

CNN-Based Malaria Parasite Detection in Thin Blood Smear Images

Pedram Mohammadalizadehbakhtevari
Department of Electrical and Computer Engineering
Queen's University
Kingston, Ontario, Canada
pedram.mohammadalizadehbakhtevari@queensu.ca

Abstract— Malaria is a serious global health problem of humans caused by the Plasmodium parasites transmitted through the bite of female Anopheles mosquito. Visual examination of blood smears under the microscope by expert technicians is the standard way of diagnosing malaria parasite infected red blood cells. This method is inefficient since the diagnostic accuracy heavily depends on the experience and the expertise of the person doing the examination. Although automatic image recognition techniques based on machine learning have been applied toward the task of malaria diagnosis of blood smears before, the delivered performance has not been sufficient so far. In this study we propose two deep learning models based on convolutional neural networks (CNNs) toward the task of automated classification of malaria infected and uninfected thin blood smear images. Using a total number of 27598 images from human blood cells our 21- and 16- layer CNNs could achieve average accuracy of 96.08% and 95.53% respectively. Comparing our CNN-based solutions with transfer learning models, we can observe superior performance of CNN models over transfer learning method along with other traditional classifiers.

Keywords—Deep Learning, Convolutional Neural Networks, Machine Learning, Malaria, Computer-aided diagnosis

I. INTRODUCTION

Malaria is a major infectious disease of humans caused by the Plasmodium parasites transmitted through the bite of the female Anopheles mosquito. According to the World Health Organization (WHO), 212 million people were affected by malaria in 2016 with more than 429,000 deaths [1]. With the advent of more accurate and efficient diagnostic methods, malaria can be prevented, controlled, and cured more effectively. The standard diagnostic method for malaria is the microscopic thick and thin blood smear examination. The presence of malaria parasites is detected by examining thick blood smears while the species of the parasite causing the infection are identified through thin blood smears [2]. However, this method is inefficient, since the diagnostic accuracy heavily depends on the

experience and expertise of the microscopists. Polymerase chain reaction (PCR) and Rapid diagnostic tests (RDT) are also widely used but they are more expensive, limited in terms of performance, and provide less information than microscopy [3, 4].

Machine Learning (ML) techniques with hand-engineered features are used in the majority of image analysis-based computer-aided diagnosis (CADx) software, for decision making purposes [5, 6, 7]. However, analyzing the variance in shape, size, background, angle, and position of the region of interest (ROI) on the images is a process in which needs expertise. Deep Learning (DL) method is used with significant success to overcome the challenges of creating hand-engineered features that capture the characteristics of the underlying data [8]. Deep learning models can automatically discover features in the data by using successions of layers of non-linear processing units. In these models, higher-level features are extracted from lower-level features which makes the learning of complex, non-linear decision-making functions, with end-to-end feature extraction and classification, possible [9]. The performance of deep learning models unlike kernel-based algorithms such as Support Vector Machines (SVMs), improves with the increase in data size and with the availability of powerful computational resources deep learning models become highly scalable [10]. Automatic image recognition technologies based on deep learning and big data have been recently applied to both thick and thin malaria blood smears for microscopic diagnosis.

In this work, we use a deep learning method to detect parasite-infected red blood cells in thin smears on standard microscope slides prepared using routine methods. We propose a 21-layer convolutional neural network (CNN) model, which is a deep learning model particularly designed for learning of two-dimensional data such as images toward classifying the parasitized and uninfected cells images to aid in improved disease screening. The advantage of a CNN model is that its hierarchical structure provides a general-purpose learning framework with end-to-end feature extraction and promises highly scalable and superior results over traditional classifiers.

II. RELATED WORK

A. Convolutional Neural Network

Convolutional neural network (CNN) is a class of deep learning models designed for image recognition and classification purposes. For images, an important source of information can be extracted from the spatial local correlation between the adjacent pixels. CNNs are designed to exploit this information by their key features such as local receptive fields, shared weights, pooling, and the use of many layers [8, 11].

The early applications of CNN can be traced back to the 1990s for speech recognition [12] and text recognition when LeCun et al. (1998) introduced a CNN model called LeNet-5 in the task of recognizing handwritten digits. This neural network was trained on grayscale images of MNIST dataset [13]. The use of CNNs is then extended to natural image recognition [14].

Convolutional neural networks have achieved significant performance in large-scale image recognition recently [11, 13, 18, 19, 20]. This achievement has become possible due to high-performance computing systems (HPC), such as GPUs or large-scale distributed clusters [15], and the large public image repositories, such as ImageNet. The ImageNet project is a large visual database designed for use in visual object recognition software research. More than 14 million images have been hand-annotated by this project to indicate what objects are pictured and in at least one million of the images, bounding boxes are also provided. ImageNet contains more than 20,000 categories [16]. In particular, ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) has played a pivotal role in the improvement of deep visual recognition architectures [17]. This challenge has served as a testbed for a few generations of large-scale image classification systems.

In 2012, AlexNet, a CNN based deep learning model proposed by Alex Krizhevsky et al., won the ILSVRC-2012 competition. This model enhanced the performance of CNN models significantly, toward the task of natural image classification. AlexNet was composed of five convolutional layers followed by two fully connected layers and used relatively large receptive fields and stride in the first convolution layers (11 x 11 with stride 4). The AlexNet Network has near 60M trainable parameters and was trained on two GPUs. This Network is considered a breakthrough application of CNNs to multi-categorical classification which successfully classified the ILSVRC-2012 sets with 10,184 categories and 8.9 million images with a top-5-error of 15.3% [11]. Following this initial success, several representative CNNs demonstrated significant improvements in succeeding ILSVRC annual challenges. In 2013, M. D. Zeiler et al., proposed ZFNet which used smaller receptive fields and stride in the first convolution layers (7 x 7 with stride 2) compared to

AlexNet architecture and achieved the top-5 error of 14.8% [18], then in 2014 GoogLeNet was proposed by Szegedy et al., which introduced Inception module (layer). In an inception layer, multiple convolutional layers with different number of filters and different filter sizes (1 x 1, 3 x 3, and 5 x 5) along with a pooling layer are used together to reduce the number of hyper-parameters and make the network's increase in depth feasible. Using this module, GoogLeNet was able to achieve the top-5 error of 6.7% [19].

In 2015, Simonyan et al., proposed VGGNet, which has 19 weight layers and uses very small 3 x 3 receptive fields throughout the whole network. This way, they could achieve the top-5 error of 6.8% by a relatively simple yet deep architecture [20].

Then in 2016, ResNet was proposed by He et al., which introduced Residual blocks. Residual blocks implement skip connections, which can take the activation from one layer and directly feed it to another layer deeper in the neural network. Using Residual blocks, ResNet could implement a very deep neural network without facing the vanishing and exploding gradient types of problems, and therefore, could reduce the top-5 error to 3.6% [21].

Later, Xception was proposed by Chollet et al., which outperformed the Inception-V3 model, a derivative of GoogLeNet architecture (Szegedy et al., 2016), on the ImageNet data classification task by using depth-wise separable convolutions [22]. Huang et al. have proposed the latest CNN variant called Densely Connected Convolutional Networks (DenseNet), which in their network architecture, each layer is directly connected to every later layer. This model has achieved notable improvements compared to the state-of-the-art models while significantly decreasing the number of parameters along with the computational cost [23].

B. Automatic Malaria Blood Smear Classification

Machine learning has been used by Sio et al. (2007) for rapid and accurate detection of parasitemia in images of Giemsa-stained blood smears [24]. This early study introduced an image analysis software for the task of parasitemia detection but did not provide any parasitized and uninfected red blood cells classification. In 2009, Diaz et al., applied a support vector machine (SVM) to preprocessed red blood cells images to classify infected and uninfected cells. Although this proposed classifier performed well, in terms of specificity and sensitivity, on a dataset consisting of 450 malaria images, its performance drops when it is applied to blood images in the infection stage [25].

Recently, researchers have begun to apply deep learning methods into a wide variety of medical image analyses and have achieved decent results [26, 27]. Malaria parasite detection from the blood cells images is one of the areas in medical image analysis that has benefited from the deep learning methods. In 2017, Dong et al., compared the performance of pre-trained CNN-based

deep learning models such as LeNet [13], AlexNet [11], and GoogLeNet [19] with an SVM classifier in terms of accuracy toward the task of classifying infected and uninfected cells. The authors reported more than 95% classification accuracy for CNN-based models, outperforming the SVM classifier delivering 92% accuracy [28]. Liang et al. (2017) proposed a 16-layer CNN model toward the task of classifying the uninfected and parasitized cells and compared the performance of their proposed model with a pre-trained AlexNet model. This pre-trained model was trained on the CIFAR-100 dataset and was used as the feature extractor for transfer learning. The extracted features were then classified using an SVM classifier. The authors reported that their proposed model outperformed the pretrained model in terms of accuracy, sensitivity, and specificity [29]. Bibin, Nair & Punitha (2017) proposed a method to identify the presence of malaria parasites in human blood cells using a 6-layer deep belief network. The study reported 96.4% accuracy toward the task of classifying 4,100 peripheral blood smear images into infected and non-parasitized classes [30]. Rajaraman et al. (2018) evaluated the performance of pre-trained CNN based deep learning models as feature extractors toward classifying the parasitized and uninfected cells to aid in improved disease screening. The authors presented a comparative analysis of the performance of their customized model and pre-trained deep learning models such as AlexNet, VGG-16, ResNet-50, Xception, and DenseNet-121 as feature extractors toward Malaria detection. The authors reported ResNet-50 outperforming the other models with the accuracy of 95.7% and then analyzed the optimal layer in the pre-trained models to extract features from the underlying data [31].

III. MODEL AND EXPERIMENT

A. Deep Model

We propose our own sequential 21-layer CNN model shown in Figure 1, for the task of classifying parasitized and uninfected cells towards malaria parasite detection. The proposed CNN has seven convolutional layers and four fully connected layers. The input image to the model constitutes segmented cells of (44 x 44) pixels resolution, with three channels for RGB colors. The convolutional layers except the last convolutional layer which uses a (2 x 2) filter with 1-pixel stride, use (3 x 3) filters with 1-pixel stride. The first and second convolutional layers have 128 filters, the third and fourth convolutional layers have 256 filters, the fifth, sixth, and the last Conv layers have 512 filters. The sandwich design with one convolutional layer followed by one rectified linear unit (ReLU) and proper weight initialization enhances the learning process [11, 32, 33]. Max-pooling layers with a pooling window of (2 x 2)

and 2 pixels strides follow second, fourth, and sixth convolutional layers to merge semantically similar features into one. The output of the seventh convolutional plus ReLU layer is fed to the first and second fully connected layers that have 512 neurons each. The third fully connected layer with 64 neurons, then feeds into the softmax classifier. Dropout regularization with a dropout ratio of 0.2 is applied to the outputs of fully connected layers to prevent the model from overfitting [10]. The model is trained by optimizing the softmax regression objective using adaptive moment estimation (Adam) with beta-one and beta-two parameters of 0.9 and 0.999 respectively [34]. We used L2 regularization to reduce the effect of model overfitting and therefore, converging to a better solution [20]. For the learning rate and L2 regularization parameter we used 0.00005 and 0.0001 respectively.

B. Shallow Model

We also propose another model to evaluate the effect of filter size of convolution layers along with the depth of the model, on the overall performance of the neural network in terms of accuracy. Therefore, we implemented a less deep CNN with five convolution layers with filter size of (5 x 5) and four fully connected layers. This network was derived from the deep model presented in the previous section by replacing each (3 x 3) convolution layer with a (5 x 5) convolution layer. This 16-layer CNN model is shown in Figure 2.

C. Experiment

1) Data Source

We used archived segmented cells of thin blood smear slide images from the Malaria Screener research activity provided by National Library of Medicine. This dataset consists of a total number of 27598 images from human blood cells that contain malaria or not. There are 13799 cell images for each of the malaria infected and uninfected cells. Each sample in this dataset is a 139 by 142 pixels image of human blood cells.

2) CNN Model Training

For training and evaluation of our proposed model, images were normalized and reshaped to 44 by 44 pixels resolution. Then, we used 80% of the whole dataset for training, 10% for validation and the remaining 10% of images for testing. We executed the training process for 100 epochs.

3) Result

The results are presented in Table 1. As we can see our proposed deep model outperforms the pre-trained models introduced in [31]. Our deep model shows superiority over the transfer learning by delivering 96.08% accuracy compared to the best result of transfer learning models which is 95.7% by

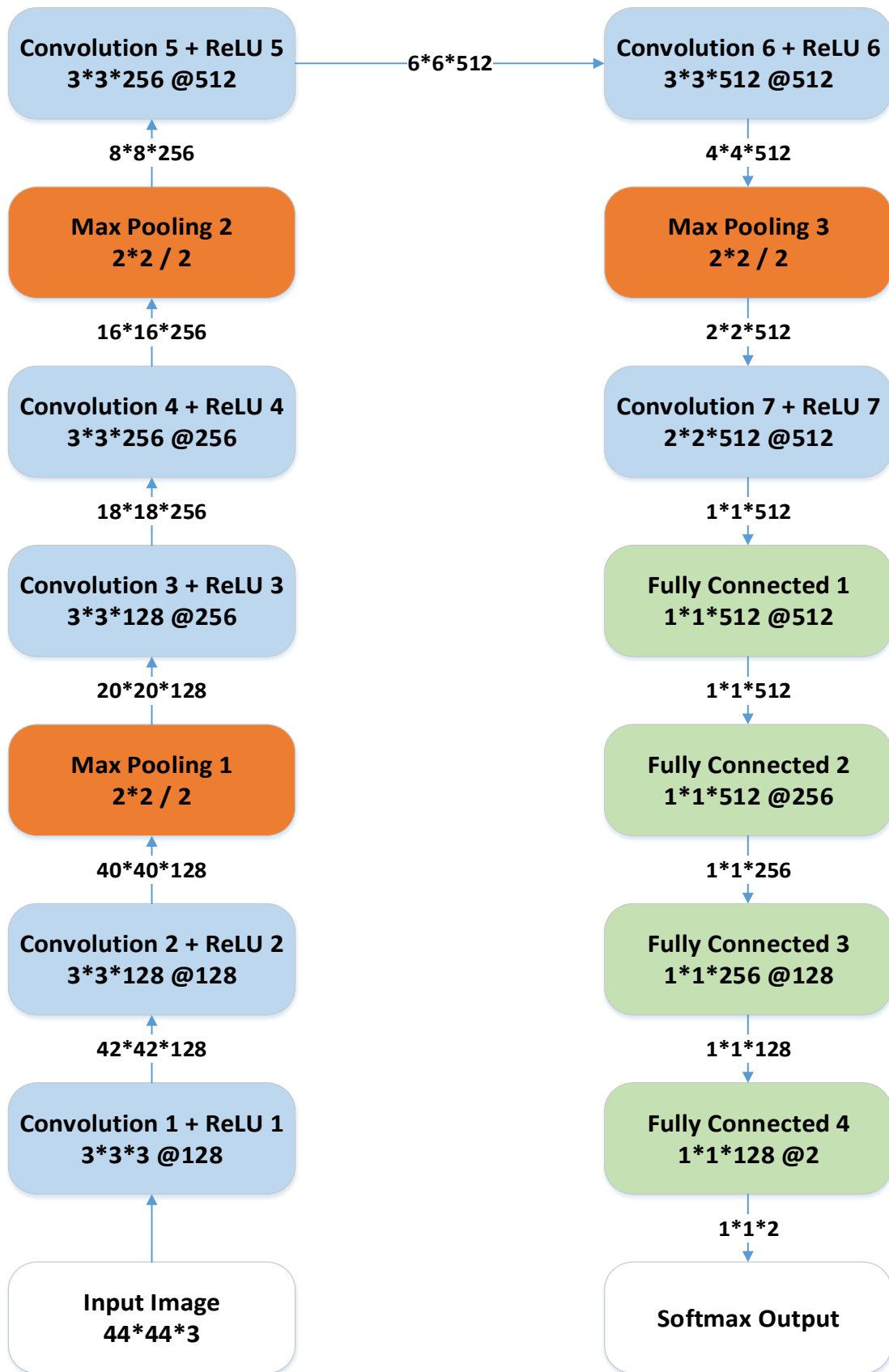


Figure 1. Deep Model Architecture

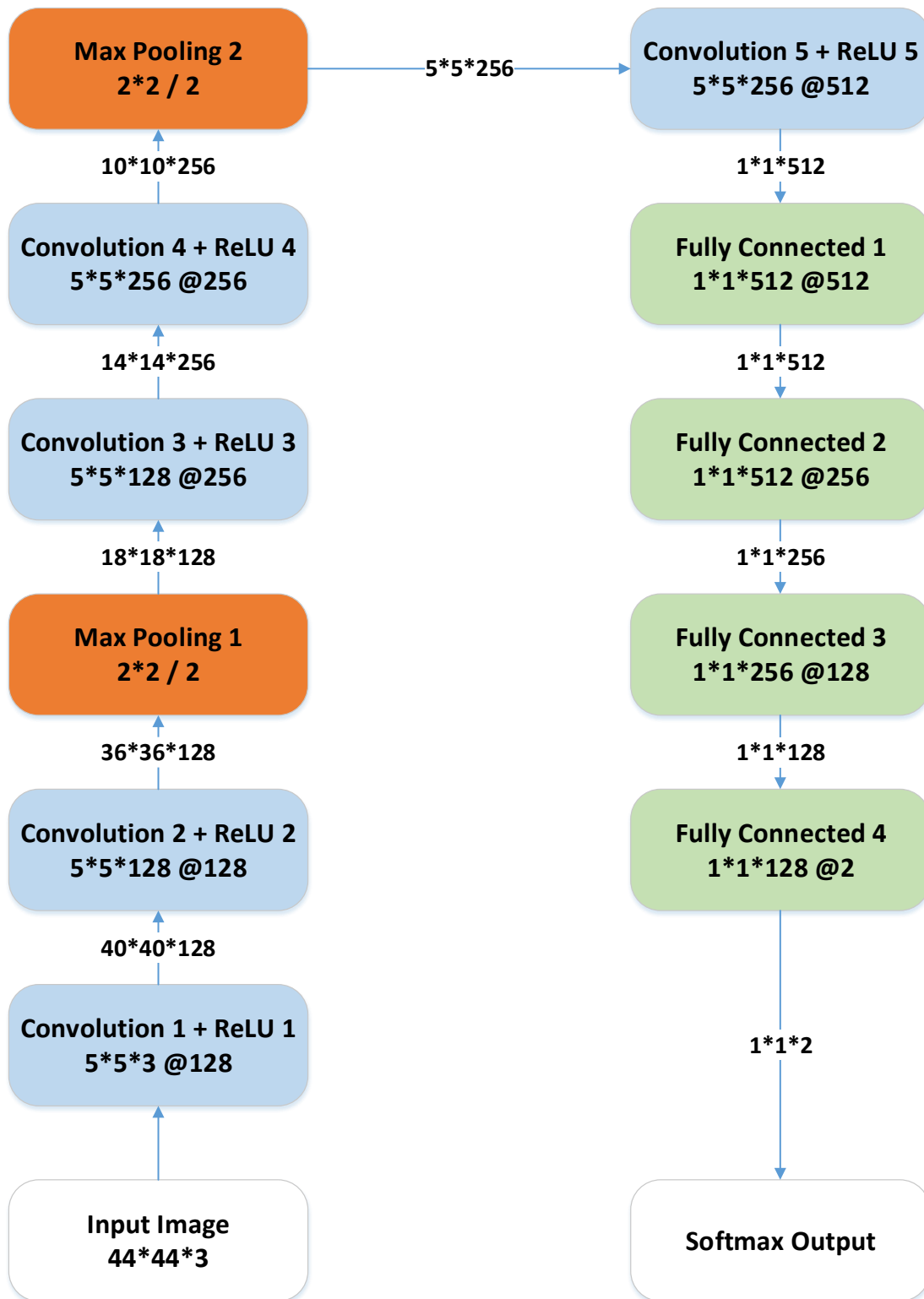


Figure 2. Shallow Model Architecture

Table 1. Performance comparison of models

Model	Accuracy	F1-score
Deep Model	96.08%	95.9%
Shallow Model	95.57%	95.53%
AlexNet [31]	93.7%	93.7%
VGG-16 [31]	94.5%	94.5%
ResNet-50 [31]	95.7%	95.7%
Xception [31]	89%	89.5%
DenseNet-121 [31]	93.1%	93.1%

ResNet-50. This shows that the trained CNN model is a better representation of the training samples compared to the transfer learning models that use the extracted features from a different dataset. In addition, we can see among our proposed models that the deep model which uses smaller receptive fields, outperforms the less deep model with bigger filter sizes.

IV. CONCLUSION AND DISCUSSION

Based on the above results we can conclude that using convolutional neural network models is a suitable approach toward automated malaria detection in blood smear images. Our proposed convolutional neural network model delivers competitive performance in terms of accuracy compared to transfer learning method using pre-trained models such as AlexNet, VGG-16, ResNet-50, Xception, and DenseNet-121 [31], along with other CNN-based solutions.

The architecture of the convolutional neural network and the size of the training dataset are the two main factors affecting the overall performance of the deep learning model. Therefore, we can further improve the performance of our model by increasing the number of training samples by data augmentation techniques [35]. In addition, we showed that using smaller filter sizes in convolution layers which leads to deeper networks, can improve the performance of CNN-based deep neural networks in terms of accuracy while adding to the computational cost.

All in all, using deep learning methods can promise state-of-the-art performance in terms of efficiency and accuracy toward the task of malaria blood smear classification and it can be also applied to other health-related applications.

REFERENCES

- [1] World Malaria Report, World Health Organization, 2016.
- [2] Centers for Disease Control and Prevention. CDC—Malaria. 2012.
- [3] Hommelsheim CM, Frantzeskakis L, Huang M, Ülker B. PCR amplification of repetitive DNA: a limitation to genome editing technologies and many other applications. *Scientific Reports* 4:5052. 2014.
- [4] Hawkes M, Katsuva J, Masumbuko C. Use and limitations of malaria rapid diagnostic testing by community health workers in war-torn Democratic Republic of Congo. *Malaria Journal* 8(1):308. 2009.
- [5] Ross NE, Pritchard CJ, Rubin DM, Dusé AG. Automated image processing method for the diagnosis and classification of malaria on thin blood smears. *Medical & Biological Engineering & Computing* 44:427–436. 2006.
- [6] Das DK, Ghosh M, Pal M, Maiti AK, Chakraborty C. Machine learning approach for automated screening of malaria parasite using light microscopic images. *Micron* 45:97–106. 2013.
- [7] Poostchi M, Silamut K, Maude RJ, Jaeger S, Thoma GR. Image analysis and machine learning for detecting malaria. *Translational Research* 194:36–55. 2018.
- [8] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 521:436–444. 2015.
- [9] Schmidhuber J. Deep Learning in neural networks: an overview. *Neural Networks* 61:85–117. 2015.
- [10] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* 15:1929–1958. 2014.
- [11] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. New York: ACM. 2012.
- [12] Waibel A., Hanazawa T., Hinton G. E., Shikano K. and Lang K. Phoneme recognition using time-delay neural networks. *IEEE Trans. Speech Signal Process*, 37, pp. 328–339. 1989.
- [13] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, “Gradient-based learning applied to document recognition.” *Proceedings of the IEEE*. 1998 Nov; 86(11), pp. 2278–324.
- [14] R. Vaillant, C. Monrocq and Y. LeCun, “Original approach for the localisation of objects in images.” *IEEE Proceedings-Vision, Image and Signal Processing*. 1994 Aug, 141(4), pp. 245–50.
- [15] Dean, J., Corrado, G., Monga, R., Chen, K., Devin, M., Mao, M., Ranzato, M., Senior, A., Tucker, P., Yang, K., Le, Q. V., and Ng, A. Y. Large scale distributed deep networks. In *NIPS*, pp. 1232–1240, 2012.
- [16] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In *Proc. CVPR*, 2009.
- [17] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., and Fei-Fei, L. ImageNet large scale visual recognition challenge. *CoRR*, abs/1409.0575, 2014.
- [18] Zeiler, M. D. and Fergus, R. Visualizing and understanding convolutional networks. *CoRR*, abs/1311.2901, 2013. Published in *Proc. ECCV*, 2014.
- [19] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. Going deeper with convolutions. *CoRR*, abs/1409.4842, 2014.
- [20] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2015.
- [21] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), 770–778. 2016.
- [22] Chollet F. Xception: deep learning with separable convolutions. 1–14. 2016.
- [23] Huang G, Liu Z, Weinberger KQ, Van der Maaten L. Densely connected convolutional networks. 2016.
- [24] S.W. Sio, W. Sun, S. Kumar, W.Z. Bin, S.S. Tan, S.H. Ong, H. Kikuchi, Y. Oshima and K.S. Tan, “Malaria Count: an image analysis-based program for the accurate determination of parasitemia.” *Journal of microbiological methods*. 2007 Jan 31;68(1), pp. 11–8.
- [25] G. Díaz, F.A. González and E. Romero, “A semi-automatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images.” *Journal of Biomedical Informatics*. 2009 Apr 30;42(2), pp. 296–307.
- [26] Rajaraman S, Antani S, Xue Z, Candemir S, Jaeger S. 2017. Visualizing abnormalities in chest radiographs through salient network activations in Deep Learning. In: *Life Sciences Conference (LSC)*, 2017 IEEE. Sydney, NSW, Piscataway: IEEE, 71–74.
- [27] Suzuki K. 2017. Overview of deep learning in medical imaging. *Radiological Physics and Technology* 10:257–273.

- [28] Dong Y, Jiang Z, Shen H, David Pan W, Williams LA, Reddy VVB, Benjamin WH, Bryan AW. 2017. Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells. In: 2017 IEEE EMBS international conference on biomedical and health informatics, BHI 2017. Piscataway: IEEE, 101–104.
- [29] Liang Z, Powell A, Ersoy I, Poostchi M, Silamut K, Palaniappan K, Guo P, Hossain MA, Sameer A, Maude RJ, Huang JX, Jaeger S, Thoma G. 2017. CNN-based image analysis for malaria diagnosis. In: Proceedings—2016 IEEE international conference on bioinformatics and biomedicine, BIBM 2016. Piscataway: IEEE, 493–496.
- [30] Bibin D, Nair MS, Punitha P. 2017. Malaria parasite detection from peripheral blood smear images using deep belief networks. *IEEE Access* 5:9099–9108.
- [31] Rajaraman et al. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *PeerJ* 6:e4568; DOI 10.7717/peerj.4568. 2018.
- [32] A. Hyvarinen and E. Oja, “Independent component analysis: algorithms and applications,” *Neural networks*, 2000 Jun 30, 13(4), pp.411–430.
- [33] Shang W, Sohn K, Almeida D, Lee H. 2016. Understanding and improving convolutional neural networks via concatenated rectified linear units. In: Proceedings of 33rd international conference on machine learning (ICML2016), vol. 48. Princeton: International Society of Machine Learning (ISML), 2217–2225.
- [34] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [35] Wang, Jason, and Luis Perez. "The effectiveness of data augmentation in image classification using deep learning." *Convolutional Neural Networks Vis. Recognit.* 2017.