-axis only and the combination of XYZ axes. It was shown that the ANN developed for the Z-axis could achieve a CA of 98.7%

In a more recent study, [Abdullah et al. \(2020\)](#) inspected six machine learning models, *viz.* SVM, kNN, ANN, Logistic Regression (LR), RF and NB in classifying five

skateboarding tricks. An amateur skateboarder participated in the study in which the accelerometer and gyro data along the XYZ axes were acquired. Different statistical time-domain features were extracted from all the signals, *i.e.*, mean, skewness, kurtosis, peak to peak, root mean square and standard deviation. A CA of 95% was reported to be attained *via* the features extracted on the LR and NB model. It could be seen from the limited literature available with regards to the employment of machine learning in classifying skateboarding tricks demonstrated commendable classification accuracy. Nevertheless, it is also evident from the literature reported that the use of CNN could mitigate the shortcomings of conventional machine learning models, particularly in acquiring significant features that would consequently yield better predictions. Therefore, this paper aims to address the gap by leveraging the use of a variation of the CNN model, *i.e.*, the transfer learning model with its fully connected layer replaced with an optimized SVM model towards the classification of skateboarding tricks. The effect of input image transformation towards classification accuracy is also investigated.

METHODOLOGY

Data collection

The skateboarding tricks signals were acquired through an instrumented inertial measurement unit (IMU) device developed. The device is embedded with an MPU6050 sensor, a Bluetooth 2.0 module, a microcontroller and a 3.7 V Lithium Polymer rechargeable battery. The device is paired together with a riser pad on the other side of the truck to give balance to the skateboard. Specifically, the device is mounted behind the front truck, and the pair are fixed with a nylon lock nuts (nyloc). The whole case and riser pad are made from ABS material printed *via* Zortrax M200 3D printer. The design of the IMU device is inspired by the works carried out by (Groh, Kautz & Schuldhaus, 2015). Figures 1 and 2 depicts the placement of the instrumented device from the frontal and rear view.

The chosen skateboarding tricks in the present investigation are Ollie (O), Nollie Frontside Shuvit (NFS), Frontside 180 (FS180), Pop Shove-It (PS) and Kickflip (KF). The selection of the tricks is non-trivial as it is the most common moves that are executed by a skateboarder in any competition (Groh, Kautz & Schuldhaus, 2015; Corrêa et al., 2017). The skateboarding tricks were performed by six 20 ± 7 years old amateur skateboarder with at least 5 years of experience and been executed successfully five times per trick. Universiti Malaysia Terengganu granted Ethical approval to carry out the study within and with its associated facilities (Ethical Application Ref: UMT/JKEPHMK/2021/53) whilst informed consent was obtained from the skateboarders participated in the present investigation.

Input signal image transformation

In general, there were six simultaneous different raw signals data collected from the device per successful trick. The raw signals data were solely taken from the IMU embedded in the device. They are x-axis linear acceleration (aX), y-axis linear acceleration (aY), z-axis linear acceleration (aZ), x-axis angular acceleration (gX), y-axis acceleration (gY), and



Figure 1 The position of the instrumented device on the skateboard (front view).


Full-size  DOI: 10.7717/peerj-cs.680/fig-1



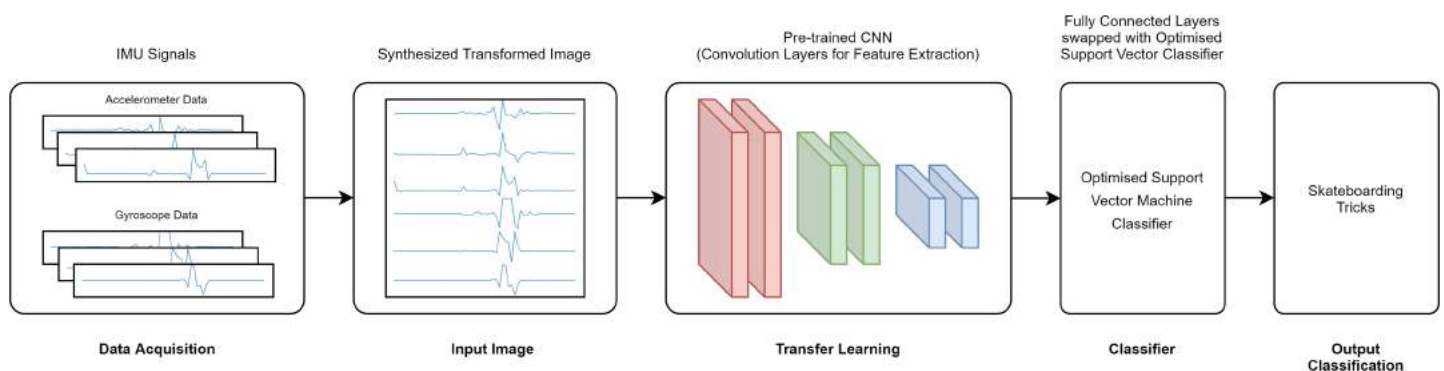
Figure 2 The placement of the instrumented device on the deck (rear view).

Full-size  DOI: 10.7717/peerj-cs.680/fig-2

z-axis acceleration (gZ). All these six raw signals data were synthesized into a single image representing one single skateboarding trick according to the default image size based on Table 1.

Table 1 Default size settings of the transfer learning models.

No.	Model	Flatten reshape	Input image	
			Height	Width
1	MobileNet	7 * 7 * 1,024	224	224
2	MobileNetV2	7 * 7 * 1,280	224	224
3	NasNetLarge	11 * 11 * 4,032	331	331
4	NasNetMobile	7 * 7 * 1,056	224	224
5	ResNet101	7 * 7 * 2,048	224	224
6	ResNet101V2	7 * 7 * 2,048	224	224


Figure 3 RAW-TL-optimized SVM pipeline.

[Full-size !\[\]\(d66ff64371a51729ac8c1cdaa685ba6f_img.jpg\) DOI: 10.7717/peerj-cs.680/fig-3](https://doi.org/10.7717/peerj-cs.680/fig-3)

Two input image transformation were chosen for this study. The basic input image was the raw transformation (RAW), where it is directly synthesized from the six raw signals stacked in a single image. The second input image transformation was a scalogram image transformed *via* Continuous Wavelet Transform (CWT). CWT is the representation of the time-frequency domain of a set of signals that have been demonstrated to be effective for non-stationary signals (Qassim *et al.*, 2012). The resolution represented through the CWT algorithm has been reported to be beneficial owing to the exploitation of the small scale of high frequencies and large scale of low frequencies (Türk & Özerdem, 2019). Moreover, it has also been reported to provide a better representation of the arrangement of the frequency domain features as compared to Fourier Transforms. The mother wavelet that was used in this research is the Morlet Wavelet. Morlet wavelet is the multiplication of the complex exponential and Gaussian window. The Morlet algorithm gives an innate link between frequency and time domain to distinguish the signals acquired *via* Fourier Transform.

Feature extraction: transfer learning

A total of six transfer learning models was used for this study. The proposed architecture investigated is depicted in Fig. 3 (RAW) and Fig. 4 (CWT), respectively. This study exploits the use of three families of pre-trained CNN models, *i.e.*, the MobileNet, NasNet and

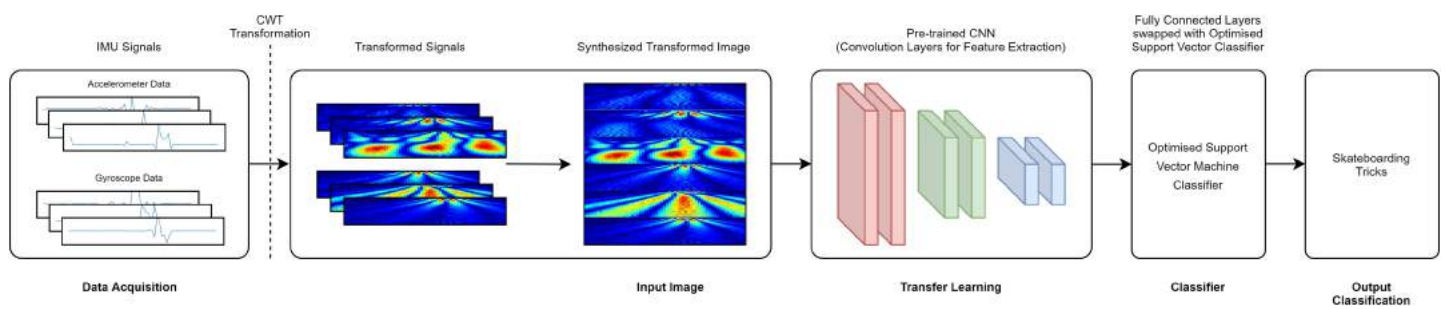


Figure 4 CWT-TL-optimized SVM pipeline.

Full-size DOI: 10.7717/peerj-cs.680/fig-4

Table 2 Hyper-parameter description and range of value.

No.	Hyper-parameter	Description	Range
1	Kernel	Type of kernel	'linear', 'poly', 'rbf'
2	Degree	Degree of polynomial function (only applicable for 'poly' kernel)	2–6
3	Gamma, γ	Kernel coefficient	0.1, 1, 10, 100
4	C	Strength of the regularization	0.01, 0.1, 1, 10, 100

ResNet families. The rationale of employing transfer learning (TL) models is to reduce the model development time as the CNN models are not required to be built from scratch (Amanpour & Erfanian, 2013; Chronopoulou, Baziotis & Potamianos, 2019; Mahendra Kumar et al., 2021). A departure from conventional means of using such models is that the present study replaces the fully connected layers that is often referred to dense layers with a conventional machine learning model, i.e., SVM. Hence, the convolution layers of the transfer learning models utilized are used exclusively for feature extraction purpose. The list of the transfer learning model and their respective parameters are tabulated in Table 1.

Classifier: support vector machine

The features extracted from the different transfer learning models based on the input images are fed into a variety of SVM models. The variation is based on the different hyperparameters evaluated, namely, the type of **kernel**, viz linear, radial basis function (rbf) and polynomial (poly); the **degree** of the polynomial function, which was varied between two to six; (2–6); the kernel coefficient or **gamma**, γ (0.1, 1, 10, 100); and strength of the regularization, **C** (0.01, 0.1, 1, 10, 100), respectively. It is worth noting that the γ parameter affects the rbf and poly-based SVM models. The loss function of the SVM classifier built-in in the scikit-learn package is the squared-hinge loss function. Table 2 lists the hyperparameters evaluated. The dataset was split into a ratio of 60:20:20 for training, testing and validation, respectively, on the 150 synthesized images per input transformation. The hyperparameters of the SVM models were tuned *via* the grid-search algorithm *via* the three-fold cross-validation technique on the training dataset. A total of 125 SVM models were developed per transfer learning model and per image input.

		Predicted Label	
		Positive (1)	Negative (0)
True Label	Positive (1)	TP	FN
	Negative (0)	FP	TN

Figure 5 An example of a binary confusion matrix.

Full-size [DOI: 10.7717/peerj-cs.680/fig-5](https://doi.org/10.7717/peerj-cs.680/fig-5)

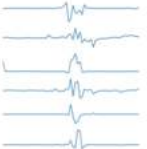

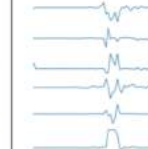
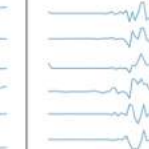
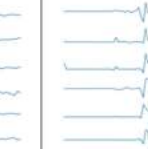
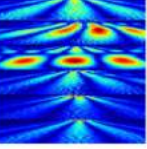
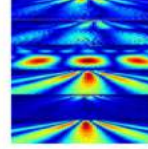
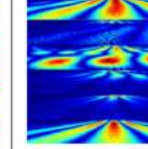
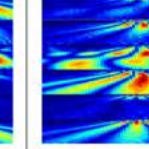
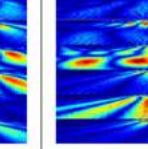
Input Image Type	RAW				
Image Example					
Trick	Ollie	NFS	FS180	PS	KF
Input Image Type	CWT				
Image Example					
Trick	Ollie	NFS	FS180	PS	KF

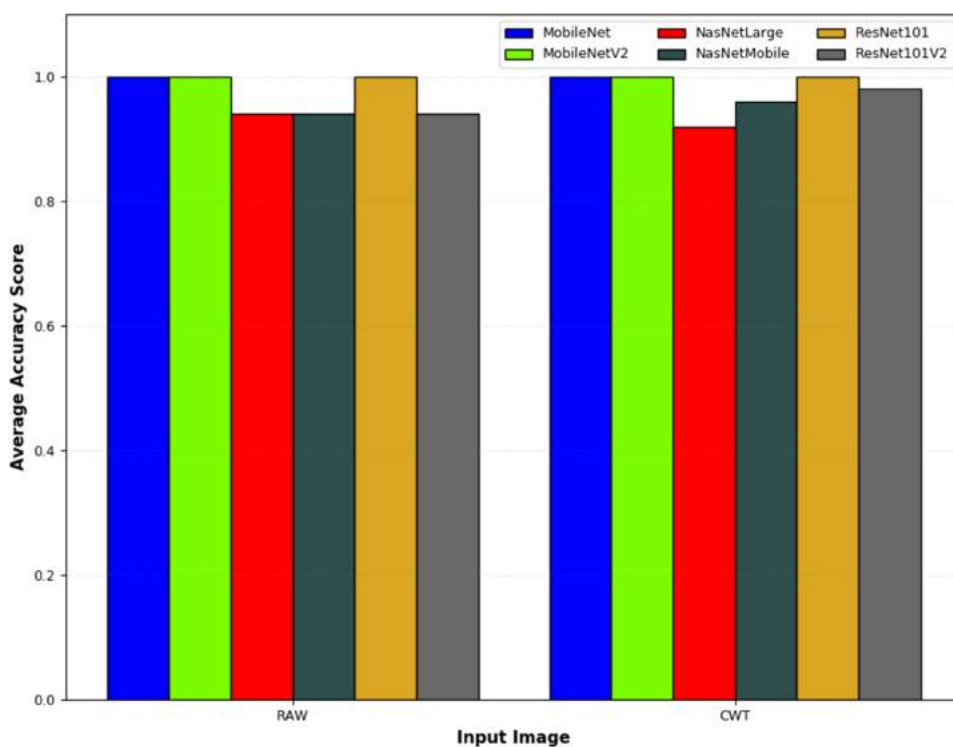
Figure 6 Example of the input image transformation.

Full-size [DOI: 10.7717/peerj-cs.680/fig-6](https://doi.org/10.7717/peerj-cs.680/fig-6)

Therefore, 1,500 pipelines (which consist of input image-transfer learning model-tuned SVM model) were evaluated in the present investigation. It is worth noting that the overall pipeline was evaluated on an Intel Core i7 4800MQ @ 2.70 GHz with 8 GB DDR3 800

Table 3 Classification accuracy of transfer learning model with SVM for the different input image.

No.	Input image	Model	Accuracy			
			Train	Validate	Test	Average
1	RAW	MobileNet	1.00	1.00	1.00	1.00
2	RAW	MobileNetV2	1.00	1.00	1.00	1.00
3	RAW	NasNetLarge	1.00	0.88	1.00	0.94
4	RAW	NasNetMobile	1.00	0.88	1.00	0.94
5	RAW	ResNet101	1.00	1.00	1.00	1.00
6	RAW	ResNet101V2	1.00	0.88	1.00	0.94
7	CWT	MobileNet	1.00	1.00	1.00	1.00
8	CWT	MobileNetV2	1.00	1.00	1.00	1.00
9	CWT	NasNetLarge	1.00	0.88	0.96	0.92
10	CWT	NasNetMobile	1.00	1.00	0.92	0.96
11	CWT	ResNet101	1.00	1.00	1.00	1.00
12	CWT	ResNet101V2	1.00	1.00	0.96	0.98


Figure 7 The average classification accuracy of different pipelines developed for the different input image evaluated.

[Full-size !\[\]\(a03a7eb2f4046e1d3c76772003e549ea_img.jpg\) DOI: 10.7717/peerj-cs.680/fig-7](https://doi.org/10.7717/peerj-cs.680/fig-7)

MHz RAM and an Intel HD Graphics 4,600 via Spyder 3.3.6, a Python IDE running on Python 3.7 along with associated libraries, *i.e.*, scikit-learn 0.22.1 and Keras 2.3.1: Tensorflow 1.14.0.

Table 4 Computational time between the evaluated pipelines.

No.	Input image	Model	Time (s)		
			Train	Validate	Test
1	RAW	MobileNet	0.5000	0.1250	0.1250
2	RAW	MobileNetV2	0.1563	0.1563	0.1563
3	RAW	NasNetLarge	6.0000	1.5938	1.4375
4	RAW	NasNetMobile	0.4844	0.1250	0.1250
5	RAW	ResNet101	1.0063	0.2656	0.2500
6	RAW	ResNet101V2	1.0938	0.2653	0.2653
7	CWT	MobileNet	0.4688	0.1094	0.1094
8	CWT	MobileNetV2	0.6094	0.1563	0.1563
9	CWT	NasNetLarge	5.7188	1.4844	1.4375
10	CWT	NasNetMobile	0.5000	0.1406	0.1250
11	CWT	ResNet101	1.0156	0.2656	0.2500
12	CWT	ResNet101V2	1.0781	0.2656	0.2656

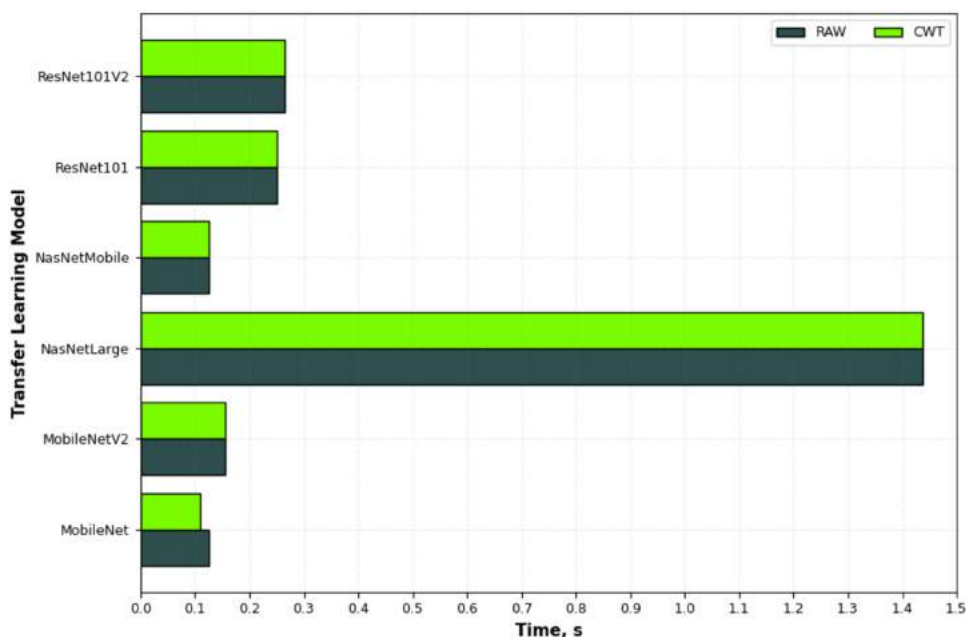


Figure 8 Prediction time of the different pipelines developed for the different input image evaluated.

Full-size DOI: 10.7717/peerj-cs.680/fig-8

Linear,

$$y_i(w \cdot x_i + b) \geq 1 - \varepsilon_i \quad i = 1, \dots, m \quad (1)$$

Polynomial, 'poly'

$$K(x, x') = (x \cdot x + c')^q \quad (2)$$

Table 5 Precision of transfer learning model with SVM for the different input image.

No.	Input image	Model	Precision		
			Train	Validate	Test
1	RAW	MobileNet	1.00	1.00	1.00
2	RAW	MobileNetV2	1.00	1.00	1.00
3	RAW	NasNetLarge	1.00	0.89	1.00
4	RAW	NasNetMobile	1.00	0.93	1.00
5	RAW	ResNet101	1.00	1.00	1.00
6	RAW	ResNet101V2	1.00	0.93	1.00
7	CWT	MobileNet	1.00	1.00	1.00
8	CWT	MobileNetV2	1.00	1.00	1.00
9	CWT	NasNetLarge	1.00	0.91	0.97
10	CWT	NasNetMobile	1.00	1.00	0.93
11	CWT	ResNet101	1.00	1.00	1.00
12	CWT	ResNet101V2	1.00	1.00	0.97

Table 6 Recall of transfer learning model with SVM for different input image.

No.	Input image	Model	Recall		
			Train	Validate	Test
1	RAW	MobileNet	1.00	1.00	1.00
2	RAW	MobileNetV2	1.00	1.00	1.00
3	RAW	NasNetLarge	1.00	0.88	1.00
4	RAW	NasNetMobile	1.00	0.88	1.00
5	RAW	ResNet101	1.00	1.00	1.00
6	RAW	ResNet101V2	1.00	0.88	1.00
7	CWT	MobileNet	1.00	1.00	1.00
8	CWT	MobileNetV2	1.00	1.00	1.00
9	CWT	NasNetLarge	1.00	0.88	0.96
10	CWT	NasNetMobile	1.00	1.00	0.92
11	CWT	ResNet101	1.00	1.00	1.00
12	CWT	ResNet101V2	1.00	1.00	0.96

Radial basis function, 'rbf'

$$K(x, x') = e^{-\frac{\|x - x'\|^2}{2\sigma^2}} \quad (3)$$

where w is the weighting vector, b is the constant, and ϵ is the nonnegative slack variable.

Performance evaluation

In the present study, a number of evaluation metrics were used. The accuracy score represents the accuracy of the model in predicting the corresponding value to the true value. The value ranges from zero to one where zero indicates a total misclassification

Table 7 F1-score of transfer learning model with SVM for different input image.

No.	Input image	Model	F1-score		
			Train	Validate	Test
1	RAW	MobileNet	1.00	1.00	1.00
2	RAW	MobileNetV2	1.00	1.00	1.00
3	RAW	NasNetLarge	1.00	0.88	1.00
4	RAW	NasNetMobile	1.00	0.87	1.00
5	RAW	ResNet101	1.00	1.00	1.00
6	RAW	ResNet101V2	1.00	0.87	1.00
7	CWT	MobileNet	1.00	1.00	1.00
8	CWT	MobileNetV2	1.00	1.00	1.00
9	CWT	NasNetLarge	1.00	0.86	0.96
10	CWT	NasNetMobile	1.00	1.00	0.92
11	CWT	ResNet101	1.00	1.00	1.00
12	CWT	ResNet101V2	1.00	1.00	0.96

transpired whilst one indicates that no misclassification transpired. It is commonly used to evaluate the accuracy of a multiclass classification problem (Foody & Mathur, 2004) and one of the most straightforward and simplest measures (Sokolova & Lapalme, 2009; Flach, 2019). This score also can be interpreted through the confusion matrix. Figure 5 illustrates an example of a binary class confusion matrix. True Positive (TP) is defined as a positive sample correctly predicted as positive. True Negative (TN) is the negative sample correctly predicted as negative. When the positive sample is incorrectly predicted as negative, it is counted toward the False Negative (FN). Conversely, False Positives (FP) is a negative sample incorrectly predicted as positive. The precision measures the percentage of correct positive predictions over the cumulative number of positive predictions. The sensitivity (often known as recall) is the number of true positive predictions divided by the sum of true positives as well as the false negatives (Vijay Anand & Shantha Selvakumari, 2019). The F1-score discloses the balance between the recall and the precision values. Whilst the specificity is essentially the proportion of actual negative values, which is forecasted as the true negative. It is also worth noting at this juncture, in the event that a tie between the classification accuracy transpire between the best pipelines, the determining factor would be based on the computational time of the pipelines.

EXPERIMENTAL RESULTS AND DISCUSSION

Figure 6 depicts an example of the synthesized input images per skateboarding tricks with respect to RAW and CWT, respectively. Table 3 reports the accuracy of the evaluated pipelines. It could be observed that both RAW and CWT input transformation could yield an accuracy of 100% on all train, test and validation dataset for both MobileNet and ResNet101 families by utilizing the optimized SVM model. The optimized hyperparameters for the pipelines are the linear kernel-based SVM model with a C and gamma, γ value of 0.01 and 0.1, respectively. A similar performance is noticed for the

Table 8 Specificity of transfer learning model with SVM for RAW input image.

No.	Input image	Model	Trick	Specificity		
				Train	Validate	Test
1	RAW	MobileNet	Ollie	1.00	1.00	1.00
			NFS	1.00	1.00	1.00
			FS180	1.00	1.00	1.00
			PS	1.00	1.00	1.00
			KF	1.00	1.00	1.00
2	RAW	MobileNetV2	Ollie	1.00	1.00	1.00
			NFS	1.00	1.00	1.00
			FS180	1.00	1.00	1.00
			PS	1.00	1.00	1.00
			KF	1.00	1.00	1.00
3	RAW	NasNetLarge	Ollie	1.00	0.95	1.00
			NFS	1.00	1.00	1.00
			FS180	1.00	0.95	1.00
			PS	1.00	0.95	1.00
			KF	1.00	1.00	1.00
4	RAW	NasNetMobile	Ollie	1.00	1.00	1.00
			NFS	1.00	1.00	1.00
			FS180	1.00	1.00	1.00
			PS	1.00	0.85	1.00
			KF	1.00	1.00	1.00
5	RAW	ResNet101	Ollie	1.00	1.00	1.00
			NFS	1.00	1.00	1.00
			FS180	1.00	1.00	1.00
			PS	1.00	1.00	1.00
			KF	1.00	1.00	1.00
6	RAW	ResNet101V2	Ollie	1.00	1.00	1.00
			NFS	1.00	1.00	1.00
			FS180	1.00	1.00	1.00
			PS	1.00	0.85	1.00
			KF	1.00	1.00	1.00

RAW-ResNet101-optimized SVM as well as the CWT-ResNet101-optimized SVM models. Figure 7 depicts the average accuracy of the pipelines evaluated. Therefore, the determining factor for which pipeline is the best would be the computational time. As shown in Table 4, based on the computational time, the CWT-MobileNet-optimized SVM is deemed to be the best pipeline owing to the reduced computational time taken as compared to the other models evaluated. Figure 8 illustrates the aforesaid prediction time.

The precision, recall, F1-score, specificity on both input images across different evaluated pipelines are tabulated in Tables 5–9, respectively. The confusion matrix of

Table 9 Specificity of transfer learning model with SVM for CWT input image.

No.	Input image	Model	Trick	Specificity		
				Train	Validate	Test
1	CWT	MobileNet	Ollie	1.00	1.00	1.00
			NFS	1.00	1.00	1.00
			FS180	1.00	1.00	1.00
			PS	1.00	1.00	1.00
			KF	1.00	1.00	1.00
2	CWT	MobileNetV2	Ollie	1.00	1.00	1.00
			NFS	1.00	1.00	1.00
			FS180	1.00	1.00	1.00
			PS	1.00	1.00	1.00
			KF	1.00	1.00	1.00
3	CWT	NasNetLarge	Ollie	1.00	0.95	1.00
			NFS	1.00	1.00	0.95
			FS180	1.00	1.00	1.00
			PS	1.00	0.90	1.00
			KF	1.00	1.00	1.00
4	CWT	NasNetMobile	Ollie	1.00	1.00	1.00
			NFS	1.00	1.00	0.95
			FS180	1.00	1.00	0.95
			PS	1.00	1.00	1.00
			KF	1.00	1.00	1.00
5	CWT	ResNet101	Ollie	1.00	1.00	1.00
			NFS	1.00	1.00	1.00
			FS180	1.00	1.00	1.00
			PS	1.00	1.00	1.00
			KF	1.00	1.00	1.00
6	CWT	ResNet101V2	Ollie	1.00	1.00	1.00
			NFS	1.00	1.00	0.95
			FS180	1.00	1.00	1.00
			PS	1.00	1.00	1.00
			KF	1.00	1.00	1.00

the best pipeline, *i.e.*, the CWT-MobileNet-optimized SVM on the test dataset, is depicted in Fig. 9. The present study has demonstrated that through the proposed pipeline, a better classification accuracy could be achieved as compared to the conventional means reported in the literature, particularly with regards to the classification of skateboarding tricks. The encouraging results reported suggests that the proposed pipeline could be beneficial in providing an objective-based judgment in The findings of the present investigation are in agreement with other studies that have employed such a technique in different applications, for instance, [Lee, Yoon and Cho \(2017\)](#), [Rangasamy et al. \(2020\)](#) as well as [Mahendra Kumar et al. \(2021\)](#). Nonetheless, it is worth noting that the efficacy of

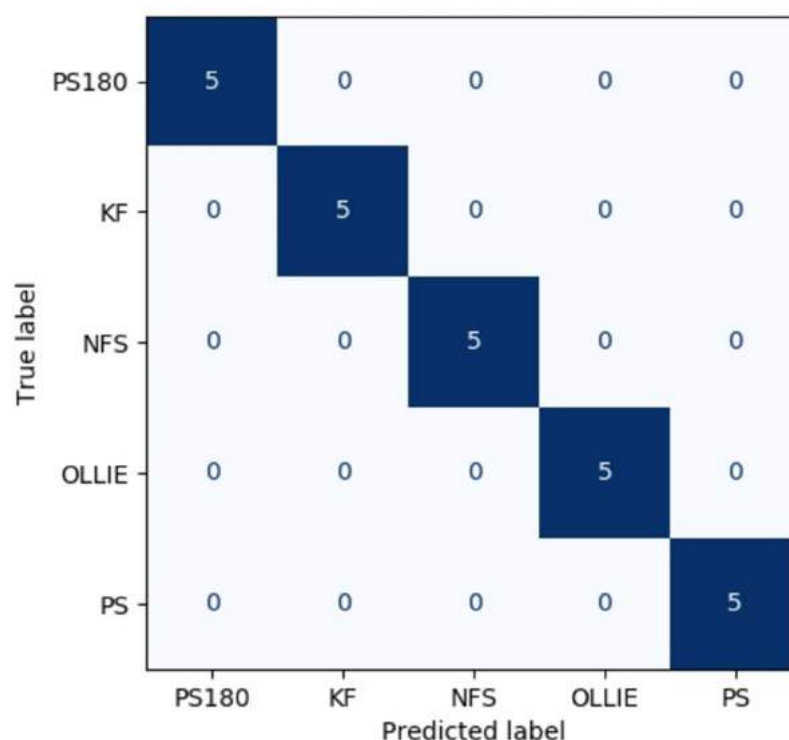


Figure 9 Confusion matrix of the best pipeline on the test dataset.

Full-size DOI: 10.7717/peerj-cs.680/fig-9

the pipelines is highly dependent on the dataset utilized, and the performance may vary. Future studies shall explore the use of different feature input transformation, other transfer learning as well as machine learning models.

CONCLUSION

The present study investigated the efficacy of different transfer learning pipeline towards the classification of skateboarding tricks. It was shown that the study that best pipeline identified is the CWT-MobileNet-optimized SVM as it could yield the fastest computational time. It could be seen that the convolution part of the pre-trained CNN models or transfer learning models could effortlessly extract significant features, regardless of the input image provided. The findings are non-trivial in the realization of an objective-based judgement in a skateboarding competition. Future studies shall evaluate other types of input image transformation methods and transfer learning models as well as their effect towards other classifiers that have yet been investigated in the present study.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work is funded by the Ministry of Education, Malaysia through the Fundamental Research Grant Scheme (FRGS/1/2019/TK03/UMP/02/6) and Universiti Malaysia Pahang

(RDU1901115). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

Ministry of Education, Malasiya: FRGS/1/2019/TK03/UMP/02/6.

Universiti Malaysia Pahang: RDU1901115.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Muhammad Amirul Abdullah conceived and designed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Muhammad Ar Rahim Ibrahim performed the experiments, prepared figures and/or tables, and approved the final draft.
- Muhammad Nur Aiman Shapiee performed the experiments, prepared figures and/or tables, and approved the final draft.
- Muhammad Aizzat Zakaria performed the computation work, prepared figures and/or tables, and approved the final draft.
- Mohd Azraai Mohd Razman conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Rabi Muazu Musa performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Noor Azuan Abu Osman performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Anwar P.P. Abdul Majeed conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Ethics

The following information was supplied relating to ethical approvals (*i.e.*, approving body and any reference numbers):

Universiti Malaysia Terengganu granted Ethical approval to carry out the study within and with its associated facilities (Ethical Application Ref: UMT/JKEPHMK/2021/53).

Data Availability

The following information was supplied regarding data availability:

Raw data and sample code are available in the [Supplemental Files](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.680#supplemental-information>.

REFERENCES

- Abdullah MA, Ibrahim MAR, Shapiee MNAB, Mohd Razman MNA, Musa RM, Abdul Majeed APP. 2020.** The classification of skateboarding trick manoeuvres through the integration of IMU and machine learning. In: *Lecture Notes in Mechanical Engineering*. Berlin: Springer, 67–74.
- Akula A, Shah AK, Ghosh R. 2018.** ScienceDirect deep learning approach for human action recognition in infrared images. *Cognitive Systems Research* **50**(8):146–154 DOI [10.1016/j.cogsys.2018.04.002](https://doi.org/10.1016/j.cogsys.2018.04.002).
- Amanpour B, Erfanian A. 2013.** Classification of brain signals associated with imagination of hand grasping, opening and reaching by means of wavelet-based common spatial pattern and mutual information. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Piscataway: IEEE, 2224–2227.
- Anlauff J, Weitnauer E, Lehnhardt A, Schirmer S, Zehe S, Tonekaboni K. 2010.** A method for outdoor skateboarding video games. In: *ACM International Conference Proceeding Series*. 40–43.
- Chen Y, Xue Y. 2015.** A deep learning approach to human activity recognition based on single accelerometer. In: *2015 IEEE International Conference on Systems, Man, and Cybernetics*. Piscataway: IEEE, 1488–1492.
- Chronopoulou A, Baziotis C, Potamianos A. 2019.** An embarrassingly simple approach for transfer learning from pretrained language models. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Vol. 1. 2089–2095 DOI [10.18653/v1/n19-1213](https://doi.org/10.18653/v1/n19-1213).
- Corrêa NK, de Lima JCM, Russomano T, dos Santos MA. 2017.** Development of a skateboarding trick classifier using accelerometry and machine learning. *Research on Biomedical Engineering* **33**(4):362–369 DOI [10.1590/2446-4740.04717](https://doi.org/10.1590/2446-4740.04717).
- Flach P. 2019.** Performance evaluation in machine learning: the good, the bad, the ugly, and the way forward. *Proceedings of the AAAI Conference on Artificial Intelligence* **33**:9808–9814 DOI [10.1609/aaai.v33i01.33019808](https://doi.org/10.1609/aaai.v33i01.33019808).
- Foody GM, Mathur A. 2004.** A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing* **42**(6):1335–1343 DOI [10.1109/TGRS.2004.827257](https://doi.org/10.1109/TGRS.2004.827257).
- Groh BH, Fleckenstein M, Kautz T, Eskofier BM. 2017.** Classification and visualization of skateboard tricks using wearable sensors. *Pervasive and Mobile Computing* **40**(6):42–55 DOI [10.1016/j.pmcj.2017.05.007](https://doi.org/10.1016/j.pmcj.2017.05.007).
- Groh BH, Kautz T, Schuldhaus D. 2015.** IMU-based trick classification in Skateboarding. In: *KDD Workshop on Large-Scale Sports Analytics*.
- Lee SM, Yoon SM, Cho H. 2017.** Human activity recognition from accelerometer data using convolutional neural network. In: *2017 IEEE International Conference on Big Data and Smart Computing (BigComp)*. Piscataway: IEEE, 131–134 DOI [10.1109/BIGCOMP.2017.7881728](https://doi.org/10.1109/BIGCOMP.2017.7881728).
- Mahendra Kumar JL, Rashid M, Muazu Musa R, Mohd Razman MA, Sulaiman N, Jailani R, Abdul Majeed APP. 2021.** The classification of EEG-based winking signals: a transfer learning and random forest pipeline. *PeerJ* **9**(2):e11182 DOI [10.7717/peerj.11182](https://doi.org/10.7717/peerj.11182).
- Pappalardo A. 2014.** How do you judge a skateboard contest? Available at <http://www.jenkemmag.com/home/2014/04/07/how-do-you-judge-a-skateboard-contest/>.
- Qassim YT, Cutmore T, James D, Rowlands D. 2012.** FPGA implementation of Morlet continuous wavelet transform for EEG analysis. In: *International Conference on Computer and Communication Engineering*. 59–64.

- Rangasamy K, As'ari MA, Rahmad NA, Ghazali NF. 2020.** Hockey activity recognition using pre-trained deep learning model. *ICT Express* **6(3)**:170–174 DOI [10.1016/j.ict.2020.04.013](https://doi.org/10.1016/j.ict.2020.04.013).
- Sokolova M, Lapalme G. 2009.** A systematic analysis of performance measures for classification tasks. *Information Processing and Management* **45(4)**:427–437 DOI [10.1016/j.ipm.2009.03.002](https://doi.org/10.1016/j.ipm.2009.03.002).
- Stein M, Janetzko H, Lamprecht A, Breitzkreutz T, Zimmermann P, Goldlücke B, Schreck T, Andrienko G, Grossniklaus M, Keim DA. 2018.** Bring it to the pitch: combining video and movement data to enhance team sport analysis. *IEEE Transactions on Visualization and Computer Graphics* **24(1)**:13–22 DOI [10.1109/TVCG.2017.2745181](https://doi.org/10.1109/TVCG.2017.2745181).
- Türk Ö, Özerdem MS. 2019.** Epilepsy detection by using scalogram based convolutional neural network from EEG signals. *Brain Sciences* **9(5)**:115 DOI [10.3390/brainsci9050115](https://doi.org/10.3390/brainsci9050115).
- Vijay Anand S, Shantha Selvakumari R. 2019.** Noninvasive method of epileptic detection using DWT and generalized regression neural network. *Soft Computing* **23(8)**:2645–2653 DOI [10.1007/s00500-018-3630-y](https://doi.org/10.1007/s00500-018-3630-y).

Two-stage training algorithm for AI robot soccer

Taeyoung Kim¹, Luiz Felipe Vecchietti¹, Kyujin Choi¹, Sanem Sariel² and Dongsoo Har¹

¹Cho Chun Shik Graduate School of Green Transportation, Korea Advanced Institute of Science and Technology, Daejeon, South Korea

²Department of Computer Engineering, Istanbul Technical University, Istanbul, Turkey

ABSTRACT

In multi-agent reinforcement learning, the cooperative learning behavior of agents is very important. In the field of heterogeneous multi-agent reinforcement learning, cooperative behavior among different types of agents in a group is pursued. Learning a joint-action set during centralized training is an attractive way to obtain such cooperative behavior; however, this method brings limited learning performance with heterogeneous agents. To improve the learning performance of heterogeneous agents during centralized training, two-stage heterogeneous centralized training which allows the training of multiple roles of heterogeneous agents is proposed. During training, two training processes are conducted in a series. One of the two stages is to attempt training each agent according to its role, aiming at the maximization of individual role rewards. The other is for training the agents as a whole to make them learn cooperative behaviors while attempting to maximize shared collective rewards, e.g., team rewards. Because these two training processes are conducted in a series in every time step, agents can learn how to maximize role rewards and team rewards simultaneously. The proposed method is applied to 5 versus 5 AI robot soccer for validation. The experiments are performed in a robot soccer environment using Webots robot simulation software. Simulation results show that the proposed method can train the robots of the robot soccer team effectively, achieving higher role rewards and higher team rewards as compared to other three approaches that can be used to solve problems of training cooperative multi-agent. Quantitatively, a team trained by the proposed method improves the score concede rate by 5% to 30% when compared to teams trained with the other approaches in matches against evaluation teams.

Submitted 1 April 2021
Accepted 26 August 2021
Published 17 September 2021

Corresponding author
Dongsoo Har, dshar@kaist.ac.kr

Academic editor
Alma Alanis

Additional Information and
Declarations can be found on
page 19

DOI 10.7717/peerj-cs.718

© Copyright
2021 Kim et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Agents and Multi-Agent Systems, Algorithms and Analysis of Algorithms, Artificial Intelligence, Robotics

Keywords Multi-agent reinforcement learning, Heterogeneous agents, Centralized training, Deep learning, Robotics

INTRODUCTION

Recently, deep reinforcement learning (DRL) has been widely applied to deterministic games (*Silver et al., 2018*), video games (*Mnih et al., 2015; Mnih et al., 2016; Silver et al., 2016*), sensor networks (*Kim et al., 2020*), and complex robotic tasks (*Andrychowicz et al., 2017; Hwangbo et al., 2019; Seo et al., 2019; Vecchietti et al., 2020; Vecchietti, Seo & Har, 2020*). Despite the breakthrough results achieved in the field of DRL, deep learning

in multi-agent environments that require both cooperation and competition is still challenging. Promising results have been for cooperative-competitive multi-agent games such as StarCraft (Vinyals et al., 2019) and Dota (Berner et al., 2019). For multi-agent problems such as multi-robot soccer (Liu et al., 2019), security (He, Dai & Ning, 2015; Klima, Tuyls & Oliehoek, 2016), traffic control (Chu et al., 2019; Zhang et al., 2019), and autonomous driving (Shalev-Shwartz, Shammah & Shashua, 2016; Sallab et al., 2017), non-stationarity, partial observability, multi-agent training schemes, and heterogeneity can be challenging issues (Nguyen, Nguyen & Nahavandi, 2020). To solve these challenges, multi-agent reinforcement learning (MARL) techniques (Lowe et al., 2017; Sunehag et al., 2017; Foerster et al., 2018; Vinyals et al., 2019; Liu et al., 2019; Samvelyan et al., 2019; Rashid et al., 2020) have been intensively investigated.

When using the MARL, several works have used the centralized training in decentralized execution (CTDE) framework (Lowe et al., 2017; Sunehag et al., 2017; Foerster et al., 2018; Rashid et al., 2020). In the CTDE framework, local observations of agents, global state of the environment, and joint-actions taken by the agents at each time step are available during training to the centralized policy network, while only the local observations of agents are available during execution. In other words, each agent selects its action, that is the output of a policy network, without considering the full information of the environment. To address the non-stationarity problem, multi-agent deep deterministic policy gradient (MADDPG) (Lowe et al., 2017) was proposed using a CTDE framework and the deep deterministic policy gradient (DDPG) actor-critic algorithm for continuous action spaces (Lillicrap et al., 2015). When cooperative behavior is to be achieved, representing that there is a cooperative reward that should be maximized by multiple agents, credit should be assigned accordingly to each agent based on its contribution. To address this problem, counterfactual multi-agent (COMA) (Foerster et al., 2018), value decomposition networks (VDN) (Sunehag et al., 2017), and monotonic value function factorization (QMIX) (Rashid et al., 2020) have been proposed, using the CTDE framework combined with value-based algorithms such as deep Q networks (DQN) (Mnih et al., 2013), deep recurrent Q networks (DRQN) (Hausknecht & Stone, 2015), and dueling Q networks (Wang et al., 2016).

In this paper, a novel training method for MARL of heterogeneous agents, in which each agent should choose its action in a decentralized manner, is proposed. The proposed method addresses how to provide an optimal policy and maximize the cooperative behavior among heterogeneous agents. To this end, during training, two training stages are conducted in a series. The first stage is for making each agent learn to maximize its individual role reward while executing its individual role. The second one is for making the agents as a whole learn cooperative behavior, aiming at the maximization of team reward. The proposed method is designed to be applied to MARL with heterogeneous agents in cooperative or cooperative-competitive scenarios. In this paper, a cooperative-competitive Artificial Intelligence (AI) robot soccer environment is used for experiments. The environment can be described in relation to 5 versus 5 robot soccer game described in Hong et al. (2021). In the robot soccer game, two teams of five robots capable of kick and jump behaviors compete against each other, similarly to the StarCraft, so the game can be seen as a micro-management problem. The policy for the proposed method and other methods

for comparisons are trained by using self-play ([Heinrich, Lanctot & Silver, 2015](#); [Lanctot et al., 2017](#); [Silver et al., 2017](#)). Self-play in a competitive environment is used so that the opponent team is kept at an appropriate level of difficulty at each training stage.

The main contributions of this paper are as follows

1. A framework for novel training method called two-stage heterogeneous centralized training (TSHCT) aiming at centralized training of heterogeneous agents is proposed. In the proposed method, there are two training stages that are conducted in a series. The first stage is responsible for training individual behaviors by maximizing individual role rewards. The second stage is for training cooperative behaviors by maximizing a shared collective reward.
2. Experiments are conducted to compare the performance of the proposed method with other baseline methods, COMA, VDN, and QMIX. The proposed method and the baseline methods are trained with self-play. To compare the results obtained from the experiments, total rewards (during training) and score/concede rates (against different opponent teams) are presented. From the comparisons, we will show better performance of the proposed method during game.
3. The proposed method aims at MARL with heterogeneous agents in cooperative and cooperative-competitive scenarios. For experiments, a cooperative-competitive AI robot soccer environment, where there are 5 robots with 3 different roles in each team (one goalkeeper, two defenders, and two forwards), is used.

The remainder of this paper is organized as follows. ‘Background’ presents the concept of the MARL, system modeling, and other methods which are used as baselines for comparisons in the experiments. ‘Proposed Method’ introduces the proposed method in details. ‘Simulation Results’ presents the simulation environment, ablation studies, and game results of the AI robot soccer. ‘Conclusion’ concludes this paper.

BACKGROUND

In this section, the mathematical modeling of the proposed method is presented. Also, other methods for cooperative MARL using the CTDE framework are presented.

System modeling

The cooperative-competitive multi-agent problem, specifically applied in this paper to AI robot soccer, is modeled as a decentralized partially observable Markov decision process (Dec-POMDP) ([Oliehoek & Amato, 2016](#)) that each agent has its own observation of the environment. The Dec-POMDP can be formulated by an 8-tuple $G = \langle S, U, P, r, Z, O, n, \gamma \rangle$. The set of states and the set of actions are represented by S and U respectively. Each team contains n agents. The observation function $O(s, a)$, where s and $a \in \{1, \dots, n\}$ are state and n agents, determines the observation $z \in Z$ that each agent perceived individually at each time step. At each time step, the n agents choose their actions $u^a \in U$, which is an action taken by the a -th agent, based on their action-observation history. In this modeling, as recurrent neural networks (RNN) ([Hochreiter & Schmidhuber, 1997](#)) is used by the MARL algorithm, the policy is conditioned on the joint action-observation history as well as the current agent observation z . The state of the environment changes

according to a transition probability P . Unlike the partially observable stochastic game, all agents in Dec-POMDP share a collective reward and an individual reward drawn from the reward function $r(s, \mathbf{u})$, where \mathbf{u} is a joint-action which is a set of each agent's action. The discount factor of the MARL algorithm is represented by γ .

In MARL, as multiple agents act simultaneously in the environment based only on their own action-observation history and do not know about the individual policy of each agent, there exists a non-stationarity problem. The behaviors of other agents are changing during training and can influence the reward received by each agent. To address this issue, the system is modeled using a centralized training in decentralized execution (CTDE) framework. In the CTDE framework, the full state of the environment can be accessed in the training procedure to get the state-action value. On the other hand, only the local observation can be accessed by the agent during execution. The joint-action from all agents is also available during the training procedure by the centralized policy to alleviate the non-stationarity issue.

In this paper, we focus on value-based MARL algorithms applied in environments where a sense of cooperation is needed between agents, meaning that they share a collective reward. The proposed algorithm is to be combined with deep recurrent Q-networks (DRQN) (Hausknecht & Stone, 2015) and dueling deep Q-networks (Wang et al., 2016). The DRQN algorithm, as proposed in Hausknecht & Stone (2015), addresses single-agent with partially observable environments. The architecture consists of the DQN (Mnih et al., 2015) combined with RNN. The DRQN approximates the state-action value function $Q(s, u)$, where s and u are a state and an action of single agent, with RNN to maintain an internal state and aggregate observations over time. It also can be taken to approximate $Q(s_t, h_{t-1}, u)$, where s_t and h_{t-1} represent the observation at time step t and the hidden state at time step $t - 1$, which has information of previous states and acts as a memory. The proposed method is also to be combined with the dueling deep Q-networks (Wang et al., 2016). The dueling deep Q-networks is a neural network architecture designed for value-based RL that has two streams in the computation of the state-action value. One stream is for approximating the value function $V(s)$ and the other is for approximating the advantage function $A(s, u)$. The value function $V(s)$ depends only on state and presents how good a state is. The advantage function $A(s, u)$ depends on both state and action and presents how advantageous it is to take an action u in comparison to the other actions at the given state s . The value and the advantage are merged to get the final state-action value $Q(s, u)$ as follows

$$Q(s, u) = V(s) + A(s, u) - \frac{\sum_{u'} A(s, u')}{N}, \quad (1)$$

where u' represents each possible action and N is the number of actions. In this paper, the dueling deep Q-networks is combined with the RNN to handle the action-observation history used as the input of the policy. In the architecture of dueling deep Q-networks with the RNN, e.g., Dueling DRQN, the RNN is inserted right before the crossroad of streams of computation. The dueling DRQN is compared with the DRQN as an ablation study in 'Simulation Results'.

In the following subsections, other methods relevant to comparisons are presented. In this paper, we focus on methods that can be combined with off-policy value-based algorithms and focus on the maximization of a joint state-action value, trying to assign proper credit to individual agents on the shared reward received.

Counterfactual multi-agent policy gradients

Counterfactual multi-agent (COMA), introduced by [Foerster et al. \(2018\)](#), utilizes a single centralized critic to train decentralized actors and deals with the challenge of the multi-agent credit assignment problem. In the cooperative environments that are the main target for the COMA, it is difficult to determine the contribution of each agent to the shared collective reward received by the team. The centralized critic has access to the global state and the actions of the agent to model the joint state-action value function.

Value decomposition network

The value decomposition network (VDN) ([Sunehag et al., 2017](#)) aims at learning a joint-action value function $Q_{tot}(\tau, \mathbf{u})$, where τ is a joint-action observation history and \mathbf{u} is a joint-action. The $Q_{tot}(\tau, \mathbf{u})$ can be expressed as a sum of a -th agent's individual value functions $Q_a(\tau^a, u^a; \theta^a)$ as follow

$$Q_{tot}(\tau, \mathbf{u}) = \sum_{a=1}^n Q_a(\tau^a, u^a; \theta^a), \quad (2)$$

where each $Q_a(\tau^a, u^a; \theta^a)$ is a utility function of the a -th agent and θ^a is the policy of the a -th agent. The loss function for the VDN is the same as that of the deep Q-network (DQN) ([Mnih et al., 2015](#)), where Q is replaced by $Q_{tot}(\tau, \mathbf{u})$.

QMIX

QMIX ([Rashid et al., 2020](#)) is a deep multi-agent reinforcement learning method to be trained using CTDE. It uses the additional global state information that is the input of a mixing network. The QMIX is trained to minimize the loss, just like the VDN ([Sunehag et al., 2017](#)), given as

$$\mathcal{L}(\theta) = \sum_{i=1}^b [(y_i^{tot} - Q_{tot}(\tau, \mathbf{u}, s; \theta))^2], \quad (3)$$

where b is the batch size of transitions sampled from the replay buffer and Q_{tot} is output of the mixing network and the target $y_i^{tot} = r + \gamma \max_{\mathbf{u}'} Q_{tot}(\tau', \mathbf{u}', s'; \theta^-)$, and θ^- are the parameters of a target network. The QMIX allows learning of joint-action-value functions, which are equivalent to the composition of optimal Q-values of each agent. This is achieved by imposing a monotonicity constraint on the mixing network. Monotonicity can be enforced by the constraint on the relationship between Q_{tot} and individual Q value functions, given as

$$Q_a : \frac{\partial Q_{tot}}{\partial Q_a} \geq 0, \forall a \in A. \quad (4)$$

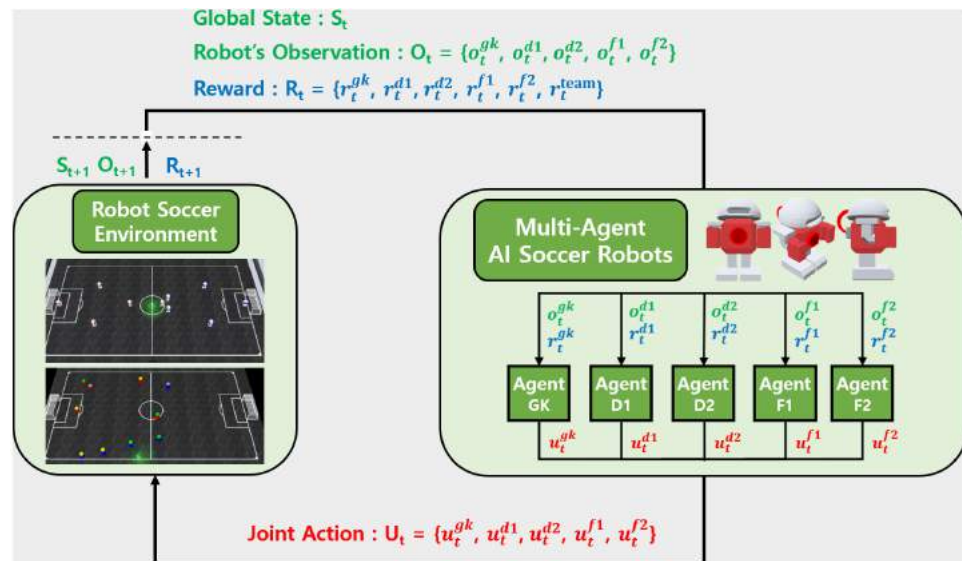


Figure 1 MARL structure for AI robot soccer.

Full-size [DOI: 10.7717/peerjcs.718/fig-1](https://doi.org/10.7717/peerjcs.718/fig-1)

PROPOSED METHOD

In heterogeneous multi-agent reinforcement learning, the main challenge can be described as how to provide an optimal policy and maximize cooperative behavior in a heterogeneous multi-agent environment. In this scenario, the agents act independently and maximize not only the individual reward but also a shared reward. To tackle this problem, a novel training method called two-stage heterogeneous centralized training is proposed and described in this section and to be applied to 5 *versus* 5 AI robot soccer.

MARL structure for AI robot soccer

The MARL structure in 5 *versus* 5 AI robot soccer is presented in Fig. 1. In the AI robot soccer each robot has its role. The roles are goalkeeper, defender 1, defender 2, forward 1, and forward 2 which are denoted as GK(gk), D1(d1), D2(d2), F1(f1), and F2(f2), respectively. Each robot has individual observations and individual rewards according to its role in soccer game. Each robot receives its individual observation o_t^{role} , $role \in \{gk, d1, d2, f1, f2\}$ at each time step t and selects its action u_t^{role} according to a policy network which is trying to maximizing individual role rewards r_t^{role} and team reward r_t^{team} . The policy network also takes into consideration past individual observations and actions taken. The concatenation of individual actions of the 5 robots forms a joint-action set U_t . By performing this joint-action in the AI robot soccer environment, the simulator calculates the next global state S_{t+1} , robot observation O_{t+1} , and reward R_{t+1} . It is noted that the global state is available only during training.

TSHCT architecture

As shown in Fig. 2, the training procedure is divided into two stages. In the first stage, agents of the same type (homogeneous agents, e.g., two agents as defenders) are trained.

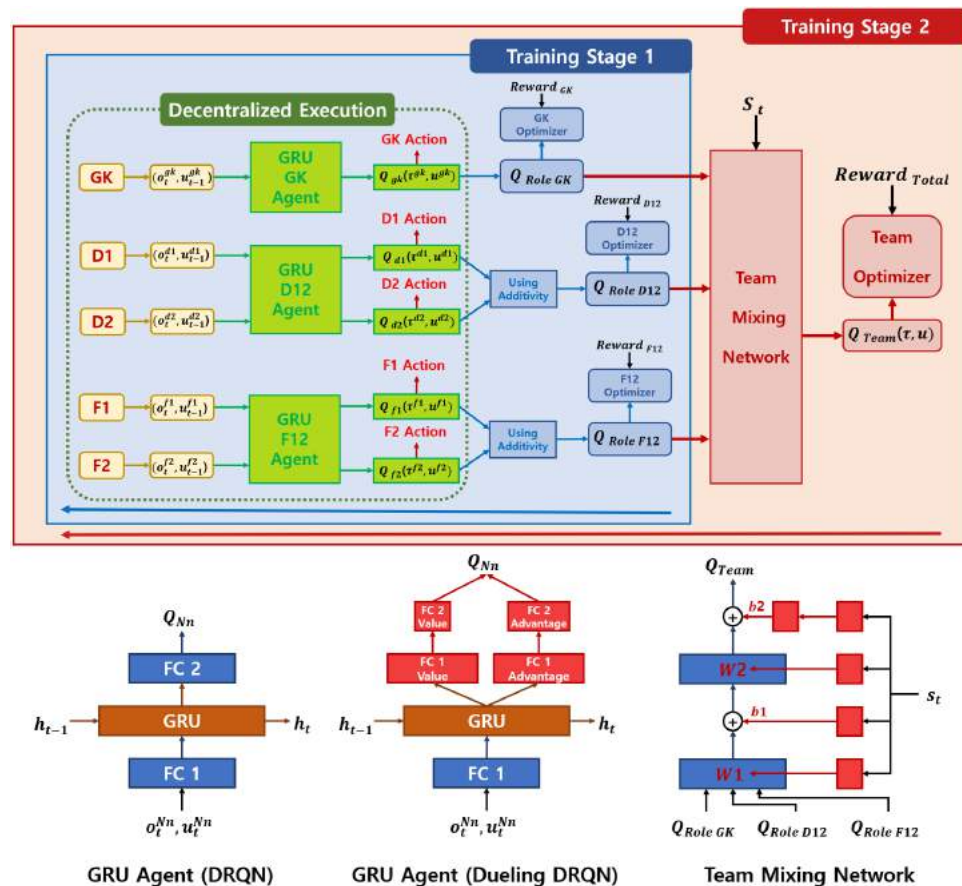


Figure 2 Overall architecture of two-stage heterogeneous centralized training. In the first stage, the agents are trained using their individual role rewards, goalkeeper reward, defender reward, and forward reward. A shared policy is used by defenders and by forwards. In the second stage, the agents are trained using a collective team reward. The global state of the environment s_t is also used as an input of the team mixing network, following the structure of a hypernetwork (Ha, Dai & Le, 2016).

Full-size [DOI: 10.7717/peerjcs.718/fig-2](https://doi.org/10.7717/peerjcs.718/fig-2)

Decentralized execution is used during inference and a shared policy is used by the agents of the same type. In the second training stage, all heterogeneous agents are trained jointly. These two stages are executed in a serial learning structure.

To model each agent's policy, the structure of DQN with gated recurrent unit (GRU) (Chung et al., 2014) or the structure of Dueling Q-Networks with GRU is used in the experiments. The policy network receives as input 40 subsequential frames with the current individual observation of the agent $o_t^{(N_n)}$ and the last action chosen $u_{t-1}^{(N_n)}$, where N_n is the n -th agent of the N -role (type). The output of the policy network is the state-action value Q_{N_n} . The action with the highest Q-value is chosen at each time step with epsilon greedy exploration.

In training stage 1, the $Q_{(RoleN)} \forall N \in \{GK(goalkeeper), D12(defenders), F12(forwards)\}$ is calculated by adding Q-values Q_{N_n} from the homogeneous agent network. In training stage 2, the team mixing network combines the individual role rewards into the shared

collective reward. The mixing network is modeled as a hypernetwork (Ha, Dai & Le, 2016), using feed-forward layers. The hypernetwork is conditioned on the global state S_t of the environment and takes the values of $Q_{(RoleGK)}$, $Q_{(RoleD12)}$, and $Q_{(RoleF12)}$ produced in training stage 1 as inputs. The output of the mixing network is Q_{Team} .

TSHCT learning equations

The proposed method is used to minimize the losses through the entire training. In training stage 1, each role optimizer updates the weights of the policy network to minimize the loss $\mathcal{L}_{RoleN}(\theta)$ in relation to the target y^{RoleN} . The target y^{RoleN} is calculated based on the Bellman equation (Bellman, 1954) with the sum of the individual role rewards $Reward_{RoleN}$ for the current time step and the Q-value estimated for the next state. The target and the loss are given as follows

$$y^{RoleN} = Reward_{RoleN} + \gamma \max_{u'} Q_{RoleN}(\tau', u', s'; \theta^-),$$

$$\mathcal{L}_{RoleN}(\theta) = \sum_{i=1}^b [(y_i^{RoleN} - Q_{RoleN}(\tau, u, s; \theta))^2], \quad (5)$$

where γ and θ^- are the parameters of a target network, the discount factor and policy, similar to the ones presented in DQN (Mnih et al., 2015) to stabilize the training procedure and b is the batch size of episodes sampled from the replay buffer. In training stage 2, the team optimizer updates the weights of mixing network and policy networks to minimize the team loss in relation to the team target y^{Team} calculated with the total shared reward $Reward_{Total}$, which is the sum of sparse cooperative team rewards and dense individual role rewards. The team loss $\mathcal{L}_{Team}(\theta)$ is given as follows


$$y^{Team} = Reward_{Total} + \gamma \max_{u'} Q_{Team}(\tau', u', s'; \theta^-),$$

$$\mathcal{L}_{Team}(\theta) = \sum_{i=1}^b [(y_i^{Team} - Q_{Team}(\tau, u, s; \theta))^2]. \quad (6)$$

Equations (5) and (6) are analogous to the minimum squared loss used in Mnih et al. (2015). Using additivity (Sunehag et al., 2017) and monotonicity (Rashid et al., 2020), the TSHCT trains heterogeneous agents by maximizing Q_{Team} in stage 2, while learning multiple roles by maximizing the Q-value of each individual role $Q_{(RoleGK)}$, $Q_{(RoleD12)}$, and $Q_{(RoleF12)}$ in stage 1.

TSHCT curriculum learning through self-play

To train a robust policy in a competitive-cooperative scenario that can work well against multi-agent in the opponent team, curriculum learning is needed. In this paper, we use self-play as a form of the implicit curriculum with the objective of learning robust AI robot soccer strategies. The implicit self-play curriculum is implemented by updating the opponent team when the number of episodes reaches a particular number. The opponent team is kept updated and reference policies take turns. Using self-play, it is possible to keep the opponent team at an appropriate level of competitiveness, not too strong so that the policy allows good behavior and not too easy so that the policy avoids weak behaviors. The soccer strategy learned through self-play tends to lead to acceptable game performance



Specifications	Roles		
	Goalkeeper (GK)	Defender (D1, D2)	Forward (F1, F2)
Robot Length/Width	15 cm		
Robot Height	28 cm		
Wheel Radius	4 cm		
Axle Length	14 cm		
Robot Body Mass	2.5 kg	2.0 kg	1.5 kg
Robot Body Center of Mass	1.5 cm above the ground		
Wheel Mass	0.15 kg per Wheel		
Wheel Center of Mass	Center of the Wheel		
Slider Mass	0.5 kg per Slider		
Max Linear Velocity Available	1.8 m/s	2.1 m/s	2.55 m/s
Max Torque Available	0.8 N*m	1.2 N*m	0.4 N*m

Figure 3 Specifications of the AI robot soccer environment. Robots with different roles, goalkeeper, defender, or forward, have different mass, maximum linear velocity, and maximum torque.

Full-size  DOI: [10.7717/peerjcs.718/fig-3](https://doi.org/10.7717/peerjcs.718/fig-3)

(Heinrich, Lanctot & Silver, 2015; Lanctot et al., 2017; Silver et al., 2017) as the result of the automated curriculum.

SIMULATION RESULTS

In this section, the MARL environment used in the experiments and the results obtained by the TSHCT and other baseline methods are described.

AI robot soccer MARL environment

To demonstrate the performance of the TSHCT, experiments are conducted in an AI robot soccer environment with specifications presented in Fig. 3, which is developed with Webots robot simulation software (Michel, 2004) and based on the environment described in Hong et al. (2021). In this AI Soccer simulation game, two teams compete similarly to a real soccer game, trying to kick the ball into the opponent's goal area to score and to win the game against the opponent team. In each team, there are 5 robots with 3 different roles (one goalkeeper, two defenders, and two forwards). The AI robot soccer game is divided into two 5 minute-long halves. For training, the game is divided up into episodes of 40 sequential frames. An episode is over whenever 40 sequential frames are processed.

Global state and observations

The global state, available only during centralized training and used as input to the mixing network, contains information of all the soccer robots and the ball. Specifically, the state vector contains the coordinates and orientations of all soccer robots, including robots of the opponent team, and the ball coordinates. The coordinates are relative to the center of the field (origin). The individual local observations of each robot are their relative positions in the field and relative distances and orientations to other robots and to the ball within their range of view. These observations are used as inputs of the policy networks.

Action

The basic actions committed by the robots are move, jump and kick. They are achieved by giving continuous control variables to the feet and legs. To achieve these behaviors a discrete set of 20 actions is designed which is allowed to be taken by the agent at each time step. A discrete set of actions is used so that the DRQN and the Dueling DRQN can be used as the off-policy value-based algorithms for the experiments. The discrete action set consists of actions of forward motion, backward motion, 6 directions of forward turns, 4 directions of backward turns, clockwise and counterclockwise turns, 2 kinds of forward turn combined with kick, 2 kinds of forward motion combined with kick, stop combined with kick, and stop.

Reward

To train AI soccer robots to perform their roles and cooperative behavior, individual role rewards and a shared team reward are defined. Individual role rewards are a combination of dense rewards associated with two pieces of information. One is the ball information relative to the robot, such as distance, velocity, and angle. The other is the information of the expected position which is defined for each role, *i.e.*, default position where the robot should be to play its role. The team reward is a combination of a sparse reward related to scoring and dense rewards related to the distance and velocity between the ball and the opponent's goal.

Equations (7) and (8) show the mathematical modeling of the individual role reward and the team reward. In Eq. (7), d_{rp} is the distance between the robot and its expected role position, θ_{rb} and v_{rb} are the relative angle and relative velocity between the robot and the ball, $d_{bg,pre/cur}$ is the distance between the ball and the opponent goal center at previous/current time step, and $isTouch$ is a boolean that is true when the robot touched the ball within the last 10 time steps. In Eq. (8), d_{bg} and v_{bg} are distance and velocity between the ball and the opponent goal center and $isScore$ is 100 if the team scored against the opponent team.

$$r^{role} = e^{-d_{rp}} + 0.5e^{-\theta_{rb}} + 0.5(1 - e^{-v_{rb}}) + 50(d_{bg,pre} - d_{bg,cur}) \times isTouch. \quad (7)$$

$$r^{team} = 5e^{-d_{bg}} + 5(1 - e^{-v_{bg}}) + isScore. \quad (8)$$

Network hyperparameters

The neural network hyperparameters used in the experiments are as follows

- DRQN architecture: 2 layers with 128 hidden units, 1 layer of GRU with 128 hidden units, and ReLU non-linearities.
- Dueling DRQN architecture: 1 layer with 128 hidden units, 1 layer of GRU with 128 hidden units, 2 layer with 128 hidden units for value prediction, 2 layer with 128 hidden units for advantage prediction, and ReLU non-linearities.
- Mixing network architecture: 1 layer of mixing network with 32 hidden units, 2 layers of hypernetworks with 32 hidden units, and ReLU non-linearities.

- ADAM optimizer (*Kingma & Ba, 2014*) with learning rate set to 4×10^{-5} for both policy and mixing networks.
- Discount factor γ set to 0.99.
- Target networks updated every 16,000 iterations.
- Epsilon used for exploration decreased by 0.025 every 10^4 iterations until it is kept at 0.05 at the end of training.
- Buffer size set to store 5×10^3 episodes.
- Batch size set to 64.

Results

Evaluation of TSHCT and baselines, COMA, VDN, and QMIX

In this section, the evaluation of the TSHCT and baseline methods, COMA, VDN, and QMIX, are presented. The proposed method and baseline methods are trained for a total of 200k episodes using epsilon greedy exploration with self-play. The evaluation is conducted by comparing the performances of 4 algorithms, TSHCT, COMA, VDN, and QMIX. The performances are measured by matches against three evaluation teams, noted as Evaluation Team 1, 2, and 3. As the result of the evaluation, comparisons of rewards and score-concede rates are given. The “score” term means a goal scored by own team while the “concede” term represents a goal scored by the opponent team. The score-concede rate is defined as the percentage of the number of scores divided by the sum of the number of scoring and conceding.

In the first evaluation, the performances of the TSHCT and the baselines are obtained by playing against the Evaluation Team 1, which is a team trained for 200k episodes with COMA. The experimental result shows that the TSHCT is superior to COMA, VDN, and QMIX algorithms after 80k episodes, as shown in [Fig. 4](#), where the total reward is defined as the sum of three individual rewards and the team reward. When the maximum average total reward is defined as the maximum value of the average of total reward of sequential 10,000 episodes, the maximum average total rewards of TSHCT, COMA, VDN, and QMIX are 5.92, 4.63, 4.83, and 5.05, respectively. The score-concede rate is defined as the maximum value of the averages of score-concede rates obtained over 10 sequential games. The score-concede rates of TSHCT, COMA, VDN, and QMIX are 79.01%, 50.40%, 64.21%, and 67.30%, respectively, as shown in [Fig. 5](#). It is observed that the TSHCT improves the score-concede rate by 28.61% as compared to that of COMA.

For the second evaluation, the performances of the TSHCT and the baselines are measured by playing against the Evaluation Team 2, which is a team trained for 200k episodes with VDN. Experiment results presented in [Fig. 6](#) show that the TSHCT is superior to the baseline algorithms after 80k episodes. The maximum average total rewards of TSHCT, COMA, VDN, and QMIX are 6.05, 4.50, 4.89, and 5.08, respectively. The maximum averages of score-concede rate of TSHCT, COMA, VDN, and QMIX are 62.85%, 32.27%, 50.97%, and 60.85%, respectively, as shown in [Fig. 7](#). It is observed that the TSHCT improved the score-concede rate by 11.88% as compared to that of VDN.

For the third evaluation, the performances of the TSHCT and the baselines are obtained by playing against the Evaluation Team 3, which is a team trained for 200k episodes with

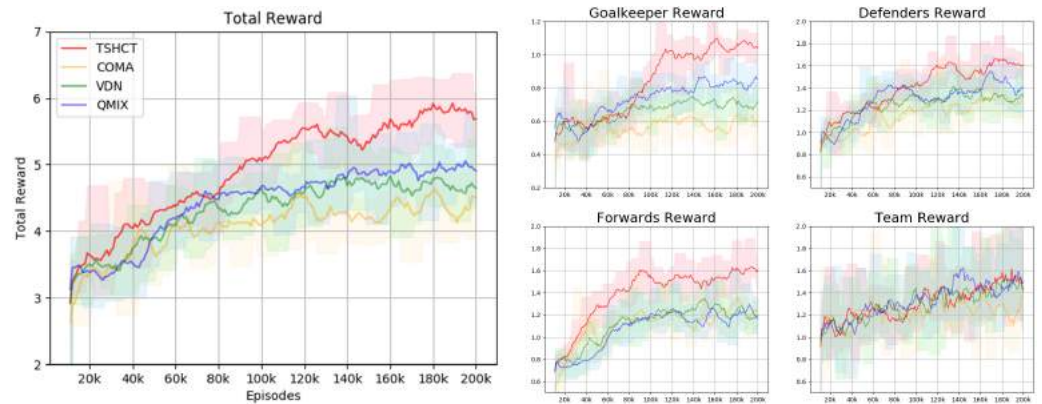


Figure 4 Total reward obtained during training by TSHCT, COMA, VDN, and QMIX. It is evaluated against Evaluation Team 1.

Full-size [DOI: 10.7717/peerjcs.718/fig-4](https://doi.org/10.7717/peerjcs.718/fig-4)

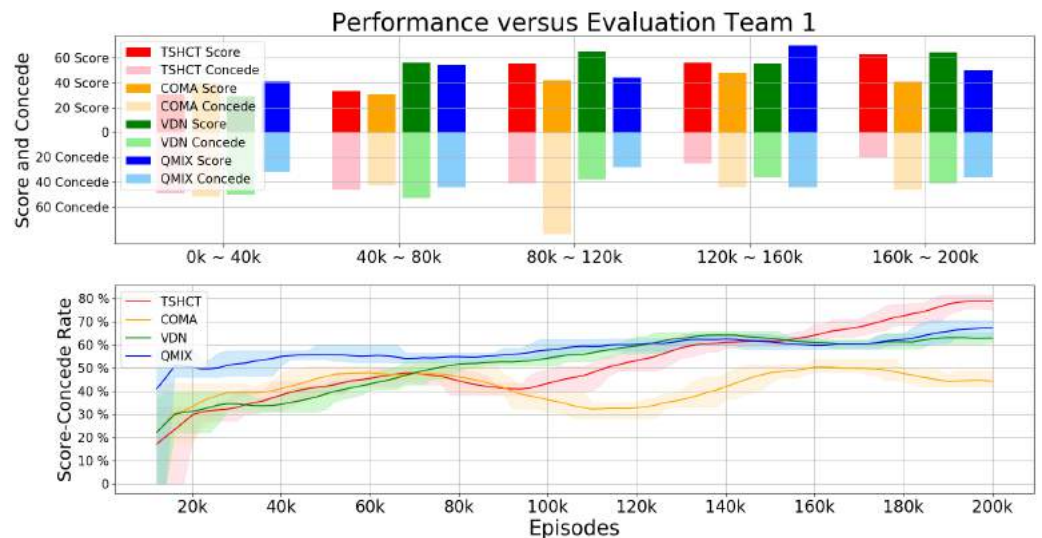


Figure 5 Comparison of score, concede, and score-concede rate obtained during training by TSHCT, COMA, VDN, and QMIX. It is evaluated against Evaluation Team 1.

Full-size [DOI: 10.7717/peerjcs.718/fig-5](https://doi.org/10.7717/peerjcs.718/fig-5)

QMIX. Experiment results show that TSHCT outperforms the baseline algorithms after 60k episodes, as shown in Fig. 8. The maximum average total rewards of TSHCT, COMA, VDN, and QMIX are 5.92, 4.50, 4.95, and 4.98, respectively. The maximum averages of score-concede rate of TSHCT, COMA, VDN, and QMIX are 52.08%, 29.99%, 48.84%, and 46.63%, respectively, as shown in Fig. 9. It is seen that the TSHCT improved the performance by 5.45% as compared to that of QMIX. It is important to mention that QMIX is the algorithm with the best performance when compared with the other baseline methods, VDN and COMA. s

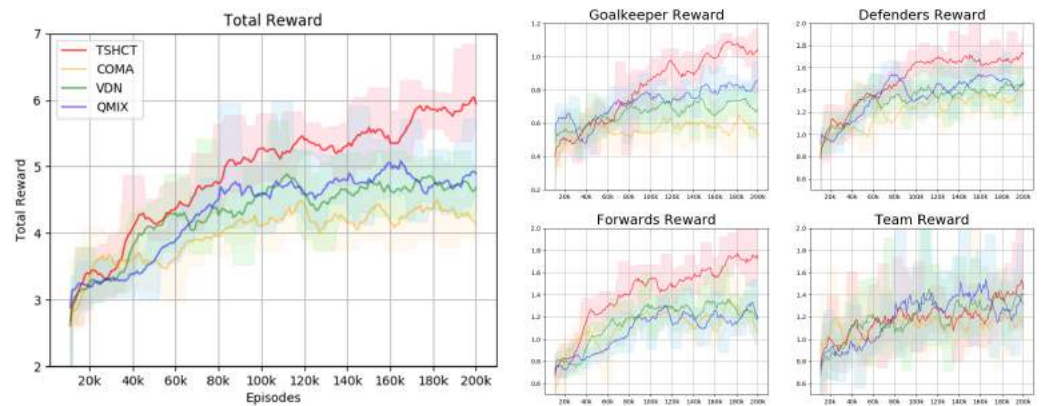


Figure 6 Total reward obtained during training by TSHCT, COMA, VDN, and QMIX. It is evaluated against Evaluation Team 2.

Full-size [DOI: 10.7717/peerjcs.718/fig-6](https://doi.org/10.7717/peerjcs.718/fig-6)

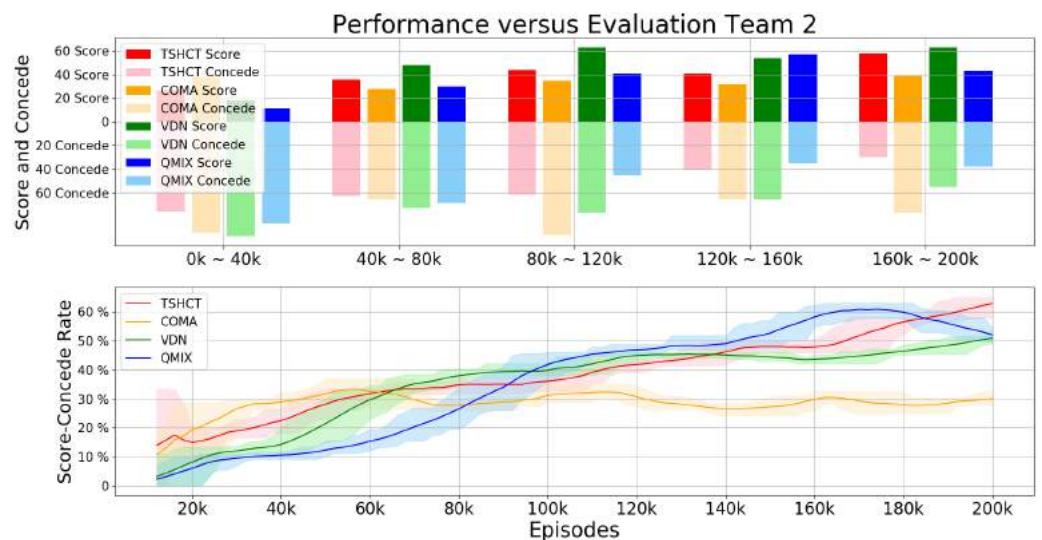


Figure 7 Comparisons of score, concede, and score-concede rate obtained during training by TSHCT, COMA, VDN, and QMIX. The score, concede, and score-concede rate are evaluated against Evaluation Team 2.

Full-size [DOI: 10.7717/peerjcs.718/fig-7](https://doi.org/10.7717/peerjcs.718/fig-7)

The final performances of the policies trained by the proposed method and the baseline methods are compared by conducting 10 min matches. Table 1 summarizes the results and statistics of these matches.

Ablation study: DRQN vs dueling DRQN

In AI robot soccer, several different sequences of actions can lead to similar reward values. From this observation, an ablation study is conducted by combining the TSHCT with dueling Q-network. Dueling Q-network often leads to better policy in the presence of distinct actions leading to similar reward values (Wang et al., 2016). In this ablation study, the traditional dueling Q-network architecture is combined with the RNN, which is named

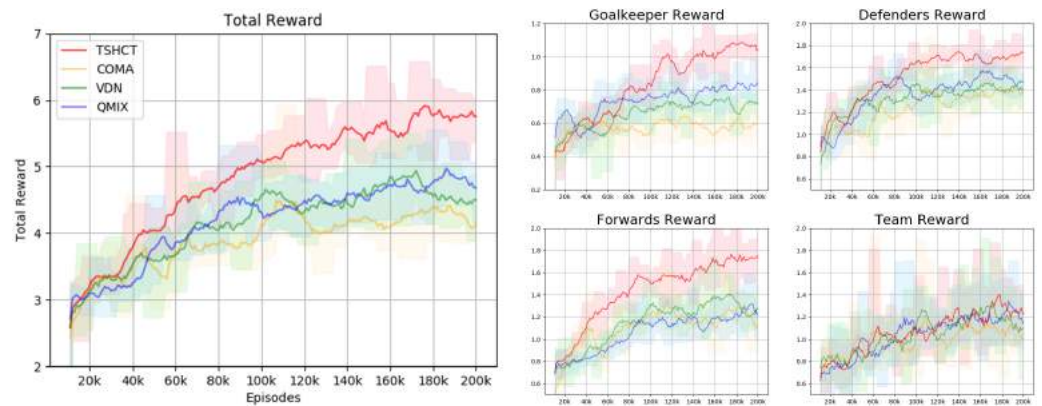


Figure 8 Total reward obtained during training by TSHCT, COMA, VDN, and QMIX evaluated against Evaluation Team 3.

Full-size [DOI: 10.7717/peerjcs.718/fig-8](https://doi.org/10.7717/peerjcs.718/fig-8)

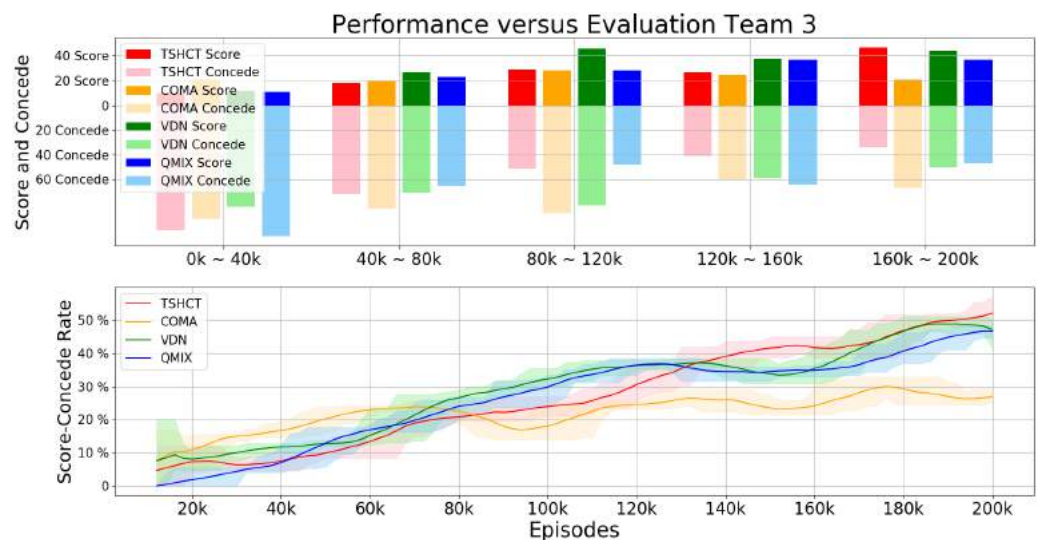


Figure 9 Comparison of score, concede, and score-concede rate obtained during training by TSHCT, COMA, VDN, and QMIX. The score, concede, and score-concede rate are evaluated against Evaluation Team 3.

Full-size [DOI: 10.7717/peerjcs.718/fig-9](https://doi.org/10.7717/peerjcs.718/fig-9)

here as Dueling DRQN. The proposed method combined with the Dueling DRQN is compared with the TSHCT combined with the DRQN. The TSHCT with Dueling DRQN is trained with 200k episodes using epsilon greedy exploration with self-play, similar to the cases shown in Figs. 5, 7 and 9. For comparisons of rewards and score-concede rates, game matches between the team trained by the TSHCT with DRQN, TSHCT-DRQN, and the team trained by the TSHCT with Dueling DRQN, TSHCT-Dueling DRQN, are conducted. The results of these matches are presented in Table 2.

In Fig. 10, the rewards obtained by the TSHCT with DRQN and the TSHCT with Dueling DRQN are presented. Figure 10 shows the increasing trends of rewards. It is seen

Table 1 Results and statistics of evaluation matches for TSHCT against the baseline methods.

	TSHCT	vs COMA	vs VDN	vs QMIX
100K episodes trained policy	Score	7.09 ± 1.83	3.92 ± 1.38	3.82 ± 1.70
	Concede	3.27 ± 1.54	4.23 ± 2.04	4.55 ± 0.89
	Score difference	3.82	-0.30	-0.73
	Score concede rate	68.4%	48.1%	45.6%
	Winning rate	100%	50%	20%
200k episodes trained Policy	Score	5.55 ± 2.23	5.00 ± 1.13	3.45 ± 1.44
	Concede	2.18 ± 1.59	2.82 ± 1.70	3.00 ± 1.41
	Score difference	3.37	2.18	0.45
	Score concede rate	71.8%	63.9%	53.5%
	Winning rate	100%	90%	80%

Table 2 Results and statistics of evaluation matches for TSHCT-Dueling DRQN against TSHCT-DRQN.

TSHCT-Dueling DRQN		vs TSHCT-DRQN	
		100k episodes trained policy	200k episodes trained policy
100k episodes trained policy	Score	5.09 ± 2.07	3.91 ± 1.78
	Concede	3.82 ± 2.03	4.64 ± 2.19
	Score difference	1.27	-0.73
	Score concede rate	57.1%	45.7%
	Winning rate	60%	30%
100k episodes trained policy	Score	8.36 ± 2.64	5.82 ± 2.48
	Concede	1.55 ± 1.30	2.27 ± 1.14
	Score difference	6.81	3.55
	Score concede Rate	84.4%	71.9%
	Winning rate	100%	80%

that the TSHCT with Dueling DRQN leads to a higher total reward as compared to the TSHCT with DRQN. The maximum average score-concede rates of the team trained by the TSHCT with Dueling DRQN against a team trained by the TSHCT with DRQN and three evaluation teams are 65.59%, 81.49%, 81.52%, and 64.67%, respectively, as shown in Fig. 11. The TSHCT with Dueling DRQN demonstrates improved score-concede rates over Evaluation Team 1, 2, and 3 by 2.48%, 18.67%, and 12.59% as compared to that obtained by the TSHCT with DRQN.

The policies trained by the TSHCT combined with DRQN and by the TSHCT combined with Dueling DRQN are compared with game results. Table 2 lists the results of these evaluation matches. For policies trained with the same number of training episodes, the TSHCT combined with Dueling DRQN outperforms the TSHCT combined with DRQN, achieving 60% and 80% winning rates with 100k episodes and 200k episodes, respectively. For the cases in which one algorithm is trained with two times the number of episodes of the opponent, *i.e.*, 200k *versus* 100k, the algorithm that was trained for a longer time achieves a higher winning rate. However, even for this case, the trained policy using the

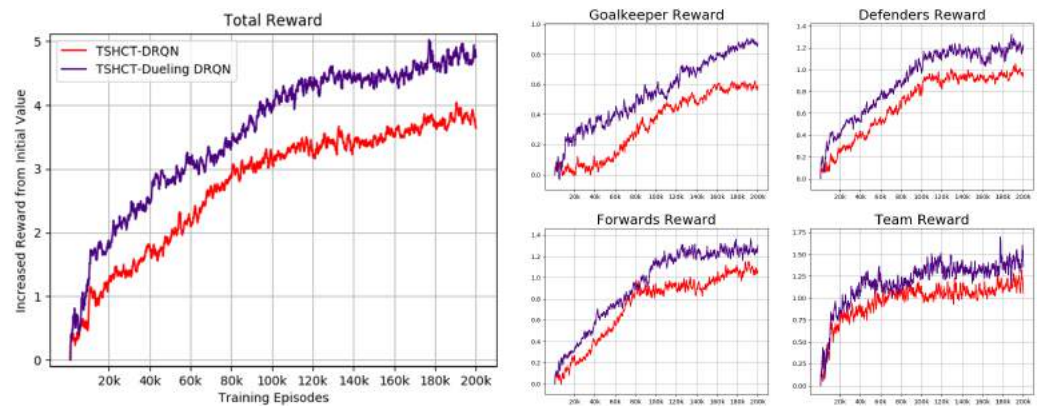


Figure 10 Rewards of TSHCT with DRQN and Dueling DRQN during training for 200k episodes with self-play.

Full-size [DOI: 10.7717/peerjcs.718/fig-10](https://doi.org/10.7717/peerjcs.718/fig-10)

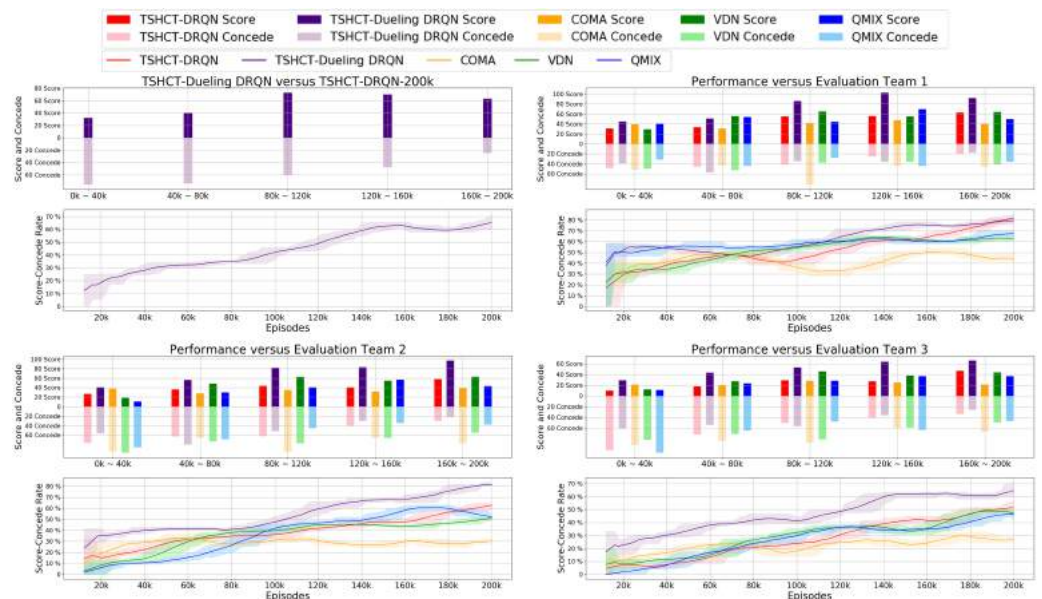


Figure 11 Comparison of score, concede, and score-concede rate obtained during training by TSHCT-DRQN, TSHCT-Dueling DRQN, COMA, VDN, and QMIX. The score, concede, and score-concede rate are evaluated against a team trained by the TSHCT with DRQN and three evaluation teams.

Full-size [DOI: 10.7717/peerjcs.718/fig-11](https://doi.org/10.7717/peerjcs.718/fig-11)

TSHCT combined with Dueling DRQN is more robust, achieving a 30% winning rate and a score-concede rate of 45.7%.

Discussion

Efficient exploration and reward modeling remain a big challenge in complex multi-agent environments. In a game such as robot soccer, using only team rewards, *e.g.*, a sparse score/concede reward or a sparse win/lose reward after the game is finished, is not enough for the agents to learn robust behavior. To deal with this problem, [Vinyals et al. \(2019\)](#)

use data from professional players at the beginning of the training in a supervised fashion to train a Starcraft 2 agent. Without this supervised data, it is difficult for the models to achieve a level capable of playing against good players and exploiting game strategies. In this aspect, by the results obtained in the results section, additional information in the form of individual role rewards that can be provided or learned unsupervised improves the policies.

In relation to the improvement of team rewards during training, the results obtained in the simulations indicate that it is difficult to train for cooperative behavior while performing multiple roles. The results obtained by the proposed method and the baseline methods, COMA, VDN, and QMIX, suggest that techniques that assign the contribution of each robot in the reward received as well as the techniques that train individualized roles that lead to stronger agents is needed. This can be addressed by the proposed method using two training stages. The stage 1 induces the learning of individual roles while stage 2 causes the learning of cooperative behavior and maximizing team rewards. From the observation of graphs of the individual role rewards, as shown by reward plots in Figs. 4, 6 and 8, the TSHCT achieves role rewards higher than those obtained by other algorithms. As robot soccer is a game played against an opponent team, the main objective, more than having high rewards during training, is to train multi-agent that performs well against opponents. Observing the matches against evaluation teams, as shown in Figs. 5, 7 and Fig. 9, it is noted that the proposed method is able to achieve substantially higher score-concede rate when compared with other methods. These results suggest that the proposed method in general works better than other methods. In the aspect of computational load, among the proposed method and the baselines, the proposed method takes the second-longest time to train the team for the same number of iterations because of its two stages.

In AI robot soccer, the policies are trained to maximize both individual role rewards and a shared team reward. Individual role rewards are designed for the robots to learn their roles, specifically to learn how to position and to learn how to control the ball to perform passing and shooting. Team rewards are designed for the team to learn how to score against the opponent team, avoid conceding, and also learn how to put pressure on opponent robots during the game (keeping the ball near the opponent goal area as much as possible during the game). The results obtained from simulation, using these rewards, have shown that the robots are able to learn individual role rewards while trying to act collaboratively. The GK learns to move to protect the goal against kicks of the opponent team while trying to kick away if the ball is reachable. The defenders act mostly if the ball is in the own field and try to recover the ball and kick the ball away from goal. When the ball is in the opponent field, defenders mostly try to position themselves in the field to avoid counter-attacks. Forwards are the most active players in the trained policies, trying to always be near the ball and kick the ball along right direction into the opponent's goal.

It is important to mention that, despite the results being obtained only in a simulated environment, the final goal of the RL approaches is to transfer the policy in a simulation to a real world scenario, such as playing a real robot competition in the RoboCup (Kitano *et al.*, 1995) contest. It is necessary to create a framework with the sensors available in the real robots in real-time so that the work learned by simulation can be transferred to real

robots without the need of re-training or with little re-training. Important research works have already been investigated while transferring the results obtained from simulation to real robots (*Peng et al., 2018*). To train robust models, the most important aspects are to respect the partially observable modeling of the robot soccer environment and to consider the variability of real world scenarios. For such purpose, noise that affects the state, the action, and the physics modeling should be added to the simulation environment so that less fine tuning is needed when deploying the trained policy.

CONCLUSION

This paper deals with multi-agent reinforcement learning with heterogeneous agents. The classic way to solve this problem is using the CTDE framework. However, the CTDE framework is less efficient for heterogeneous agents in learning individual behaviors. This paper presents the TSHCT, a novel heterogeneous multi-agent reinforcement learning method that allows heterogeneous agents to learn multiple roles for cooperative behavior. In the proposed method, there are two training stages that are conducted in a serial manner. The first stage is for training individual behavior through maximizing individual role rewards, while the second stage is for training cooperative behavior while maximizing a shared team reward. The experiments are conducted with 5 *versus* 5 AI robot soccer which is relevant to the cooperative-competitive multi-agent environment. The proposed method is compared with other baseline methods that maximize the shared reward to achieve cooperative behavior. The proposed method and baseline methods, COMA, VDN, and QMIX, are combined with value-based algorithms, such as DQN and dueling Q-networks.

Comparisons of total rewards and score-concede rates are presented in the paper. The results show that the TSHCT training method is superior to other baseline algorithms in role training and learning cooperative behavior. The maximum average score-concede rates of the TSHCT in comparison with the COMA, VDN, and QMIX are 79.01%, 62.85%, and 52.08%, respectively, representing the improvement achieved by the TSHCT in competitive AI robot soccer matches.

Because similar action-observation history leads to similar rewards in AI robot soccer, the training process can be unstable. To address this issue, an ablation study comparing the TSHCT combined with Dueling DRQN and DRQN is conducted. The performances of the TSHCT with DRQN and Dueling DRQN are measured by total rewards, score-concede rates, and match results. As a result, the TSHCT combined with Dueling DRQN achieves better performance when compared to the TSHCT combined with DRQN. The maximum average score-concede rate of the TSHCT with Dueling DRQN in comparison with the COMA, VDN, and QMIX are 81.49%, 81.52%, and 64.67%, respectively. This result represents an improvement of 2.48%, 18.67%, and 12.59% as compared to the case of the TSHCT combined with DRQN.

Simulation results show that the TSHCT is able to train an AI robot soccer team effectively, achieving higher individual role rewards and higher total rewards, as compared to other approaches that can be used for training to get cooperative behavior in a multi-agent environment. As future work, this framework is to be combined with actor-critic

policy-based multi-agent algorithms that can be applied in environments with continuous actions.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the Institute for Information communications Technology Promotion (IITP) grant funded by the Korean government (MSIT) (No.2020-0-00440, Development of Artificial Intelligence Technology that continuously improves itself as the situation changes in the real world). There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
 Institute for Information communications Technology Promotion (IITP).
 Korean government (MSIT): 2020-0-00440.
 Development of Artificial Intelligence Technology.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Taeyoung Kim conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Luiz Felipe Vecchietti conceived and designed the experiments, analyzed the data, performed the computation work, authored or reviewed drafts of the paper, and approved the final draft.
- Kyujin Choi conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Sanem Sarel and Dongsoo Har analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The code files are available at GitHub: https://github.com/ngng9957/TSHCT_dueling

REFERENCES

- Andrychowicz M, Wolski F, Ray A, Schneider J, Fong R, Welinder P, McGrew B, Tobin J, Abbeel P, Zaremba W. 2017. Hindsight experience replay. ArXiv preprint. [arXiv:1707.01495](https://arxiv.org/abs/1707.01495).

- Bellman R. 1954.** The theory of dynamic programming. Technical report, Randcorp santa monica ca.
- Berner C, Brockman G, Chan B, Cheung V, Dębiak P, Dennison C, Farhi D, Fischer Q, Hashme S, Hesse C. 2019.** Dota 2 with large scale deep reinforcement learning. ArXiv preprint. [arXiv:1912.06680](https://arxiv.org/abs/1912.06680).
- Chu T, Wang J, Codecà L, Li Z. 2019.** Multi-agent deep reinforcement learning for large-scale traffic signal control. *IEEE Transactions on Intelligent Transportation Systems* 21(3):1086–1095.
- Chung J, Gulcehre C, Cho K, Bengio Y. 2014.** Empirical evaluation of gated recurrent neural networks on sequence modeling. ArXiv preprint. [arXiv:1412.3555](https://arxiv.org/abs/1412.3555).
- Foerster J, Farquhar G, Afouras T, Nardelli N, Whiteson S. 2018.** Counterfactual multi-agent policy gradients. In: *Proceedings of the AAAI conference on artificial intelligence*.
- Ha D, Dai A, Le QV. 2016.** Hypernetworks. ArXiv preprint. [arXiv:1609.09106](https://arxiv.org/abs/1609.09106).
- Hausknecht M, Stone P. 2015.** Deep recurrent q-learning for partially observable mdps. ArXiv preprint. [arXiv:1507.06527](https://arxiv.org/abs/1507.06527).
- He X, Dai H, Ning P. 2015.** Improving learning and adaptation in security games by exploiting information asymmetry. In: *2015 IEEE conference on computer communications (INFOCOM)*. Piscataway: IEEE, 1787–1795.
- Heinrich J, Lanctot M, Silver D. 2015.** Fictitious self-play in extensive-form games. In: *International conference on machine learning, vol. 35. 7-9 July 2015*. Lille, France: PMLR, 805–813.
- Hochreiter S, Schmidhuber J. 1997.** Long short-term memory. *Neural Computation* 9(8):1735–1780 DOI 10.1162/neco.1997.9.8.1735.
- Hong C, Jeong I, Vecchietti LF, Har D, Kim JH. 2021.** AI world cup: robot soccer-based competitions. *IEEE Transactions on Games* Epub ahead of print 2021 11 March DOI 10.1109/TG.2021.3065410.
- Hwangbo J, Lee J, Dosovitskiy A, Bellicoso D, Tsounis V, Koltun V, Hutter M. 2019.** Learning agile and dynamic motor skills for legged robots. *Science Robotics* 4(26).
- Kim T, Vecchietti LF, Choi K, Lee S, Har D. 2020.** Machine learning for advanced wireless sensor networks: a review. *IEEE Sensors Journal*.
- Kingma DP, Ba J. 2014.** Adam: a method for stochastic optimization. ArXiv preprint. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- Kitano H, Asada M, Kuniyoshi Y, Noda I, Osawa E. 1995.** RoboCup: the robot world cup initiative. Available at <https://www.robocup.org/> (accessed on 15 June 2021).
- Klima R, Tuyls K, Oliehoek F. 2016.** Markov security games: learning in spatial security problems. In: *NIPS workshop on learning, inference and control of multi-agent systems*. 1–8.
- Lanctot M, Zambaldi V, Gruslys A, Lazaridou A, Tuyls K, Pérolat J, Silver D, Graepel T. 2017.** A unified game-theoretic approach to multiagent reinforcement learning. ArXiv preprint. [arXiv:1711.00832](https://arxiv.org/abs/1711.00832).
- Lillicrap TP, Hunt JJ, Pritzel A, Heess N, Erez T, Tassa Y, Silver D, Wierstra D. 2015.** Continuous control with deep reinforcement learning. ArXiv preprint. [arXiv:1509.02971](https://arxiv.org/abs/1509.02971).

- Liu S, Lever G, Merel J, Tunyasuvunakool S, Heess N, Graepel T. 2019. Emergent coordination through competition. ArXiv preprint. [arXiv:1902.07151](https://arxiv.org/abs/1902.07151).
- Lowe R, Wu Y, Tamar A, Harb J, Abbeel P, Mordatch I. 2017. Multi-agent actor-critic for mixed cooperative-competitive environments. ArXiv preprint. [arXiv:1706.02275](https://arxiv.org/abs/1706.02275).
- Michel O. 2004. Cyberbotics Ltd. Webots: professional mobile robot simulation. *International Journal of Advanced Robotic Systems* 1(1):5 DOI 10.5772/5618.
- Mnih V, Badia AP, Mirza M, Graves A, Lillicrap T, Harley T, Silver D, Kavukcuoglu K. 2016. Asynchronous methods for deep reinforcement learning. In: *International conference on machine learning*, vol. 48 20-22 June 2016. New York, New York, USA: PMLR, 1928–1937.
- Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, Riedmiller M. 2013. Playing atari with deep reinforcement learning. ArXiv preprint. [arXiv:1312.5602](https://arxiv.org/abs/1312.5602).
- Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G. 2015. Human-level control through deep reinforcement learning. *Nature* 518(7540):529–533 DOI 10.1038/nature14236.
- Nguyen TT, Nguyen ND, Nahavandi S. 2020. Deep reinforcement learning for multi-agent systems: a review of challenges, solutions, and applications. *IEEE Transactions on Cybernetics* 50(9):3826–3839 DOI 10.1109/TCYB.2020.2977374.
- Oliehoek FA, Amato C. 2016. A concise introduction to decentralized POMDPs. Cham: Springer.
- Peng XB, Andrychowicz M, Zaremba W, Abbeel P. 2018. Sim-to-real transfer of robotic control with dynamics randomization. In: *2018 IEEE international conference on robotics and automation (ICRA)*. Piscataway: IEEE, 3803–3810.
- Rashid T, Samvelyan M, De Witt CS, Farquhar G, Foerster J, Whiteson S. 2020. Monotonic value function factorisation for deep multi-agent reinforcement learning. *Journal of Machine Learning Research* 21(178):1–51.
- Sallab AE, Abdou M, Perot E, Yogamani S. 2017. Deep reinforcement learning framework for autonomous driving. *Electronic Imaging* 2017(19):70–76.
- Samvelyan M, Rashid T, De Witt CS, Farquhar G, Nardelli N, Rudner TG, Hung C.-M., Torr PH, Foerster J, Whiteson S. 2019. The starcraft multi-agent challenge. ArXiv preprint. [arXiv:1902.04043](https://arxiv.org/abs/1902.04043).
- Seo M, Vecchietti LF, Lee S, Har D. 2019. Rewards prediction-based credit assignment for reinforcement learning with sparse binary rewards. *IEEE Access* 7:118776–118791 DOI 10.1109/ACCESS.2019.2936863.
- Shalev-Shwartz S, Shammah S, Shashua A. 2016. Safe, multi-agent, reinforcement learning for autonomous driving. ArXiv preprint. [arXiv:1610.03295](https://arxiv.org/abs/1610.03295).
- Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M. 2016. Mastering the game of Go with deep neural networks and tree search. *Nature* 529(7587):484–489 DOI 10.1038/nature16961.
- Silver D, Hubert T, Schrittwieser J, Antonoglou I, Lai M, Guez A, Lanctot M, Sifre L, Kumaran D, Graepel T. 2018. A general reinforcement learning algorithm that

masters chess, shogi, and Go through self-play. *Science* **362**(6419):1140–1144
[DOI 10.1126/science.aar6404](https://doi.org/10.1126/science.aar6404).

Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A. 2017. Mastering the game of go without human knowledge. *Nature* **550**(7676):354–359 [DOI 10.1038/nature24270](https://doi.org/10.1038/nature24270).

Sunehag P, Lever G, Gruslys A, Czarnecki WM, Zambaldi V, Jaderberg M, Lanctot M, Sonnerat N, Leibo JZ, Tuyls K. 2017. Value-decomposition networks for cooperative multi-agent learning. ArXiv preprint. [arXiv:1706.05296](https://arxiv.org/abs/1706.05296).

Vecchietti LF, Kim T, Choi K, Hong J, Har D. 2020. Batch prioritization in multigoal reinforcement learning. *IEEE Access* **8**:137449–137461
[DOI 10.1109/ACCESS.2020.3012204](https://doi.org/10.1109/ACCESS.2020.3012204).

Vecchietti LF, Seo M, Har D. 2020. Sampling rate decay in hindsight experience replay for robot control. *IEEE Transactions on Cybernetics* Epub ahead of print 2020 21 May [DOI 10.1109/TCYB.2020.2990722](https://doi.org/10.1109/TCYB.2020.2990722).

Vinyals O, Babuschkin I, Czarnecki WM, Mathieu M, Dudzik A, Chung J, Choi DH, Powell R, Ewalds T, Georgiev P. 2019. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* **575**(7782):350–354
[DOI 10.1038/s41586-019-1724-z](https://doi.org/10.1038/s41586-019-1724-z).

Wang Z, Schaul T, Hessel M, Hasselt H, Lanctot M, Freitas N. 2016. Dueling network architectures for deep reinforcement learning. In: *International conference on machine learning*. PMLR, 1995–2003.

Zhang H, Feng S, Liu C, Ding Y., Zhu Y, Zhou Z, Zhang W, Yu Y, Jin H, Li Z. 2019. Cityflow: a multi-agent reinforcement learning environment for large scale city traffic scenario. In: *The world wide web conference*. 3620–3624.

