RESEARCH ARTICLE

Malaria Disease Cell Classification With Highlighting Small Infected Regions

MILKISA YEBASSE[ID]1, **KYUNG JOO CHEOI**[ID]2, **AND JAEPIL KO**[ID]1
[1]Department of Computer Engineering, Kumoh National Institute of Technology, Seowon-gu, Gumi-si, Gyeongbuk 39177, South Korea
[2]Department of Computer Science, Chungbuk National University, Seowon-gu, Cheongju-si, Chungbuk 28644, South Korea

Corresponding author: Jaepil Ko (nonezero@kumoh.ac.kr)

ABSTRACT Deep learning-based methods have become an active research area in medical imaging. Malaria is diagnosed by testing red blood cells. Deep learning methods can be used to distinguish malaria infected cell images from non-infected cell images. The small number of malaria dataset may limit the application of deep learning. Moreover, the infected area in the cell images is generally vague and small, requiring more complex models and a larger dataset to train on. Motivated by the tendency of humans to highlight important words when reading, we propose a simple neural network training strategy for highlighting the infected pixel regions that are mainly responsible for malaria cell classification. In our experiments on the NIH(National Institutes of Health) malaria dataset available in public domain, the proposed method significantly improved classification accuracy for our four different sized models, ranging from simple to complex including Resnet and Mobilenet. Our proposed method significantly improved classification accuracy. The result indicate that approach achieves a classification accuracy of 97.2%, compared to 94.49% for a baseline model. In addition, we show the superiority of the proposed strategy by providing an analysis on the magnitude of weight parameters in terms of regularization.

INDEX TERMS Image processing, convolutional neural network, biomedical imaging, image classification, machine learning, image segmentation.

I. INTRODUCTION

Malaria is a fatal disease caused by Plasmodium parasites that infect red blood cells (RBCs) [1]. Infected mosquitoes transmit the parasite through their bites to humans. Nearly all malaria cases occur in developing countries, primarily in Sub-Saharan Africa. More than 290 million people are infected with malaria annually, and more than 400,000 die [2]. Malaria is diagnosed by testing the red blood cells. In order to determine the presence of malaria, centrifugal machines are used to isolate RBC and WBC so that only RBC can be used for analysis by blood films. Blood smears are used to diagnose malaria and are a standard laboratory test [3]. Deep learning methods can be used to distinguish malaria infected cells from noninfected cells. However, the success of such services depends on the availability of the dataset used. Moreover, the difference between normal and defective images is generally

small (as shown in Figure 1). In such cases, it is difficult to train a neural network, and the process can be more complex and take a long time.

Currently, deep learning methods require a large number of training samples and a significant amount of time to label the training samples [4]. A DNN is rarely suitable for tasks with a small amount of training data. The diversity and size of the data are important factors in improving classification performance [5]. Typically, large datasets lead to a better classification performance, and small datasets may trigger an overfitting [6]. Nevertheless, in the real world, particularly with medical datasets, there are a relatively small number of datasets available. It can be expensive or simply impossible to gather a large dataset. In such cases, it is critical to make the most accurate predictions possible. Motivated by humans' tendency to draw attention to important phrases in reading or writing, we propose a simple neural network training strategy based on highlighting malaria-infected pixels that are mainly responsible for classification. Spurred by highlighting

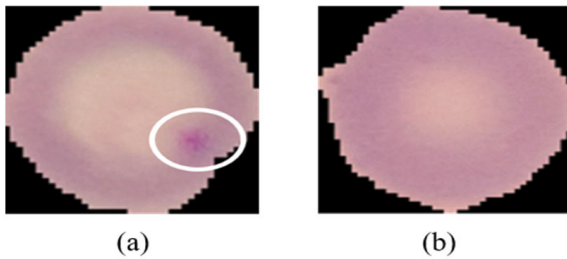The associate editor coordinating the review of this manuscript and approving it for publication was Yongming Li[ID].

FIGURE 1. Sample images: (a) infected: The white boundary shows the infected region. (b) Uninfected red blood cells.

advantage, Serra et al. developed [7] a method to transform sketched images into real images. This method is in line with our highlighting strategy, except that it emphasizes the output layer instead of the input layer.

The current approaches to image classification and object recognition requires the use of attention models [8]. The intuition behind this can be best explained using human biological systems. The human brain does not respond to an entire scene. Instead, humans focus their attention selectively on parts of the visual space to acquire information while ignoring other irrelevant information [9], guiding future eye movements and decision making. Minh proposed a model that uses recurrent neural networks to extract useful information from an image by adaptively selecting a sequence of regions and processing only those regions [10]. The use of a self-attention-based architecture, notably a transformer, has become the de facto approach for natural language processing (NLP). Vision transformers have attained excellent results when trained with larger datasets (14–300 million images). However, transformers lack some of the advantages of a convolutional neural networks (CNN), such as translation equivariance and locality, and therefore do not generalize well when trained on insufficient numbers of data [11]. As a result, it is difficult to apply an attention-based model to small-sized medical datasets. Highlighting serves the same purpose as attention models. Attention models, however, emphasize features in the intermediate layers. Our proposed highlighting strategy differs in the way important regions are highlighted at the input rather than the intermediate layer. We propose a method for guiding a neural network from an input by highlighting the diseased region. By using the highlighting method, we are able to use a light model with fewer parameters. Our model has been tested in four different size modes, including Resnet and Mobilnet. Resnets are one of the most efficient Neural Network Architectures, as they help in maintaining a low error rate much deeper in the network. Hence, proved to perform really well where deep neural networks are required, such as feature extraction, semantic segmentations, various Generative Adversarial Network architectures. We demonstrate the effectiveness of our approach in the deepest neural networks. Mobilenet is a lightweight deep neural network with fewer parameters than Resnet. Mobilenet represents the deepest yet lightest model in our study. To summarize, our contributions are as follows:

1. This paper proposes a simple neural network training strategy of highlighting motivated by humans highlighting important phrases when reading or writing.
2. In this paper, we demonstrate that the proposed method significantly improved classification accuracy for our four different sized models, ranging from simple to complex including Resnet and Mobilenet. The highlighting method allows us to use light model of small number of parameters.
3. This paper also demonstrates that the superiority of the proposed strategy by providing analysis on the magnitude of weight parameters in terms of the regularization.

II. RELATED WORK

In this section, traditional machine learning methods and optimization strategies for deep learning methods are reviewed. In a traditional machine learning pipeline, the process of malaria diagnosis involves four steps: image preprocessing, cell segmentation, feature selection, and classification of infected and non-infected cells. An image preprocessing method improves the quality of blood smear images. There are a variety of smothering filters, including median, geometric mean, and Gaussian filters that can be used to reduce noise in microscopy images [12], [13]. In addition, a morphological operator can also improve the cell contours, removing impurities and suppressing noise by covering holes in the cell [14]. Cell segmentation is the most significant step in traditional automated malaria detection systems. In [15], the segmentation of RBCs from enhanced images is achieved using the Otsu threshold. In malaria detectors such as in [16], a zach thresholding is applied to microscopic images to segment cells. In addition, various feature extraction methods have also been used to extract features from cell images [17], [18]. Feature extraction methods selects the most relevant features for a model to predict the target variable. Channel selection [19], [20] and feature extraction [21] are also important in signal classification. Traditional malaria detection algorithms have adopted a linear Euclidean distant classifier with a Poisson distribution [22], a support vector machine and an artificial neural network [23], and K-means clustering [24] for classification. In contrast to traditional detection methods, deep learning (DL) tends to solve problems end-to-end, bypassing a feature selection step. CNN models [25] can recognize a pattern in microscopic images with a much higher accuracy rate than other traditional approaches. Several methods have recently been introduced to enable a faster and more stable training in deep learning [26], [27], [28]. In [29], the authors proposed transfer learning for malaria disease classification. The use of novel data augmentation techniques [30] and various CNN schemes have also been investigated for disease diagnosis in blood smear images [31]. A number of important considerations must be taken into account, including enhancing model weight initialization by transfer learning and utilizing dropout as a method of regularization in order

FIGURE 2. Sample Pipeline of proposed approach.
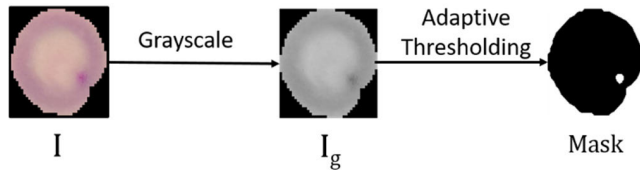

FIGURE 3. Shows process of segmentation.

to combat overfitting during model training [32], [33], [34]. There have been Regularization aims to enhance the training of neural networks by stabilizing the distribution of the layer inputs [35]. The use of traditional methods as guidance for deep neural networks has not been sufficiently explored. Thus, we propose a method for guiding a neural network from an input by highlighting the diseased region through image processing techniques.

The possibilities of using traditional methods as guidance for deep neural networks have not been sufficiently explored. In this paper, we present a highlighting approach as guidance to the CNN. Therefore, we highlight the diseased region of a malaria cell to guide a neural network, enabling a faster training process.

III. PROPOSED HIGHLIGHTING STRATEGY

The main goal of our proposed approach is to guide the neural network to classify malaria infected cells from healthy non-Infected cells. Guiding neural network can improve the classification accuracy [35]. The proposed approach comprises several stages: The proposed method consist of three parts: Segment infected region, highlighting the region of interest, and classification. The details of each stage are described in the following subsections. Figure 2 shows the architecture of the proposed system.

A. SEGMENT INFECTED REGIONS

The adaptive threshold is a segmentation method that helps us to identify the areas of an image corresponding to specific diseases in an easy to understand manner. To distinguish objects from backgrounds, we use the difference in intensity between object pixels and background pixels. A thresholding algorithm segments an image according to a certain characteristic of its pixels (for instance, intensity). Adaptive thresholding algorithm determines pixel based threshold on a small region around it. Different thresholds are applied to different regions of the same image, which gives better results for images with varied illumination. The dataset contained images of normal RBCs and images of RBCs infected with malaria. Compared

---

Algorithm 1 Highlighting Strategy

Input folder: A path to main folder of malaria dataset
Output: Highlighted image
For (image < input-folder) do
    Grayscale_image ← change image to Grayscale image
    Masked image ← apply adaptive threshold on Grayscale_image
    Extracted_image ← apply bitwise operation between image and masked image
    I_b, I_g, I_r ← split image
    Ir' ← add scaler value to r channel to highlight the selected area of masked image
    Highlighted image ← merge I_b, I_g and I_r'.
end

---

to an image of a healthy cell, an image of an infection displays darker red anomaly regions, indicating the infection.

B. HIGHLIGHT INFECTED REGIONS

In this study, we propose simple highlighting technique tailored to malaria disease classification. Following the extraction of the regions of interest, the masked images were used to highlight the diseased area. Our goal was to increase the intensity of the Red channel to highlight the diseased region. H is scaler value and is computed from mean and standard deviation of the dataset. First, we split the original image into r, g, and b channels. We add a constraint in the R channel to highlight the selected area based on the masked image, as shown in Eq.1. A scale will first be multiplied by the masked region, then it will be added to the R channel of the image. A scalar value (S) is added to the red channel of the image based on the masked area.

$$I'_R(x, y) = Min((I_R(x, y) + S * M(x, y)), 255) \qquad (1)$$

where $I'_R(x, y)$ the highlighted pixel intensity of (x,y) on the R channel, $I_R(x, y)$ non-highlighted pixel intensity of (x,y) on the R channel, S is highlighting scale, M(x,y) is segmentation of infected pixels and Min() represent operator of taking minimum.

We chose the r channel for adding the scale factor since the malaria dataset has a red infected region. The scale factor is determined by the mean and standard deviation. Following completion of the above steps, we finally merge the b, g, and r channels (Figure 4). Algorithm 1 shows the entire process for the highlighting steps.
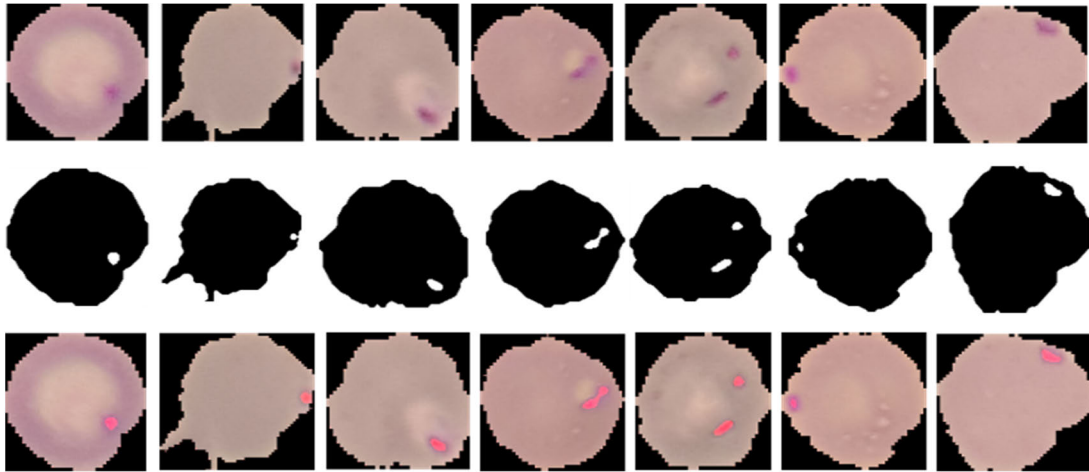
FIGURE 4. Example of infected cell images. First row: Infected images, Second row: extracted diseased area and third row: highlighted image.
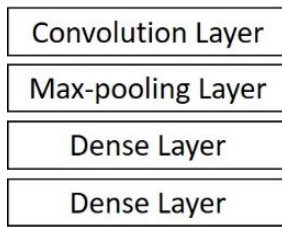

FIGURE 5. Layers of Model-S.

C. CLASSIFICATION

Large deep neural networks have achieved remarkable success with a good performance, particularly in real-world scenarios with large-scale data and extremely complex models. However, medical datasets contain relatively small numbers of data. Moreover, the deployment of deep models in mobile devices and embedded systems is a significant challenge because of the limited computational capacity and memory of such devices. In this study, we guided the neural network by highlighting the regions of interest. Thus, we were able to achieve our goal using a shallow network. Categorical cross-entropy was used as a loss function for our model.

In this study, we developed and employed four different size models to demonstrate the effectiveness of our highlighting strategy.

1. Small Model of few parameters
2. Medium Model of moderate number of parameters.
3. Resnet of huge number of parameters
4. Mobilnet of light model

1) SMALL MODEL: MODEL-S

The CNN model consists of two components: feature extraction and classification. Feature extraction is performed by the convolution layer, while Classification is performed by fully connected layers. The network consists of only one convolutional layer for feature representation and two dense layers (Figure 5). The final dense layers serve as the classification

TABLE 1. The weights and parameter of the Model-S.

layer	Output Size	Parameters
Input	3*150*150	
Conv(cout=3,k=7,s=a)	3*144*144	444
Maxpool(k=2)	3*72*72	
Dense layer(15552→64	64	995392
Dense layer(64→2)	2	130
Total		995966

layers. Categorical cross-entropy was used as a loss function for our model. Our model was trained using various hyperparameters, i.e., a batch size of 20, a learning rate of 0.001, 25 epochs, an ADAM optimizer, and categorical cross-entropy loss.

The small model has three filters in the convolution layer. After the conv layer, there is a maxpooling layer followed by two final dense layers as shown in Table 1.

2) MEDIUM MODEL: MODEL-M

Model-M has deeper convolution and maxpooling layers than Model-S, resulting in a higher number of parameters. Model-M is composed of three convolutional layers and three maxpooling layers (Figure 6). Our model was trained using various hyperparameters, i.e., a batch size of 20, a learning rate of 0.001, 25 epochs, an ADAM optimizer, and categorical cross-entropy loss.

The medium model has 64 filters in conv layer, which is significantly higher than small model. Two dense layers are added final maxpool layer in which the final dense layer is used as classification layer. Model-M is slightly more complex than Model-S, but it is lighter when compared against both Resnet and Mobilenet(as shown in Table 2).

TABLE 2. Weight and parameters of the Mode-M.

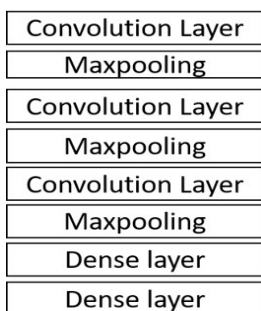| layer	Output Size	Parameters
Input	3*150*150	
Conv(cout=64,k=3,s=1)	3*144*144	1792
Maxpool(k=2)	3*72*72	
Conv(cout=64,k=3,s=1)	3*144*144	36,928
Maxpool(k=2)	3*72*72	
Conv(cout=64,k=3,s=1)	3*144*144	36,928
Maxpool(k=2)	3*72*72	
Dense layer(18496→64	64	1,183,808
Dense layer(64→2)	2	130
Total		1,259,586

FIGURE 6. Layers of Model-M.

3) DEEP MODEL: RESNET

In our study, we chose Resnet [37] to represent the deepest network. There are several variations of Resnet that are based on the same concept, but have varying numbers of layers. We have used Resnet-50 which is a convolutional neural network that has 50 layers. Images in the medical domain have different structure than those in the normal image domain. Thus, the entire model is trained with the newly added classifier. Resnet was trained using a batch size of 20, a learning rate of 0.0001, 25 epochs, and RMSprop optimizer, and categorical cross-entropy loss.

4) LIGHT MODEL: MOBILENET

Mobilenet [38] represents the deepest yet lightest model in our study. Mobilenet-v2 is a convolutional neural network that is 53 layers deep. The Mobilenet v2 architecture is based on an inverted residual structure where the input and output of the residual block are thin bottleneck layers opposite to traditional residual models, which use expanded representations in the input. Mobilenet v2 uses lightweight depthwise convolutions to filter features in the intermediate expansion layer.

IV. EXPERIMENTS, RESULTS, AND DISCUSSION

In this section, we provide the results of an experiment conducted to validate the proposed method. Our study involves
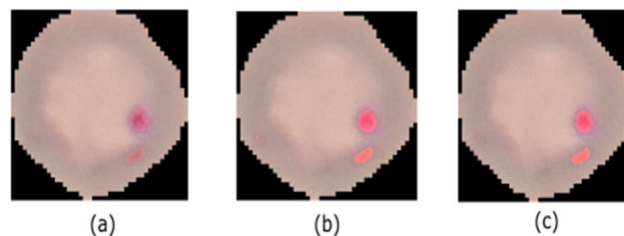

FIGURE 7. Shows process of segmentation. $H_1 = \mu - \sigma = 35$, (b). $H_2 = \mu = 113$, and (c). $H_3 = \mu + \sigma = 180$.

classifying microscopic images of smeared thin blood that had already been verified by trained microscopists as infected or uninfected by malaria. We divided our experiments into three parts. To begin with, we experimented on how to choose scale factor, then we demonstrated the speed of convergence in loss slope, followed by the magnitude of weight, and finally compared accuracy with different models. This research used 19000 images. We evaluated the predictive models through a five-fold cross-validation over five different test sets. We evaluated the predictive models through a five fold cross validation over five different test sets. The training data were randomly partitioned into five equal sized subsets, with one subset used for validation testing, and the remaining four subsets used for training. The cross validation process was then repeated five times for the proposed model, with each of the five subsets used exactly once as the validation data. The results of the validations were averaged to produce a single score.

A. DATASET

In this paper, we propose a guided model for malaria parasite detection that achieves a higher accuracy. Malaria is a fatal disease caused by Plasmodium parasites that infect red blood cells (RBC). Our experiment was conducted on images of parasitized (infected) and uninfected RBCs from the NIH Malaria dataset [39], which were collected from 201 patients and is classified into two groups, malaria-infected and uninfected, for which there are equal numbers of instances in each cell. Cells with P. falciparum were placed on a conventional light microscope at Chittagong Medical College Hospital, Bangladesh, and photographed using a smartphone.

B. SCALE FACTOR FOR HIGHLIGHTING

Our primary objective with the proposed system is to guide neural networks by highlighting the diseased regions. To highlight the diseased areas, we must add a scaler factor (H) to the diseased region. We experimented on proposed approach 3 times by varying the scaler factor. We selected the scaler factor (H) based on the mean and standard deviation of the dataset. $\mu = 113$ and $\sigma = 78$. We experimented on three values. The highest accuracy was obtained in H3. These values are added to the diseased area to highlight the region of interest. Figure 7 shows the highlighted image with respect to scaling.
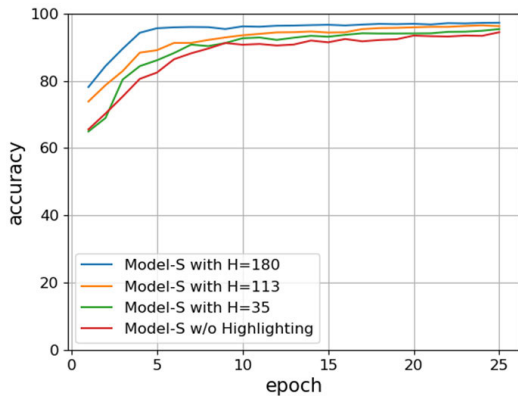
FIGURE 8. Accuracy of highlighted approach with different scale factor.

TABLE 3. Training accuracy and test accuracy in Model-S with different scale factor.

Highlighting scale	Training Accuracy%	Test Accuracy%
Model-S with H=180	98.6±0.5	97.21± 0.3
Model-S with H=130	98.0± 0.6	96..49± 0.5
Model-S with H=35	96.7±0.6	94.88± 0.9
Model-S w/o Highlighting	96.3±0.8	94.49± 1.1

TABLE 4. Comparison of model.

	Model-S	Model-M	Mobilenet	Resnet
Total Parameter	995,966	3,504,872	3,504,872	25,608,258
Parameter size (MB)	3.8	4.88	13.37	97.69
Estimated Total Size(MB)	4.53	18.85	20.96	177.85

We highlighted or marked the diseased regions on the cells. By doing so, we made the problem easier and the network is guided to focus on the most significant regions of the image for classification. Then, with each approach, we trained our network. First, the accuracy of each highlighted dataset with a different scale factor is compared. The highest accuracy was obtained in H3 (Figure 8). For the rest of the experiment, we will use H3 for our proposed approach unless stated otherwise. We compared our results with baseline in order to validate our method. The test accuracy of the proposed approach was 97.21%, which was significantly higher than that of the baseline(Model-s w/o Highlighting) 94.8% (Table 3). We demonstrated the importance of our model by comparing it with baseline. In all aspects, our proposed approach outperforms the baseline.

TABLE 5. Accuracy of the model-S, model-M, mobilenet and Resnet.

Highlighting	Model-S %	Model- M %	Mobilenet %	Resnet %
With	97.21± 0.3	97.56± 0.4	97.9± 0.32	98.07± 0.8
W/O	94.49± 1.1	95.8± 0.6	97.1± 0.5	97.4± 0.6

TABLE 6. Precision, Recall and F1 score of the Model-S.

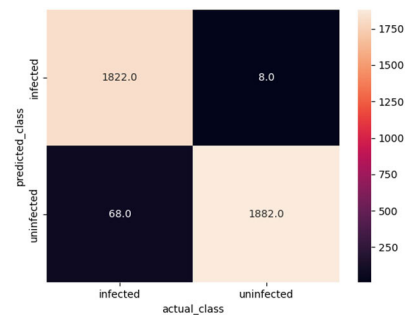| Highlighting scale	Precision%	Recall%	F1 measure%
Model-s with H=180	99.53± 0.2	94.90± 0.3	97.0± 0.2
Model-s with H=130	98.8± 0.4	93.60± 0.5	96.1±0.25
Model-s with H=35	96.02±0.6	93.51 ± 0.5	94.7± 0.4

FIGURE 9. Confusion matrix.

C. ACCURACY WITH DIFFERENT MODEL

The accuracy and parameter values of our 4 models were compared. The Model-S has the smallest parameter and depth. The Model-M is more complex than Model-S. The weight values for models is described below. We have used Resnet to represent complex models, while Mobilenet is used to represent complex light models. The Model-S has the smallest parameter while Resnet model have largest parameter (as shown in Table 4).

The weight values for models is described below. To show efficiency of our proposed approach, we show comparison test accuracy over the complex models as well. We obtained 97.21% using the proposed method, outperforming models-s without highlighting at 94.49% and models-m without highlighting at 95.8%. Additionally, the proposed approach with Resnet and Mobilenet models outperformed the respective baseline model (as shown in Table 5). With all models representing different parameters and depths.

D. PRECISION, RECALL AND F1 SCORE

We further evaluated the proposed approach by computing the recall, precision, and F1 scores. Precision refers to the proportion of positive predictions that were actually correct (true positives). Recall measures how many positive cases
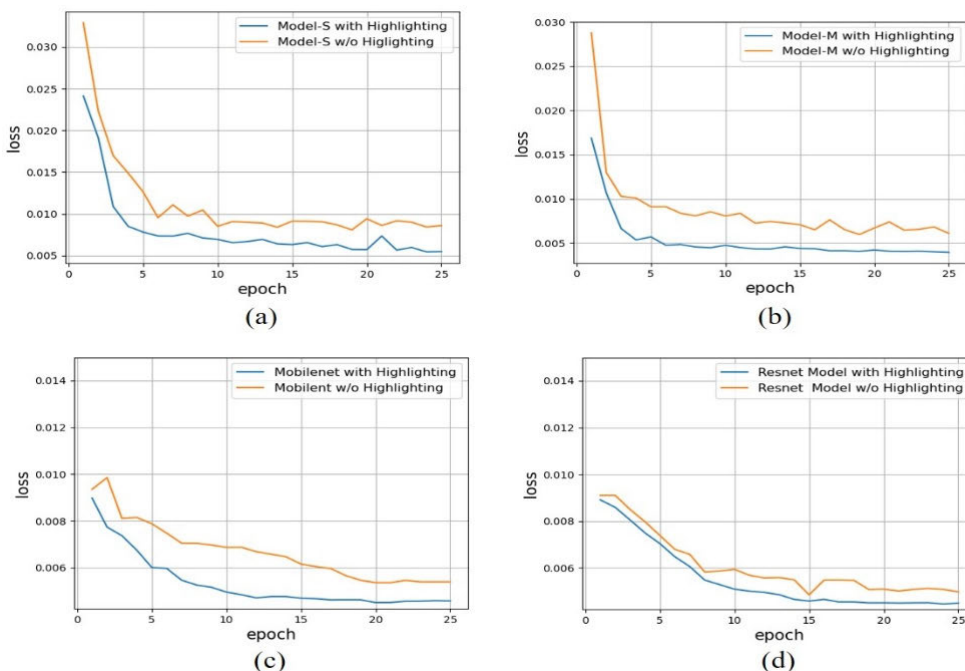
FIGURE 10. Loss graph: (a) Model-S, (b) Model-M, (c) Mobilenet model and (d) Resnet.
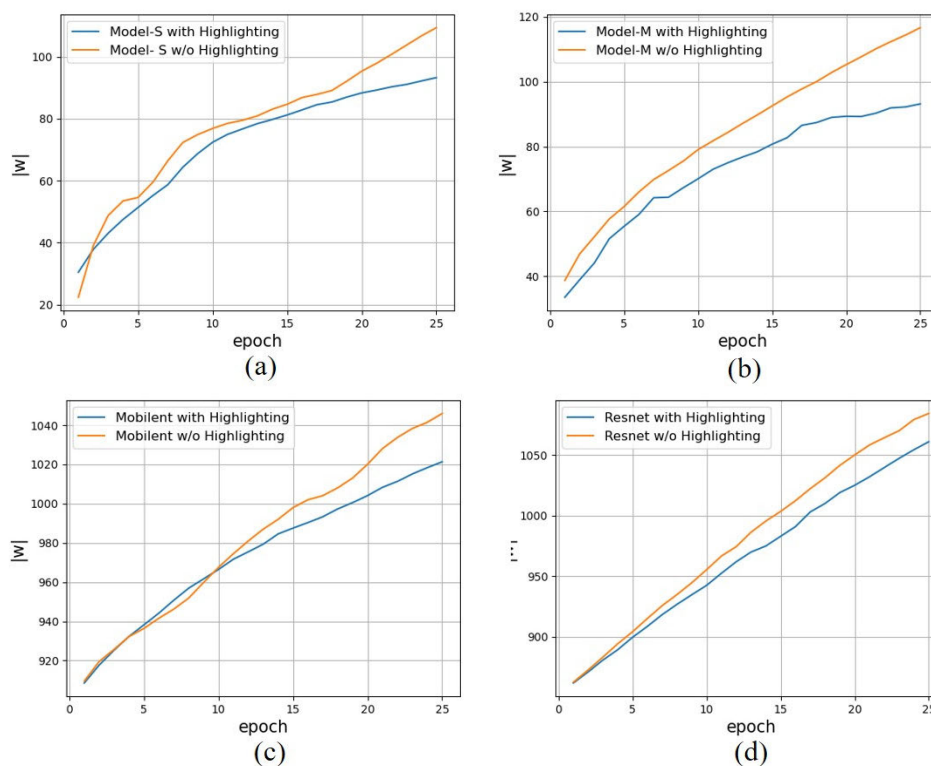

FIGURE 11. Magnitude of weight graph: (a) Model-S, (b) Model-M, (c) Mobilenet and (d) Resnet.

were correctly predicted by a classifier, compared to all positive cases in the data. F1-Score combines precision and recalculation into one measure. For Model-S, optimal results were achieved by using scale factor 180 and F1 measure rate

of 97, recall rate of 94.9%, precision rate of 99.21%, while the lowest was achieved in Model-S with scale factor of h-35, F1 score of 96.02%, precision rate of 93.5%, and recall rate of 94.7% (Table 6).

E. CONFUSION MATRIX

The confusion matrix result for the proposed system in Model-S is demonstrated in figure 9. A predicted class can be either infected or uninfected. Out of 3780 predictions, the classifier correctly predicted "infected" 1822 times and "uninfected" 1882 times.in reality, 1890 cells are "infected" and 1890 cells are "uninfected".

F. LOSS CONVERGENCE

To investigate the effect of the highlighting in the model, we report the loss behavior for different scenarios. The loss by highlighting approach, in overall, is better than the one obtained by baseline. The highlighted approach achieved the lowest loss and reached convergence much faster than the baseline. Loss graphs of the proposed and conventional approaches are shown in Figure 10. We can observe that the slope of the proposed approach is steeper than that of the baseline.

To further proof of our idea, we add our strategy to complex model, Resnet and Mobilenet. Loss difference between the proposed approach (highlighted images) and baseline (non-highlighted image) is clearly visible in Model-M and Model-S. We demonstrate that highlighting approach reaches convergence faster and have steeper slope. The proposed approach still results in lower losses for Resnet and Mobilenet. Given how powerful these models are, it is reasonable to expect that baseline can also achieve comparable performance with the proposed approach. Overall, we found that our proposed strategy can speed up training and reach convergence faster.

G. MAGNITUDE OF WEIGHT COMPARISION

Representative regularization techniques are the weight decay [34] and dropout [40]. Adding all the parameters (weights) to the loss function would be one way to penalize complexity. This process of regularization is often referred to as weight decay because it reduces the weights. Our highlighting approach also reduces weights. Therefore, it can be considered a regulation approach. The goal of regularizations is to avoid overfitting by penalizing weights in order to fit the function appropriately. The baseline weight norm grows more steadily than the proposed approach, resulting in a larger weight. Regularization encourages the weight values to decrease. Our highlighting method reduces the weight parameter, which can be regarded as regularization. We can state that the proposed approach is superior to the baseline and generalizes well. Figure 11 Shows the norm of weight of Mobilenet and Resnet.

V. CONCLUSION

In this work, we present a new training strategy based on highlighting infected regions to improve classification accuracy. The advantage of the proposed method is verified on NIH malaria dataset.

To show the efficiency of proposed approach, we have tested our approach on four different sized models, ranging from simple to complex models (Resnet). The result showed that the classification accuracy of proposed method is higher than baseline, especially with the smaller model (Model-S), obtaining 97.21% compared to the baseline of 94.49%. Moreover, our approach consistently outperformed the baseline independent of the model size. We further showed superiority of the proposed approach by analyzing weight parameters in terms of the regularization. The proposed approach reduces the weight and improves generalization. The limitation of the proposed approach is that it is customized specifically for these malaria datasets. In the future, we plan to develop a novel method for automatically highlighting a region of interest in complex images, or any images in general and also expand the study to include application of detection, segmentation and scene analysis by highlighting the detection area.

REFERENCES

[1] S. Rajaraman, S. Jaeger, and S. K. Antani, "Performance evaluation of deep neural ensembles to toward malaria parasite detection in thin-blood smear images," *PeerJ*, vol. 7, p. 6977, May 2019.

[2] L. M. Barat, A. Mills, S. Basu, N. Palmer, K. Hanson, and E. Worrall, "Do malaria control interventions reach the poor? A view through the equity lens," *Amer. J. Tropical Med. Hygiene*, vol. 71, pp. 174–178, Aug. 2004.

[3] S. Sankaran, M. Malarvel, G. Sethumadhavan, and D. Sahal, "Quantitation of malarial parasitemia in Giemsa stained thin blood smears using six sigma threshold as preprocessor," *Optik*, vol. 145, pp. 225–239, Sep. 2017.

[4] X. Sun, J. Gu, and H. Sun, "Research progress of zero-shot learning," *Int. J. Speech Technol.*, vol. 51, no. 6, pp. 3600–3614, Jun. 2021.

[5] S. Hyontai, "Performance of machine learning algorithms and diversity in data," in *Proc. MATEC Web Conf.*, vol. 210, 2018, p. 04019.

[6] M. Sordo and Q. Zeng, "On sample size and classification accuracy: A performance comparison," in *Proc. Int. Symp. Biol. Med. Data Anal.* Cham, Switzerland: Springer, 2005, pp. 193–201.

[7] E. Simo-Serra, S. Iizuka, K. Sasaki, and H. Ishikawa, "Learning to simplify: Fully convolutional networks for rough sketch cleanup," *ACM Trans. Graph.*, vol. 35, no. 4, pp. 1–11, Jul. 2016.

[8] S. Chaudhari, V. Mithal, G. Polatkan, and R. Ramanath, "An attentive survey of attention models," *ACM Trans. Intell. Syst. Technol.*, vol. 12, no. 5, pp. 1–32, Oct. 2021.

[9] R. A. Rensink, "The dynamic representation of scenes," *Vis. Cognition*, vol. 7, nos. 1–3, pp. 17–42, Jan. 2000.

[10] V. Mnih, N. Heess, and A. Graves, "Recurrent models of visual attention," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 1–9.

[11] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16×16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.

[12] T. Fatima and M. S. Farid, "Automatic detection of plasmodium parasites from microscopic blood images," *J. Parasitic Diseases*, vol. 44, no. 1, pp. 69–78, Mar. 2020.

[13] D. K. Das, M. Ghosh, M. Pal, A. K. Maiti, and C. Chakraborty, "Machine learning approach for automated screening of malaria parasite using light microscopic images," *Micron*, vol. 45, pp. 97–106, Feb. 2013.

[14] M. Poostchi, K. Silamut, R. J. Maude, S. Jaeger, and G. Thoma, "Image analysis and machine learning for detecting malaria," *Transl. Res.*, vol. 194, pp. 36–55, Apr. 2018.

[15] S. S. Savkare and S. P. Narote, "Automatic system for classification of erythrocytes infected with malaria and identification of parasite's life stage," *Proc. Technol.*, vol. 6, pp. 405–410, Jan. 2012.

[16] K. Chakrabortya, A. Chattopadhyayb, A. Chakrabarti, T. Acharyad, and A. K. Dasguptae, "A combined algorithm for malaria detection from thick smear blood slides," *J. Health Med. Inf.*, vol. 6, no. 1, pp. 645–652, 2015.

[17] Z. Zhang, L. L. S. Ong, K. Fang, A. Matthew, J. Dauwels, M. Dao, and H. Asada, "Image classification of unlabeled malaria parasites in red blood cells," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 3981–3984.

[18] N. Linder, R. Turkki, M. Walliander, A. Martensson, V. Diwan, E. Rahtu, M. Pietikäinen, M. Lundin, and J. Lundin, "A malaria diagnostic tool based on computer vision screening and visualization of *plasmodium falciparum* candidate areas in digitized blood smears," *PLoS ONE*, vol. 9, no. 8, Aug. 2014, Art. no. e104855.

[19] A. Tiwari and A. Chaturvedi, "Automatic channel selection using multi-objective X-shaped binary butterfly algorithm for motor imagery classification," *Exp. Syst. Appl.*, vol. 206, Nov. 2022, Art. no. 117757.

[20] A. Tiwari and A. Chaturvedi, "Automatic EEG channel selection for multiclass brain-computer interface classification using multiobjective improved firefly algorithm," *Multimedia Tools Appl.*, vol. 82, no. 4, pp. 5405–5433, 2022.

[21] A. Tiwari and A. Chaturvedi, "A hybrid feature selection approach based on information theory dynamic butterfly optimization algorithm for data classification," *Exp. Syst. Appl.*, vol. 196, Jun. 2022, Art. no. 116621.

[22] M. S. Suryawanshi and V. Dixit, "Improved technique for detection of malaria parasites within the blood cell images," *Int. J. Sci. Eng. Res.*, vol. 4, pp. 373–375, Jan. 2013.

[23] M. Poostchi, I. Ersoy, K. McMenamin, E. Gordon, N. Palaniappan, S. Pierce, R. J. Maude, A. Bansal, P. Srinivasan, L. Miller, and K. Palaniappan, "Malaria parasite detection and cell counting for human and mouse using thin blood smear microscopy," *J. Med. Imag.*, vol. 5, no. 4, 2018, Art. no. 044506.

[24] Y. Purwar, S. L. Shah, G. Clarke, A. Almugairi, and A. Muehlenbachs, "Automated and unsupervised detection of malarial parasites in microscopic images," *Malaria J.*, vol. 10, no. 1, pp. 1–11, 2011.

[25] N. O'Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. V. Hernandez, L. Krpalkova, D. Riordan, and J. Walsh, "Deep learning vs. traditional computer vision," in *Proc. Sci. Inf. Conf.* Cham, Switzerland: Springer, 2019, pp. 128–144.

[26] V. Sze, Y.-H. Chen, T.-J. Yang, and J. S. Emer, "Efficient processing of deep neural networks: A tutorial and survey," *Proc. IEEE*, vol. 105, no. 12, pp. 2295–2329, Dec. 2017.

[27] V. Vanhoucke, A. Senior, and M. Z. Mao, "Improving the speed of neural networks on CPUs," in *Proc. NIPS*, 2011, pp. 1–8.

[28] J. Bae, J. Lee, Y. Jin, S. Son, S. Kim, H. Jang, T. J. Ham, and J. W. Lee, "FlashNeuron: SSD-enabled large-batch training of very deep neural networks," in *Proc. 19th USENIX Conf. File Storage Technol.*, 2021, pp. 387–401.

[29] K. Sriporn, C.-F. Tsai, C.-E. Tsai, and P. Wang, "Analyzing malaria disease using effective deep learning approach," *Diagnostics*, vol. 10, no. 10, p. 744, Sep. 2020.

[30] D. O. Oyewola, E. G. Dada, S. Misra, and R. Damasevicius, "A novel data augmentation convolutional neural network for detecting malaria parasite in blood smear images," *Appl. Artif. Intell.*, vol. 36, no. 1, pp. 1–22, Dec. 2022.

[31] S. Kadry, V. Rajinikanth, D. Taniar, R. Damasevicius, and X. P. B. Valencia, "Automated segmentation of leukocyte from hematological images—A study using various CNN schemes," *J. Supercomput.*, vol. 78, no. 5, pp. 6974–6994, Apr. 2022.

[32] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," *Comput. Methods Programs Biomed.*, vol. 161, pp. 1–13, Jul. 2018.

[33] S. Lee and Y. H. Lee, "Improving emergency department efficiency by patient scheduling using deep reinforcement learning," *Healthcare*, vol. 8, no. 2, p. 77, Mar. 2020.

[34] K. Weiss, T. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *J. Big Data*, vol. 3, pp. 1–40, May 2016.

[35] J. Kukačka, V. Golkov, and D. Cremers, "Regularization for deep learning: A taxonomy," 2017, *arXiv:1710.10686*.

[36] M. Yebasse, B. Shimelis, H. Warku, J. Ko, and K. J. Cheoi, "Coffee disease visualization and classification," *Plants*, vol. 10, no. 6, p. 1257, Jun. 2021.

[37] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[38] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[39] S. Rajaraman, S. K. Antani, M. Poostchi, K. Silamut, M. A. Hossain, R. J. Maude, S. Jaeger, and G. R. Thoma, "Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images," *PeerJ*, vol. 6, p. e4568, Apr. 2018.

[40] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, Jan. 2014.

MILKISA YEBASSE received the B.S. degree from Hawassa University and the master's degree in computer engineering from the Kumoh National Institute of Technology. His research interests include computer vision, deep learning, and image processing.

KYUNG JOO CHEOI received the B.S. degree in computer science from Chungbuk National University, Cheongju, South Korea, in 1996, and the M.S. and Ph.D. degrees in computer science from Yonsei University, Seoul, South Korea, in 2002.

From 2002 to 2005, she worked as a Research Engineer with TI-Specialist-Technology of LG CNS, Seoul. Since 2005, she has been a Professor with the Computer Science Department, Chungbuk National University, South Korea. Her research interests include computer vision, image processing, and machine learning.

JAEPIL KO received the B.S., M.S., and Ph.D. degrees from Yonsei University, South Korea, in 1996, 1998, and 2004, respectively.

Since 2004, he has been a Professor with the Computer Engineering Department, Kumoh National Institute of Technology, South Korea. His research interests include computer vision, machine learning, and image processing.

• • •