Regular paper | Journal of the KIECS. pp. 483-490, vol. 17, no. 3, Jun. 30. 2022, t. 113, pISSN 1975-8170 | eISSN 2288-2189
http://dx.doi.org/10.13067/JKIECS.2022.17.3.483

전이 학습을 이용한 VGG19 기반 말라리아셀 이미지 인식

Peng Xiangshen* · 김강철**

Malaria Cell Image Recognition Based On VGG19 Using Transfer Learning

Peng Xiangshen* · Kangchul Kim**

요 약

말라리아는 기생충에 의해 발생하는 질병으로 전 세계에 퍼져있다. 말라리아 셀을 인식하는데 일반적으로 두꺼운 혈흔과 얇은 혈흔 검사 방법이 사용되지만 이러한 방법은 많은 수작업 계산이 필요하여 효율성과 정확성이 매우 낮을 뿐만 아니라 빈민국에는 병리학자가 부족하여 말라리아 치명율이 높다. 본 논문에서는 특징 추출기, 잔류 구조와 완전 연결층으로 구성되고, 전이 학습을 이용한 말라리아셀 이미지를 인식하는 모델을 제안한다. VGG-19 모델의 사전 학습된 파라미터가 사용될 때 일부 컨볼루션층의 파라미터는 고정되고, 모델의 데이터에 맞추기 위하여 미세조정이 사용된다. 그리고 제안된 모델과 비교하기 위하여 잔류 구조가 없는 말라리아셀 인식 모델을 구현한다. 실험 결과 잔류 구조를 사용한 모델이 잔류 구조가 없는 모델에 비하여 성능이 우수 하였으며, 최신 논문과 비교하여 가장 높은 97.33%의 정확도를 보여주었다.

ABSTRACT

Malaria is a disease caused by a parasite and it is prevalent in all over the world. The usual method used to recognize malaria cells is a thick and thin blood smears examination methods, but this method requires a lot of manual calculation, so the efficiency and accuracy are very low as well as the lack of pathologists in impoverished country has led to high malaria mortality rates. In this paper, a malaria cell image recognition model using transfer learning is proposed, which consists in the feature extractor, the residual structure and the fully connected layers. When the pre-training parameters of the VGG-19 model are imported to the proposed model, the parameters of some convolutional layers model are frozen and the fine-tuning method is used to fit the data for the model. Also we implement another malaria cell recognition model without residual structure to compare with the proposed model. The simulation results shows that the model using the residual structure gets better performance than the other model without residual structure and the proposed model has the best accuracy of 97.33% compared to other recent papers.

키워드

Deep Learning, Transfer Learning, Convolutional Neural Network, VGG-19, Residual Structure, Malaria Cell
딥러닝, 전이 학습, 컨볼루션 신경망, VGG-19, 잔류 구조, 말라리아 셀

* 전남대학교 대학원 컴퓨터공학과 석사과정 (pxs287604536@gamil.com)
** 교신저자 : 전남대학교 공학대학 전기컴퓨터공학부
• 접 수 일 : 2022. 05. 17
• 수정완료일 : 2022. 06. 01
• 게재확정일 : 2022. 06. 17
• Received : May. 17, 2022, Revised : Jun. 01, 2022, Accepted : Jun. 17, 2022
• Corresponding Author : Kang-Chul Kim
School of Electricity and Computer Engineering, Chonnam National University,
Email : kkc@jnu.ac.kr

Ⅰ. Introduction

Malaria is a common tropical infectious disease which affects the health of people around the world. Malaria is listed as one of the most serious infectious diseases endangering human health until now. Even with the high priority of the World Health Organization and advocacy for control, malaria also remains high in mortality. The World Health Organization has released several vital facts about malaria that nearly half of the world's population has faced a risk of malaria, with more than 200 million malaria cases and about 400,000 deaths each year [1-2]. The common methods for the recognition of malaria cells are thick and thin blood smears examination[3], Polymerase Chain Reaction(PCR)[4], and Rapid Diagnostic Test(RDT)[5]. The last two tests are commonly used when the alternatives and high-quality microscope services are unavailable[6].

In the blood smears examination method, people need to examine the blood smear at 100 times magnification and manually calculate the number of red blood cells that contain parasites in 5000 cells. This method is an intensive manual process, but the accuracy depends on human expertise, and the different thoughts of inter-observers also adversely affect diagnostic results[7].

A doctor's diagnosis is affected when the doctor works in a limited resource environment(without the appropriate medical equipment)[8]. Therefore, a more effective technique is needed to recognize malaria cells.

Deep learning is an area of machine learning which has been performed and used very well in many medical fields in recent years[9-12]. The CNN is a classic model in deep learning, which has an excellent performance in image recognition. However another challenge is that training with CNN models requires large datasets, but it is difficult to obtain large datasets in practice. Transfer learning solves these problems.

The purpose of this paper is to help doctors diagnose malaria more accurately and efficiently, especially those working in resource-limited areas. We propose a neural network model that is used to recognize malaria cells. In the proposed model, the VGG-19 model is used as a pre-trained model. All fully connected layers of the VGG-19 model are removed and the residual structure is added after the last convolutional layer of the VGG-19 model. Then the new fully connected layers are added after the residual structure. When the pre-training parameters of the VGG-19 model are imported to the proposed model, the parameters of some convolutional layers of the VGG-19 model are frozen. In order to show the advantages of the residual structure, we propose another neural network model without residual structure to compare with the proposed model. In the model without residual structure, the VGG-19 model is also used as a pre-trained model, all fully connected layers of the VGG-19 model are removed, and the new fully connected layers are added.

In this paper, section Ⅱ describes the drawbacks of blood smear methods and other models' accuracy in recognition of malaria cells. Section Ⅲ introduces deep learning, transfer learning, pre-trained model, residual structure, dataset, and proposed model. Section Ⅳ shows the results of the models. Section Ⅴ concludes this paper.

Ⅱ. Related work

Malaria is a deadly disease that is hard to recognize. And it is difficult to distinguish between parasitic and non-parasitic infections in blood smear images[13]. Cells' appearance changes with time after being infected with malaria parasites[14]. Regular manual diagnostic blood smears are an

intensive manual process that requires experience to sort and count parasitized and uninfected cells. Usually, this method may not work well and may lead to problems if there is no experience with diagnostics in some locations. By using advanced image processing and analysis techniques, some papers have provided some methods for extracting manually designed features and building a classification model based on machine learning. D.K. Das built an automated diagnostic system for malaria detection using Support Vector Machine(SVM) and Naive Bayesian machine learning classifiers with 84% and 83.5% accuracy, respectively[15]. N.E Ross et al. [16] proposed three layers convolutional neural network as a classifier for automatic diagnosis of thin blood smears, and the accuracy is 85%.

In [17], the authors built two CNN models to recognize malaria cells. One model was a Sequential CNN model, which contained three convolutional layers and one fully connected layer. The other model was a ResNet built with 13 residual structures. The accuracy of their two models was 96.00% and 97.00%. Mohanad Mohammed Qanbar used the Residual Attention Network (RAN) to build a CNN model for malaria cell recognition[18], and the accuracy was 95.51%. Also, they used the SVM model to get 83.33% accuracy. A. Sai Bharadwaj Reddy and D. Sujitha Juliet used ResNet-50 as a pre-trained model to recognize malaria cells[19]. But they only modified the last layer of the ResNet-50, and the accuracy of their model was 95.40%.

III. Methodology

3.1 Deep learning

The deep learning model has proven very effective in various computer vision tasks[20]. The first article using deep learning with convolutional neural networks(CNN) to diagnose malaria was written by Liang et al. [21]. The key layers in the CNN model include convolution layers and pooling layers.

3.2 Transfer learning

Transfer learning uses existing knowledge to solve problems in different fields. The purpose of the transfer learning method is to move the original weights to the target model and re-train the part of the weights in the target model in order to solve problems with only a small number of datasets in the target field.

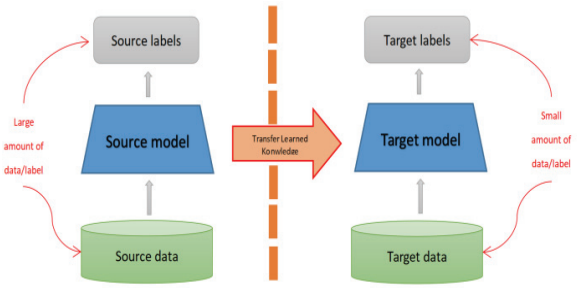

그림 1. 전이학습 블록다이어그램
Fig. 1 Transfer learning block diagram[22]

The structural framework of transfer learning is shown in Fig. 1. Fine-tuning is a transfer learning method. The idea of fine-tuning is to let the model learn some basic, generic features of the image from a large dataset and then learn more specific features from the target field. In this method, some layers of the neural network(these layers are called frozen layers, other layers called unfrozen layers) are frozen.

3.3 Pre-trained model

The pre-trained model is a neural network model that is pre-trained on a large dataset(such as the ImageNet dataset). The pre-trained model's weights and model structure in one domain are transferred to another domain through transfer learning.

In this paper, the VGG-19 model[23] is used as the pre-trained model. Fig. 2 shows the VGG-19 model architecture. In the VGG-19 model, the first few blocks(block 1 to block 4) extract more generic features, and the later blocks(block 5 to block 7) extract more specific features.

The VGG-19 model's dataset is the ImageNet dataset; this dataset is constructed for image recognition and classification. The ImageNet dataset has more than 14 million images containing more than 20,000 categories and more than a million images have explicit category annotations and annotations of object locations in the images.

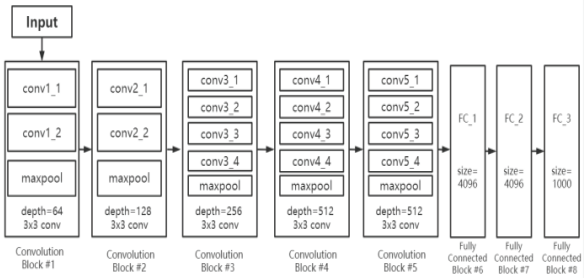

그림 2. VGG-19 모델 구조
Fig. 2 The VGG-19 model structure

3.4 Residual structure

The residual structure is a neural network model(shown in Fig. 3) which contains three convolutional layers and uses the addition function to fuse features(input's features map and the last convolutional layer output's features map). Theoretically, a deeper network has better performance. However, as the network gets deeper, a network model without residual structure is harder to train because the training error tends to increase, called network degradation[24]. But adding a residual structure helps solve the network degradation problem.

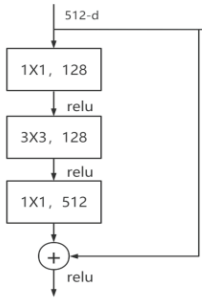

그림 3. 잔류 구조
Fig. 3 The residual structure

In order to reduce the parameter operations, the first convolutional layer does not change the size of the feature map, but only changes the depth of the feature map. The second convolutional layer only changes the size of the feature map, but does not change the depth. The third convolutional layer does not change the size of the feature map, and the depth is set to 512 in order to combine with the input feature map. The activation functions are relu functions. And the add function is used after the last convolutional kernel.

3.5 Dataset and Pre-processing

In this paper, dataset comes from the Kaggle datasets(malaria cell images dataset)[1]. This dataset contains 27,558 images and 8,000 images are randomly selected from the dataset, including 4,000 images of parasitized cells and 4,000 images of uninfected cells. Fig. 4 and Fig. 5 show images of parasitized cells and uninfected cells.

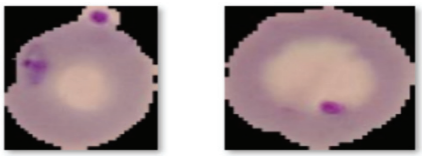

그림 4. 감염 셀
Fig. 4 Parasitized cells

1) NIH, Malaria Cell Images Dataset https://www.kaggle.com/iarunava/cell-images-for-detecting-malaria

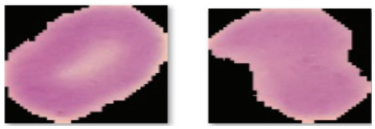

그림 5. 무감염 셀
Fig. 5 Uninfected cells

Pre-processing of the dataset plays an important role in training the model because it helps eliminate noise or redundancy and the training process becomes faster, thus improving accuracy. In this paper, firstly, 8,000 images are split into a training set and a testing set (70% for the training set and 30% for the testing set), then the images are resized to 224x224. Data Augmentation effectively prevents overfitting problems. Therefore, data augmentation techniques are used to pre-process the images. In this paper, rotation, horizontal flip, and zoom techniques are used.

3.6 Proposed model

Since the VGG-19 model is pre-trained on the ImageNet dataset, the parameters of the later blocks(from block 5 to block 7) in Fig. 2 describe the category information of the ImageNet dataset more specifically. So all the original fully connected layers of the VGG-19 model are removed and the parameters of block 1 to block 4 are frozen. The residual structure effectively reduces the network degradation problem, so the residual structure is added to our model. The GlobalAvgPool layer transforms all feature maps of the convolution layer output into a vector. Then the fully connected layer is added after the GlobalAvgPool layer, so the GlobalAvgPool layer is added to our model. Fig. 6 shows the proposed model architecture, named Model A.

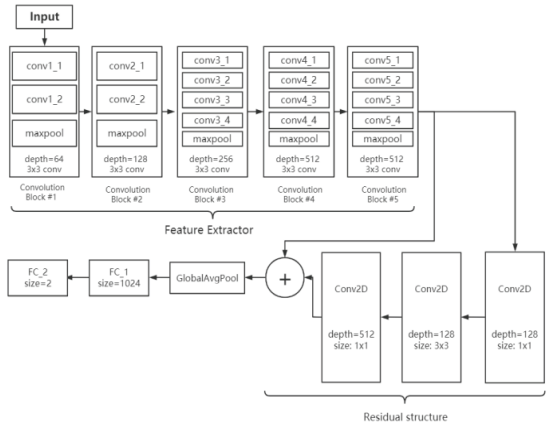

그림 6. 제안된 모델 구조(모델 A)
Fig. 6 The proposed model architecture(Model A)

Model A uses all the convolutional layers of VGG-19 as feature extractors. After the feature extractor, a residual structure(like Fig. 3) and the GlobalAvgPool layer are added. After the GlobalAvgPool layer, a fully connected layer with a size of 1024 is added. The last one is an output layer(classifier) which is a fully connected layer of size 2. In the last layer, the activation function is the softmax function.

The input data are trained with the pre-trained weights, and the parameters of these unfrozen layers are updated. The network's batch size is set to 64, the learning rate is set to 0.001, the loss function is categorical_crossentropy, and Adam is chosen as the optimizer.

Also, in order to show the residual structure to improve the performance of the deep neural network model more intuitively, we propose another model without adding the residual structure(Model B) in Fig. 7 to compare with Model A. The structure of Model B is the same as Model A, except it does not have the residual structure. The optimizer, activation function, loss function, learning rate, batch size, dataset, and data preprocessing are the same as Model A. And in Model B, the parameters of the VGG-19 model are frozen.

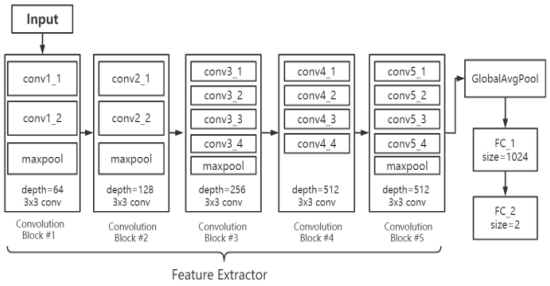

그림 7. 모델 B 구조
Fig. 7 Model B architecture

Ⅳ. Results

The accuracy, loss, precision, recall, and F1-score are used to evaluate the performance. Fig. 8 shows the progression of accuracy and loss in the training and validation sets from the beginning to the last epoch in Model A. The best accuracy and loss of the training and testing set are 97.52%, 0.0701, and 97.33%, 0.0950 respectively. The accuracy and loss difference between the testing and training sets is insignificant, so Model A is not overfitting.

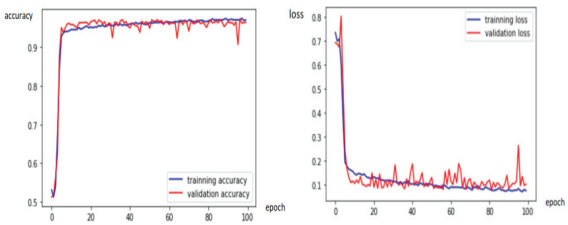

그림 8. 모델 A에 대한 정확도와 손실 과정
Fig. 8 The progression of accuracy and loss in Model A

Fig. 9 shows the accuracy and loss values for each epoch in Model B.

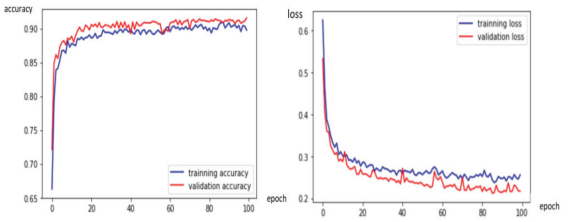

그림 9. 모델 B에 대한 정확도와 손실 과정
Fig. 9 The progression of accuracy and loss in Model B

Model A and Model B's precision, recall, accuracy and F1-score of parasitized cells and uninfected cells are shown in Table 1.

표 1. 모델 A와 B의 성능
Table. 1 Performance of Model A and Model B

Model	Cell	Precision	Recall	F1-score	Accuracy
Model A	Uninfected	0.97	0.96	0.96	0.97
Model A	Parasitized	0.96	0.97	0.97	0.97
Model B	Uninfected	0.90	0.93	0.91	0.91
Model B	Parasitized	0.93	0.90	0.92	0.91

Table 2 shows the status of the parameters of Model A and Model B. The GPU(Tesla P100) is used to train Model A and Model B. Model A takes 1 hour and 45 minutes, Model B takes 1 hour and 43minutes. So even if the trainable parameters are increased after adding the residual structure, the training time is almost the same, but the accuracy of Model A is much higher than Model B. Table 3 shows the comparison of the results of Model A and Model B.

표 2. 모델 A와 B의 매개변수
Table 2. Model A and B parameters

Model	Total parameters	Trainable parameters	Non-trainable parameters
Model A	20,833,346	10,246,658	10,586,688
Model B	20,551,746	527,362	20,024,384

표 3. 모델 A와 B의 결과 비교
Table. 3 Result comparison of the model A and B

Model	Accuracy	Loss	Training time
Model A(training set)	97.52%	0.0701	1 hour and 45 minutes
Model A(testing set)	97.33%	0.0950	
Model B(training set)	90.91%	0.2370	1 hour and 43 minutes
Model B(testing set)	91.58%	0.2168	

The above tables and figures show that the model using the fine-tuning method and adding the residual structure (Model A) gets better performance than the model B.

In [17] [18] [19], they all use the whole

image(27,558 images), but the results are similar to our results or even lower. Also, in [17] and [18] they do not use the transfer learning method. In [19], they have used the transfer learning method but the accuracy is also lower than Model A. So, 8,000 images are enough for transfer learning in this experiment. Table 4 shows the performance comparison from other models.

표 4. 다른 모델과 정확도 비교

Table 4. Accuracy comparisons with other models

Model	Accuracy
ResNet CNN model [17]	97.00%
Sequential CNN model [17]	96.00%
RAN model [18]	95.51%
SVM model [18]	83.33%
Modified ResNet-50 [19]	95.40%
Model A	97.33%
Model B	91.58%

Ⅴ. Conclusion

The results obtained in this experiment support the importance of transfer learning in malaria cell recognition. This paper aims to demonstrate that the pre-trained model offers high accuracy for malaria cell recognition utilizing transfer learning. In this paper, the VGG-19 model is used as a pre-trained model to build a network and the residual structure is added. The fine-tuning method is used to train our model, and then the accuracy of the proposed model is 97.33%.

References

[1] WHO, “World malaria report 2021,” Report, 2021.

[2] WHO, “Malaria: fact sheet,” Report, 2014.

[3] J. Duncan and L. Hartley, ''Blood Smear Examination: Normal Cells,'' *Veterinary Nursing Journal*, vol. 15, no. 6, 2014, pp. 231-234.

[4] N. J. Shah, *Introduction to Basics of Pharmacology and Toxicology*. Singapore: Springer 2019.

[5] T. Visser, J. Daily, N. Hotte, C. Dolkart, J. Cunningham, and P. Yadav, ''Rapid diagnostic tests for malaria,'' *Bulletin of the World Health Organization*, vol. 93, no. 12, 2015, pp. 862-866.

[6] WHO, ''Malaria microscopy quality assurance manual-Ver.2,'' *Report*, 2016.

[7] W. Shang, K. Sohn, D. Almeida, and H. Lee, "Understanding and improving convolutional neural networks via concatenated rectified linear units," *33rd International Conference on Machine Learning*, New York, USA, 2016.

[8] P. Moeskops, J. Wolterink, B. van der Velden, K. Gilhuijs, T. Leiner, M. Viergever, and I. Išgum, ''Deep Learning for Multi-Task Medical Image Segmentation in Multiple Modalities,'' *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Athens, Greece, 2016.

[9] K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. Van Gool, ''Deep Retinal Image Understanding,'' *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Athens, Greece, 2016.

[10] Y. Sato and G. Gerig, ''MICCAI: Medical Image Computing and Computer-Assisted Intervention,'' *Academic Radiology*, vol. 10, no. 12, 2003, pp. 1339-1340.

[11] J. H. Bong, S. H. Jeong, S. Jeong, and J. Han, ''Study on Image Use for Plant Disease Classification,'' *Journal of the Korea Institute of Electronic Communication Sciences*, vol. 17, no. 2, 2022, pp. 343-350.

[12] M. Kim, ''A Study on the Sports Rehabilitation Treatment for the Intellectual Disabilities using deep learning,'' *Journal of the Korea Institute of Electronic Communication Sciences*, vol. 15, no. 4, 2020, pp. 725-732.

[13] L. Rosado, J. M. Correia da Costa, D. Elias, and J. S. Cardoso, "A Review of Automatic Malaria Parasites Detection and Segmentation in Microscopic Images," *Anti-Infective Agents*, vol. 14, no. 1, 2016, pp. 11-22.

[14] F. B. Tek, A. G. Dempster, and I. Kale, "Parasite detection and identification for automated thin blood film malaria diagnosis," *Computer vision and image understanding*, vol. 114, no. 1, 2010, pp. 21-32.

[15] D. K. Das, R. Mukherjee, and C. Chakraborty, "Computational microscopic imaging for malaria parasite detection: a systematic review," *Journal of microscopy*, vol. 260, no. 1, 2015, pp. 1-19.

[16] N. E. Ross, C. J. Pritchard, D. M. Rubin, and A. G. Dusé, "Automated image processing method for the diagnosis and classification of malaria on thin blood smears," *Medical and Biological Engineering and Computing*, vol. 44, no. 5, 2006, pp. 427-436.

[17] S. Sinha, U. Srivastava, V. Dhiman, P. S. Akhilan, and S. Mishra, ''Performance assessment of deep learning procedures: Sequential and ResNet on malaria dataset,'' *Journal of Robotics and Control (JRC)*, vol. 2, no. 1, 2021, pp. 12-18.

[18] S. Tasdemir and M. M. Qanbar, ''Detection of Malaria Diseases with Residual Attention Network,'' *International Journal of Intelligent Systems and Applications in Engineering,* vol. 7, no. 4, 2019, pp. 238-244.

[19] A. S. B. Reddy and D. S. Juliet, ''Transfer Learning with ResNet-50 for Malaria Cell-Image Classification,'' *International Conference on Communication and Signal Processing*, Chennai, India, 2019.

[20] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, and T. Chen, ''Recent advances in convolutional neural networks,'' *Pattern Recognition*, vol. 77, 2018, pp. 354-377.

[21] Z. Liang, A. Powell, I. Ersoy, M. Poostchi, K. Silamut, K. Palaniappan, P. Guo, M. Hossain, A. Sameer, R. Maude, J. Huang, S. Jaeger, and G. Thoma, ''CNN based analysis for malaria diagnosis,'' *International conference on bioinformatics and biomedicine(BIBM),* Shenzhen China, 2016.

[22] R. Harini and N. Sheela, "Feature Extraction and Classification of Retinal Images for Automated Detection of Diabetic Retinopathy," *Second International Conference on Cognitive Computing and Information Processing*, Mysuru India, 2016.

[23] K. Simonyan and A. Zisserman, ''Very Deep Convolutional Networks for Large-Scale Image Recognition Karen,'' *American Journal of Health-System Pharmacy*, vol. 75, no. 6, 2018, pp. 398-406.

[24] O. Wichrowska, N. Maheswaranathan, M. Hoffman, S. Colmenarejo, M. Denii, N. De Freitas, and S.-D. Jascha, ''Learned optimizers that scale and generalize,'' *34th International Conference on Machine Learning,* Sydney, Australia, 2017.

저자 소개

Peng Xiangshen

2019년 WenHua College, Electronic Science and Technology 졸업(공학사)
2020년 ~현재 전남대학교 대학원 컴퓨터공학과 재학
※ 관심분야 : Image Recognition, Computer vision

김강철(Kang-Chul Kim)

1981년 서강대학교 전자공학과 학사
1983년 서강대학교 전자공학과 석사
1996년 경상대학교 전자공학과 박사
현재 전남대학교 전기전자통신컴퓨터공학부 교수
※ 관심분야 : 임베디드시스템, NoC, IoT
Pattern Recognition