

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

A Novel Stacked CNN for Malarial Parasite Detection in Thin Blood Smear Images

Muhammad Umer¹, Saima Sadiq¹, Muhammad Ahmad^{2, 3,*}, Saleem Ullah¹, Gyu Sang Choi^{4,*}, and Arif Mehmood^{5,*}

¹Department of Computer Science, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan, Pakistan

²Department of Computer Engineering, Khwaja Fareed University of Engineering and Information Technology (KFUEIT), Rahim Yar Khan, Pakistan.

³Dipartimento di Matematica e Informatica - MIFT, University of Messina, Messina, Italy

⁴Department of Information and Communication Engineering, Yeungnam University, Gyeongsan 38542, Korea

⁵Department of Computer Science & Information Technology, The Islamia University of Bahawalpur, Bahawalpur, 63100, Pakistan

Corresponding author: Gyu Sang Choi (castchoi@ynu.ac.kr), Arif Mehmood (arifnhmp@gmail.com) and Muhammad Ahmad (mahmad00@gmail.com)

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1A2C1006159), MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2020-2016-0-00313) supervised by the IITP(Institute for Information & communications Technology Promotion), The Brain Korea 21 Plus Program(No. 22A20130012814) funded by the National Research Foundation of Korea (NRF), and in part by the Fareed Computing Research Center, Department of Computer Science under Khwaja Fareed University of Engineering and Information Technology(KFUEIT), Punjab, Rahim Yar Khan, Pakistan.

ABSTRACT Malaria refers to a contagious mosquito-borne disease caused by parasite genus plasmodium transmitted by mosquito female Anopheles. As infected mosquito bites a person, the parasite multiplies in the host's liver and start destroying the red-cells. The disease is examined visually under the microscope for infected red-cells. This diagnosis depends upon the expertise and experience of pathologists and reports may vary in different laboratories doing a manual examination. Another way around, many machine learning techniques have been applied for spontaneous detection of blood smears. However, feature engineering is a challenging task that requires expertise to adjust positional and morphological features. Therefore, this study proposes a novel Stacked Convolutional Neural Network architecture that improves the automatic detection of malaria without considering the hand-crafted features. The 5-fold cross-validation process on 27, 558 cell images with equal instances of parasitized and uninfected cells on a publicly available dataset from the National Institute of health, the accuracy of our proposed model is 99.98%. Furthermore, the statistical results revealed that the proposed model is superior to the state-of-the-art models with 100% precision, 99.9% recall, and 99% f1-measure.

INDEX TERMS Convolutional Neural Network (CNN), Malaria, Blood smear images, Deep learning, Diagnostic approach

I. INTRODUCTION

MALARIA is an infectious and life-threatening disease caused by protozoa plasmodium with a minimum of seven days of the incubation period. This disease is transmitted through the bite of female mosquito Anopheles that also known as malaria vectors. Among 400 species of Anopheles mosquitos, only 30 species are malaria vectors. *P. falciparum* and *P. vivax* are the most common single-cell Plasmodium species that cause malaria and can be toxic. Initial symptoms, headache, vomiting, fever, and chills can be mild and difficult to recognize as malaria if remain untreated can cause severe

illness and may lead to death [1]. According to the World Malaria Report in 2018, a number of malarial deaths i.e., 435000 were reported in 2018 [2].

Malarial virus transmission depends on climate conditions especially after rain and it is more intense when the temperature became feasible for a longer span of life of a mosquito. This is the reason for 90% world's malaria cases occur in Africa and common in other tropical regions such as Latin America and Asia [3]. Early detection can prevent harmful consequences and a patient can be treated with proper medicine on time.

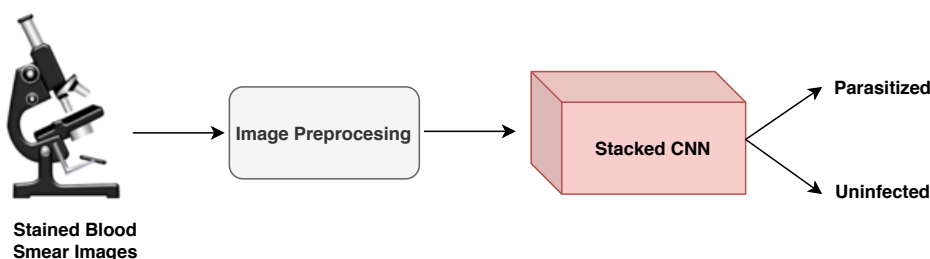


FIGURE 1: Pipeline of our proposed architecture.

To identify the malarial parasite, numerous techniques have been proposed and microscopic examination of Giemsa stain blood smear is manifest [4]. Other techniques include polymerase chain reaction and rapid diagnostic tests to detect the antigen in the blood. Although other tests outperform in malaria detection, however, microscopy is widespread due to low cost and less complexity and its efficacy depends upon pathologist expertise [5]. False diagnosis may lead to more severe malaria or the use of un-necessary malarial drugs [6]. To improve the treatment of an individual patient by automatic detection of malarial parasites is a very appealing area of research. It has two advantages; first, it will improve diagnosis even with limited resources, and secondly, it is cost-effective.

Automatic parasite detection from thin blood smear under microscope results to differentiate parasite species. The first step is to segment Red Blood Cells (RBCs) and then these segments are classified as infected or uninfected [7]. However, by applying machine learning on the medical image analysis task, feature engineering is challenging to get desired results because hand-crafted features are being used to make decisions [7], [8]. Furthermore, experienced individuals are required to adjust the size, angle, position, and region of interest (ROI) of the image. To cope with these issues, Deep Learning (DL) is being used to extract high-level features that result in end-to-end extraction of features and classification [9], [10].

For traditional image classification and analysis, the spatial correlation of neighboring pixels contains important information [11]. Convolutional Neural Network (CNN) is designed to extract such information i.e., end to end feature extraction and classification through weights and pooling [10]. However, the size of training data greatly affects the classification performance of CNN [12] as opposed to the traditional machine learning models [13]. To cope with the aforementioned issues, transfer learning has been proposed in which the features are extracted by a pre-trained network including but not limited to GoogleLeNet [14], VGGNet [15], and ResNet [16]. Transfer learning has been used as a shortcut where training time is saved by compromising performance [17].

The DenseNet architecture is a variant of CNN which is composed of dense layers in which each layer is fully connected to the later one and each layer serves as a feature

extractor [18]. DenseNet significantly improves performance for medical images without considering a large number of parameters [19]. Therefore, this work proposed a stacked CNN architecture that learns a different level of abstraction of a complex representation of malaria parasites for the classification of parasitized and uninfected cells for disease screening. A pipeline of our proposed architecture can be seen in Fig.1.

The rest of the paper is structured as follows. Section 2 describes the most relevant and state-of-the-art researches related to our proposed work. Section 3 gives a summary of the dataset, preprocessing, and related steps performed on the dataset. Section 4, presents a brief explanation of the deep learning model proposed in this work, experiment details, and machine specifications used for the experiment. In section 5 results are discussed and finally, section 6 concludes the work with possible future directions.

II. RELATED WORK

For automatic blood smear classification, traditional machine learning approaches have been used such as Diaz et al., classified blood smear images using a Support Vector Machine (SVM) to detect infected erythrocytes and their infected stage. This approach performed well with 94.0% sensitivity on a dataset containing 450 images [20]. In the structural characterization of blood cells, machine learning plays a vital role such as computer-aided learning techniques for pattern recognition that have been used by many researchers to identify the malaria parasitemia. Das et al., [21] extract features from erythrocytes textures and then apply feature selection techniques to further reduce it to 96 features and then applied statistical techniques such as Bayesian network and SVM for classification. The highest accuracy of 84.0% is achieved by a Bayesian network with the 19 most important features. Shen et al. [22] used a stacked autoencoder to learn features automatically from infected and uninfected images of cells.

Another way around, computer vision-based malarial parasite detection studies have also been proposed in the literature such as Tek et al. used a modified K-nearest neighbor (KNN) after applying normalization and color correction on input images of 9 blood films for binary classification [23]. Automatic detection and staging of infected RBCs by malarial parasite *P. falciparum* by using a quantitative phase analysis

TABLE 1: Summary table: Studies of malarial parasite detection.

Author	Dataset	Species	Staining	Blood Smear type	Features	Method	Performance
Diaz [20]	450 images	P. Falciparum	Giemsa	Thin	25 features	SVM	Sensitivity 94% Specificity 99.7%
Das [21]	Leishman stained peripheral blood smear images	P. falciparum & P. vivax	Leishman	Thin	Textual and Morphological	Naïve Bayes and SVM	Sensitivity 96.62 – 98.10%, Specificity 68.91 – 88.51%
Tek [23]	9 blood films	P. falciparum	Giemsa	Thin	color and scales	KNN	Sensitivity 74% Specificity 98%
Park [24]	Quantitative phase images of unstained cells	P. falciparum	Giemsa	Thin	23 Morphological descriptors based on the phase information	LDC, KNN, LR	LDC 99.7%, KNN 99.5% LR 99.7%
Mustafa [25]	30 Malaria cell images	P. falciparum	Giemsa	Thin	Thresholding	Bradley Fuzzy, Fuzzy C-Mean, Wolf Deghost, Bernsen Triangle	Bradley 78.43%, C-Mean 85.31%, Wolf 70.83%, Deghost 81.70%, Bernsen 65.53%, Triangle 78.12%
Liang [26]	27,578 Erythrocyte images	P. falciparum	Giemsa	Thin	NA	CNN	Sensitivity 96.99%, Specificity 97.75%
Bibin [27]	4100 Peripheral blood smear images	P. falciparum	Giemsa	Thin	Concatenated feature of color and texture	DBN	Sensitivity 97.60%, Specificity 95.92%
Gopakumar [28]	Focus stack of images	P. falciparum	Giemsa	Thin	14 features	SVM, CNN	Sensitivity 92.95 – 97.06%, Specificity 93.82 – 98.50
Mehanian [29]	Malaria blood film library	P. falciparum	Giemsa	Thin	N/A	CNN	Sensitivity 91.6%, Specificity 94.1%
Rajaraman [30]	27,558 cell images	P. falciparum	Giemsa	Thin	N/A	Pre-trained CNN	Sensitivity 0.981%, Specificity 0.992%
Rahman [31]	27,558 cell images	P. falciparum	Giemsa	Thin	N/A	CNN, VGG16 CNNEx-SVM	95.97%, VGG16 97.64%, CNNEx-SVM 94.77%

of images without staining is performed by [24]. Mustafa et al. [25] compared their work with Bradley's [32], wolf's [33], Bernsen's [34], Deghosting [35], Triangle's [36] and Fuzzy C-mean clustering method [37] for malarial parasite detection. Mustafa et al. also proposed thresholding as an important preprocessing step. Fuzzy C-mean outperformed among all the other five methods. Although reported outcomes using machine learning are reasonable, all the techniques need to prove their ability on large datasets as all these approaches have been evaluated on small sets of images. Therefore, there is a need for a deep learning approach such as Convolutional Neural Network (CNN), which has proved robustness on large datasets.

Deep neural networks such as Generative Adversarial Network (GAN) [38], Discrete-time Recurrent Neural Networks

(DRNNs) [39] and Memristive Neural Networks (MNNs) [40] have been widely used for various tasks. CNN models has been extensively used for traditional image analysis, phoneme recognition [41], document recognition [42], visual document analysis [43], face labeling [44] and object recognition [45], [46]. CNN model-based on 16 layers of malarial parasite detection was proposed in [26] which only classify the blood cells as infected or uninfected. The model was trained with approximately 27000 images and achieved 97.0% accuracy, specificity and sensitivity which is higher than transfer learning. To compensate limited resources images were resampled to the size of 44×44 pixels. The first application based on a deep belief network was proposed by Bibin et al. [27] and tested on 4100 peripheral blood smear images which achieved 89.66% F1-score.

A customized CNN architecture for the detection of plasmodium in blood smear on Leishman stained focused stacked images was proposed by Gopakumar et al., which show 97.0% sensitivity and 98.0% specificity [28]. Automatic identification of the malarial parasite was proposed in [29] which uses the patient level evaluation and thumbnails to improve the user confidence in system findings with overall 89.7% precision, 94.1% specificity and 89.7% sensitivity. Rajaraman et al., [30] evaluated a pre-trained end-to-end (i.e., feature extraction and classification) CNN model based on single-cell images. They observed that a pre-trained ResNet-50 model as an outstanding tool for diagnosis with an accuracy of 98.6%, 98.1% sensitivity, 99.2% specificity, and 95.7% F1-score. The detailed summary of existing works is shown in Table 1.

Even though the existing state-of-the-art deep learning approaches have shown promising results in malarial parasite detection but still there is room for improvement. Sometimes uninfected blood samples do not contain plasmodium but may contain other types of remnants that are wrongly classified as infected by a classifier. Therefore, color normalization techniques are needed before the classification. In this work, we evaluated the customized CNN model for feature extraction and then to classify images as infected or uninfected cells. Rajaraman et al. use 3 layers of CNN (in our case it is 5), the second thing in all three layers they applied same filter size with the same number of kernels (3×3 , @32) while in our case we vary the kernel size (4×4 , 3×3 , 2×2) with kernels ranging from 32 to 256. Sequentially reducing kernel size helps the model to get trained on small size malarial cell detection.

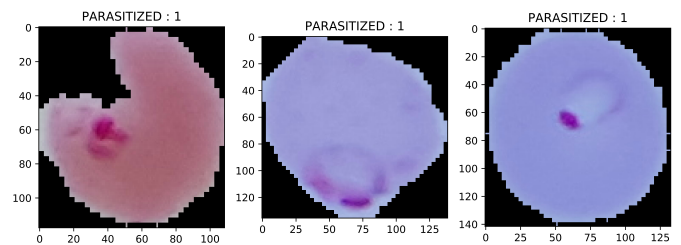
III. DATASET & PREPROCESSING

A. MALARIAL DATASET

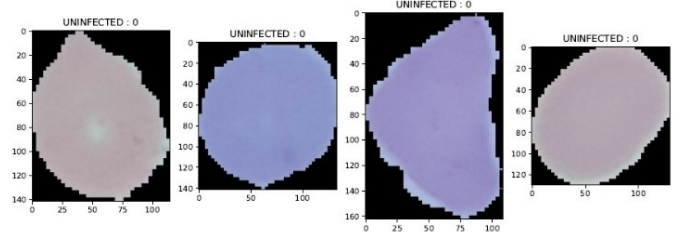
Dataset used in our study contains images based on Giemsa stained slides of thin blood smear obtained from malaria screener research activity of 50 healthy patients and 150 *P. falciparum*-infected patients. It is taken from the National Institute of Health (NIH). Images in a dataset are manually annotated by slide reader experts of Mahidol Oxford Tropical Medicine Research Unit Bangkok, Thailand, and collected at the National Library of Medicine (NLM). The dataset contains 27,558 images with the equal occurrence of infected and uninfected red blood cell images as shown in Table 2. Infected blood cell image samples contain plasmodium as shown in Fig. 2(a) and uninfected blood image samples do not contain plasmodium as shown in Fig. 2(b). Colored patches of red blood cells are of variant sizes (110 – 150 pixels), which are resampled to 120×120 according to the input requirement of classifier during preprocessing.

TABLE 2: Malaria dataset description

Total images	Parasitized images	Uninfected images
27,558	13,779	13,779



(a) Parasitized Images



(b) Uninfected Images

FIGURE 2: Microscopic view of thin blood smear

B. PREPROCESSING

The original images of the malaria dataset are captured by a mobile device therefore these images are in different sizes. Thus, before any training and testing, we resampled the images to unified image size. On the first step of preprocessing we convert all images to fix the size of 120×120 pixels. Secondly, we apply the kernel on the image to get the edges. On the third step of preprocessing we convert BGR to YUV to get the values of one luma component (Y') and two chrominance components, called U (blue projection) and V (red projection). Color variations in blood cell images exist due to the use of chemicals which can result in error margin. This problem can be solved by normalizing the image. Ciompi et al. [47] applied stain normalization on colorectal tissue classification and proved that it improves the performance. We also applied normalization in the fourth step of preprocessing to equalize the intensity values. The last step of preprocessing is to convert back the YUV image into RGB. Preprocessing steps are important to reduce noise and to improve image quality. Fig. 3(a) represents the original image before preprocessing, Fig. 3(b) shows edges obtained after applying kernel. Fig. 3(c) displays images in YUV color space to get (Y') component Fig. 3(d) shows intensity equalization and Fig. 3(e) represents images after converting back to BGR color space.

C. DATA SPLITTING AND CROSS VALIDATION

Malaria dataset is split into train/test with a ratio of 70 : 30 and to check the robustness of the model we applied a 5-fold cross-validation that is a moderate value which neither causes high bias or high variance. We randomly partitioned 5 equal subsets of our dataset; one set is used as validation and rest are used to train the model. This process is repeated five times

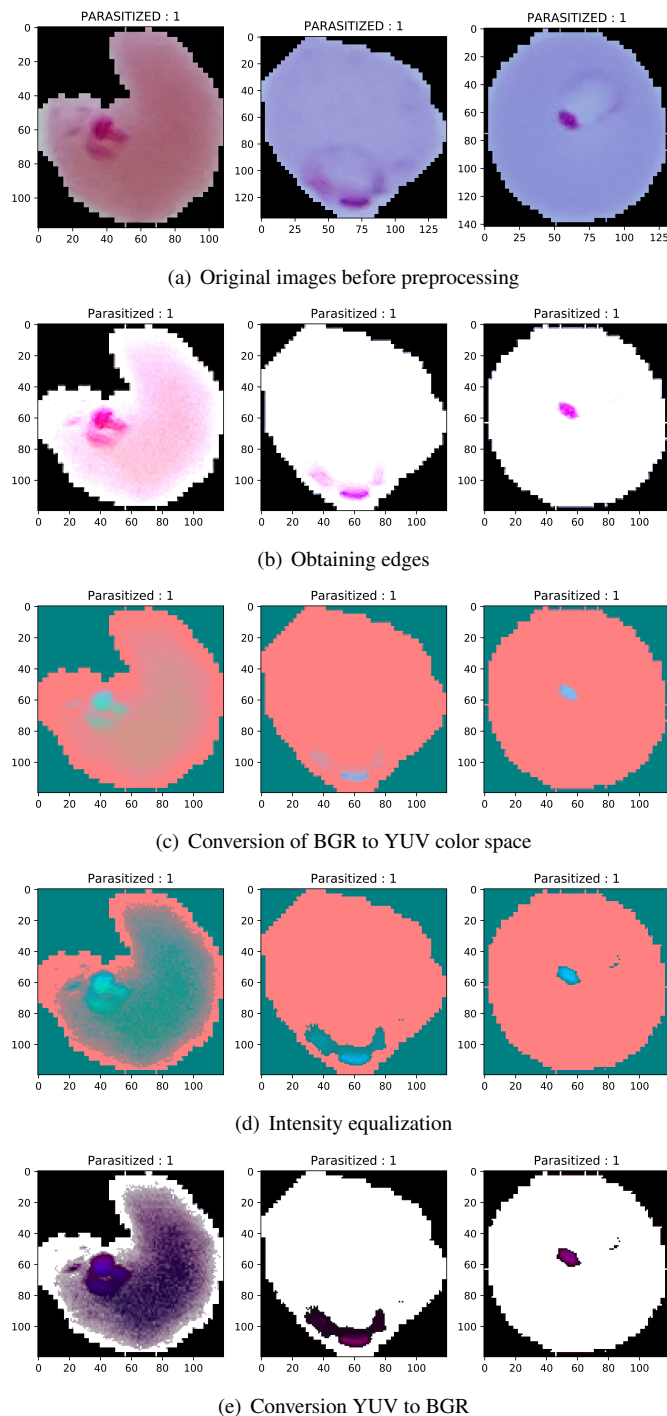


FIGURE 3: Preprocessing Steps

with each subset. Then all these five subsets are averaged and used for model evaluation.

IV. PROPOSED METHODOLOGY

A. OVERVIEW

We are using a stacked CNN to overcome the shortcomings of manual feature extraction. We apply re-sampling to extract more information to CNN with a fixed sampling pattern. In

addition, we applied stain normalization to preserve image characteristics.

The pipeline of the proposed approach composed of the following steps. First, we apply pre-processing steps to input images by re-sampling and normalizing images. Then we apply stacked CNN by fine-tuning it(filters, kernels, and strides) along with max-pooling and dropout layer. As we experimented to test different design strategies, We also created and tested models having 1,3, and 5 convolution layers, pooling layers, and a dense classification layer. We progressively increased the number of convolution layers, dropout layers, and pooling layers to see an increase in the performance of the model. We used these 1 and 3-layer models as a baseline and compare our results with these models. We used varied filter sizes and the number of layers until we achieve the best result. We could not find any other CNN architecture which showed improved performance than our proposed model. Since our 5 convolution layer stacked CNN model is the best deep learning technique for Malarial parasite classification.

B. CNN

CNN is a type of deep neural network that learns a complex hierarchy of features by convolution, nonlinear activation, and pooling layers [12]. CNN is designed for image recognition tasks as well as for image classification. Now it is commonly used in image segmentation. Traditional approach sliding window process regions independently which results in low efficiency. An alternate method is fully CNN that is trained in the end to end fashion by making computation more efficient. Fully connected layers are used at the end of the network for semantic information encoding. We have used stacked CNN as shown in Fig. 4 for detecting parasites in infected cell RGB images. CNN is a multi-layered feed-forward network inspired by biology. Filters or kernels in a layer are applied to the input of the first layer or the output of the previous layers and result in a feature map. The output of all convolutional layers is concatenated as a feature map and fed into fully connected layers. CNN has been proved as a de-facto standard by providing robust results in medical domain classification tasks. CNN has been applied for the classification of lung disease [48], brain tumor segmentation [49], chest x-rays [50], chest radiographs [51] and kidney disease [52]. In recent studies CNN has been explored for malarial parasite image classification of Giemsa stained images as parasitized or uninfected in [27], [28], [30].

The main components of CNN are convolutional layers, Rectified Linear Unit (ReLU), and pooling layer or subsampling layer. Features are extracted by the convolutional layer, ReLU is easy: it converts any negative elements of the matrix to 0 and keep the others positive constant. For the activation function, we applied the Rectified Linear Unit (ReLU).

$$y = \max(0, i) \quad (1)$$

where y is the output activation and i is the given input.

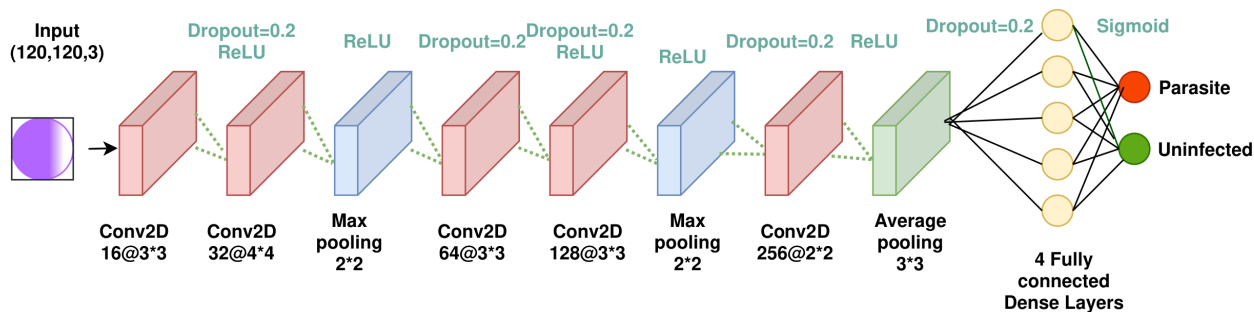


FIGURE 4: Proposed Stacked CNN architecture

During training kernel weights are applied on the input image to extract local features at convolution and subsequent layers extract high-level features from these local features. In multichannel images, CNN opens up ways in malaria diagnosis. The cross-entropy error is used as a loss function as it is used for binary classification. It is calculated as shown in equation 2.

$$\text{crossEntropy} = -(i \log(p) + (1 - i) \log(1 - p)) \quad (2)$$

where i is the binary indicator of class labels (0 or 1), a \log is a natural logarithm and p is the predicted probability. CNN is a backpropagation variant algorithm and therefore we used sigmoid output as the error function. Here N is the total number of classes in the sigmoid layer and one neuron corresponds to each class in the output layer. In our case, the number of classes is two parasitized and uninfected. CNN architecture produces output at two neurons in every case of binary classification. For an ideal case of the parasitized cell, the output will be 1 and 0 of first and second neurons respectively. For uninfected images, the output will be 0 and 1 that is the reverse of previous output. The term inside the log function computes the chance of output 1 and i_j is the output for a true class that is c for input. In our case, c is 1 for parasitized and 2 for uninfected. At testing time labels are assigned by excluding the softmax loss to maximize the response.

C. EXPERIMENTAL DETAILS

The design of stacked CNN architecture used in our experiment is shown in Fig. 4. Our proposed stacked architecture has been designed of total 22 layers, 5 convolutional layers, 2 max-pooling layers, 4 dense layers, 1 average pooling layer, 1 flatten layer, 8 layers with 20% dropout and 1 fully connected layer as shown in Fig. 4. The Rectified Linear Unit (ReLU) activation function is used in this setup.

Here, an input image of size 120×120 is resampled from 200×200 pixels which are enough to hold neighborhood details for making a final decision. In conv2D layer filter size (2×2) , (3×3) and (4×4) are applied to convolve. The kernel size used at every convolutional layer is shown in the subscript in Fig. 4. In the MaxPooling2D layer, 2×2 pool

size is used and in average pooling layer (3×3) pool size is used. Every convolutional layer is followed by a dropout layer which discards 20% of neurons. The output of the final Conv2D layer of 256 output neurons is followed by an average pooling layer. This is followed by 4 dense layers of input to other activation functions. We prefer max-pooling layers before the average pooling layer as we did not want to average the details at an early stage. As it is designed for the binary classification problem, cross-entropy function is used to calculate the error between predicted actual and predicted output. Due to the binary classification output is set to 2. After setting input and output reasonable CNN can do fair classification. We choose the appropriate deep CNN architecture not too deep nor too shallow having 5 convolutional layers in our task. However, we applied max pooling to deal with the nonlinearity of features. Adam optimizer is used to remove biases. Bias is set to 0 and random weights are randomly initialized. Batch size is set to 32 samples and continued for 13 epochs. We applied a shallow to deep CNN model with a 1 convolutional layer to 5 convolutional layers and used them as a baseline method as shown in Table 4. The complete summary of the stacked CNN model hyper-parameter values is shown in table 3.

TABLE 3: Summary of the model and hyper-parameter values

Parameter	Value
Input dimension	(120, 120, 3)
Batch size	32
Pooling	2×2
Epochs	13
Optimizer	Adam
Function	Binary cross entropy

V. RESULTS & DISCUSSIONS

All the experiments are carried out on a 2GB Dell PowerEdge T430 graphical processing unit on 2x Intel Xeon 8 Cores 2.4Ghz machine which is equipped with 32 GB DDR4 Random Access Memory (RAM). The training takes 3.5 hours to give the final result on the 'Malarial dataset'.

We evaluated our proposed architecture on Accuracy, Precision, Recall, and F1-score. The results of all these metrics are shown in Table 4. From Table 4, one can conclude that

our proposed architecture with 5 fold cross-validation outperformed with 99.964% accuracy, 100.0% precision, 99.928% recall, and 99.964% F1- score. Table 4 enlist some of the results achieved with a different number of convolutional layers are presented. Stacked CNN-1, stacked CNN-3 and Stacked CNN-5 show accuracy 50.145%, 61.412% and 99.879% respectively. It has been observed that as the number of CNN layers is increasing till 5-layers, the accuracy, precision, and F1-score also increases. The results obtained by CNN-5 layers with 5-fold cross-validation are the best among all. Detail of filter size and parameters used for tuning can be seen in Table 6. Graphical representation of the result can be seen in Figure 6. Figure 5 shows training accuracy as well as loss, precision, and recall for our stacked CNN architecture. The model is trained for 13 epochs with accuracy > 99%.

TABLE 4: Performance of Proposed Stacked CNN architecture's

Model	Accuracy	Precision	Recall	F1-score
Stacked CNN-1 layer	50.145	50.145	100.00	66.795
Stacked CNN-3 layers	61.412	77.175	100.00	87.130
Stacked CNN-5 layers	99.879	99.976	99.783	99.879
Stacked CNN-5 with 5-fold	99.964	100.00	99.928	99.964

We analyzed that it is difficult for a classifier to differentiate stains from plasmodium or any other artifact in blood. This is the reason we applied stain normalization in preprocessing steps. We reduced noise in preprocessing to improve the image quality. Results proved that these activities improved the overall result. Our proposed stacked CNN 5-layered model results are compared with preprocessing steps and without preprocessing steps in Table 5. There is a drastic difference in terms of accuracy, precision, recall, and F1- score after performing preprocessing steps.

TABLE 5: Performance of Proposed Stacked CNN with and without Preprocessing

Model	Accuracy	Precision	Recall	F1-score
Stacked CNN-5 layers with preprocessing	99.879	99.976	99.783	99.879
Stacked CNN-5 layers without preprocessing	49.61	50.14	50.14	50.14

TABLE 6: Layers structure of CNN model used in this work

CNN-1 layer	CNN-3 layer	CNN-5 layer	CNN-5 layer with 5-fold
Conv (3 × 3, 8/16), Strides(1 × 1)	Conv (3 × 3, 8/16), Strides(1 × 1)	Conv (3 × 3, 8/16), Strides(1 × 1)	Conv (3 × 3, 8/16), Strides(1 × 1)
Average Pooling (5 × 5, Strides(1 × 1))	Conv (4 × 4, 8/32), Strides(2 × 2)	Conv (4 × 4, 8/32), Strides(2 × 2)	Conv (4 × 4, 8/32), Strides(2 × 2)
Dense (120 neurons)	Max Pooling (2 × 2), Strides(2 × 2)	Max Pooling (2 × 2), Strides(2 × 2)	Max Pooling (2 × 2), Strides(2 × 2)
Dense (80 neurons)	Conv (3 × 3, 8/64), Strides(1 × 1)	Conv (3 × 3, 8/64), Strides(2 × 2)	Conv (3 × 3, 8/64), Strides(2 × 2)
Dense (10 neurons)	Average Pooling (5 × 5, Strides(1 × 1))	Conv (2 × 2, 8/128), Strides(1 × 1)	Conv (2 × 2, 8/128), Strides(1 × 1)
Sigmoid (2-class)	Dense (120 neurons)	Max Pooling (2 × 2), Strides(2 × 2)	Max Pooling (2 × 2), Strides(2 × 2)
	Dense (80 neurons)	Conv (2 × 2, 8/256), Strides(2 × 2)	Conv (2 × 2, 8/256), Strides(2 × 2)
	Dense (10 neurons)	Average Pooling (5 × 5, Strides(1 × 1))	Average Pooling (5 × 5, Strides(1 × 1))
	Sigmoid (2-class)	Dense (120 neurons)	Dense (120 neurons)
		Dense (80 neurons)	Dense (80 neurons)
		Dense (10 neurons)	Dense (10 neurons)
		Sigmoid (2-class)	Sigmoid (2-class)

Our stacked CNN model achieved optimal metric values by using five-phase extensive pre-processing, hyper-parameter optimization, different filter sizes, and dropout layers. In literature, many studies extracted features by using pre-trained CNN before classification [53] and others used customized CNN [30]. The stacked approach outperformed in the classification of parasitized and uninfected blood image cells. In our case, we identified optimal layers of CNN for feature extraction before classification for malarial parasite

detection. It accurately identifies infected cells in terms of accuracy, precision, recall, and F1-score.

We have further compared the results of the proposed stacked CNN model with state-of-the-art Deep learning models proposed in the literature. Deep CNN Models [26], [30] and [31] was chosen as baseline methods for comparison with proposed stacked CNN as these models have recently achieved best results for malarial parasite detection. The models selected for comparison purposes are tested on a dataset based on Giemsa stained thin blood smear image slides. In [26] researchers proposed a 16 layered CNN architecture containing 6 convolutional layers and obtained good accuracy of 97.37% and 97.36% F1-score. The setting used in said work is as follows; filter sizes (5 × 5, 4 × 4, 3 × 3) are used and (5 × 5) filter size is used in 4 convolutional layers out of 6 layers. It can be observed that shallow models give reasonable accuracy with a large filter size. Hence filter size plays an important role in CNN architecture. We used a small filter size and reduced it in a sequential way form (4 × 4) to (2 × 2) which helps the model to train on small spots on infected parasitized blood cell images.

The ResNet model was proposed in [30] which outperformed among all five used pre-trained deep learning models with 95.70% accuracy and their 3-layered customized CNN model achieved 94.00% accuracy. Rajaraman et. al [30] also claimed that they need to use color normalization techniques to improve accuracy. The proposed model shows good accuracy improvement over the baseline [30] by applying color normalization at the pre-processing level. The comparison of the proposed model and other state-of-the-art models are shown in Table 7.

TABLE 7: Performance Comparison between CNN based DL models in literature on malaria dataset

Model	Accuracy	Precision	Recall	F1-score
Customized [26]	97.37	96.99	97.75	97.36
Customized [30]	94.00	95.10	93.10	94.10
ResNET-50 [30]	95.70	96.90	94.50	95.70
Customized [31]	96.29	98.04	92.34	94.95
TL-VGG16 [31]	97.77	97.19	97.20	97.09
Proposed	99.96	100	99.92	99.96

Another CNN-based model [31] applied very deep CNN architecture TL-VGG16 with 16 convolutional layers and customized with 8 convolutional layers for malarial parasite detection. This model [31] achieved 96.29% by customized approach and 97.77% by TL-VGG16. Very deep architecture requires more time for training as TL-VGG16 [31] converges at 80 epochs.

To get rid of bias and to reduce overfitting we applied 5-fold cross-validation toward optimal development of stacked CNN architecture. Results of 5-fold cross-validation is represented in Table 4 with 99.964% Accuracy, 100% Precision, 99.928% Recall and 99.964% F1-score. We present results of all 5 folds in Table 8 and also compares estimated accuracy after performing 5-fold cross-validation with other best performing models in the literature as shown in Table 9.

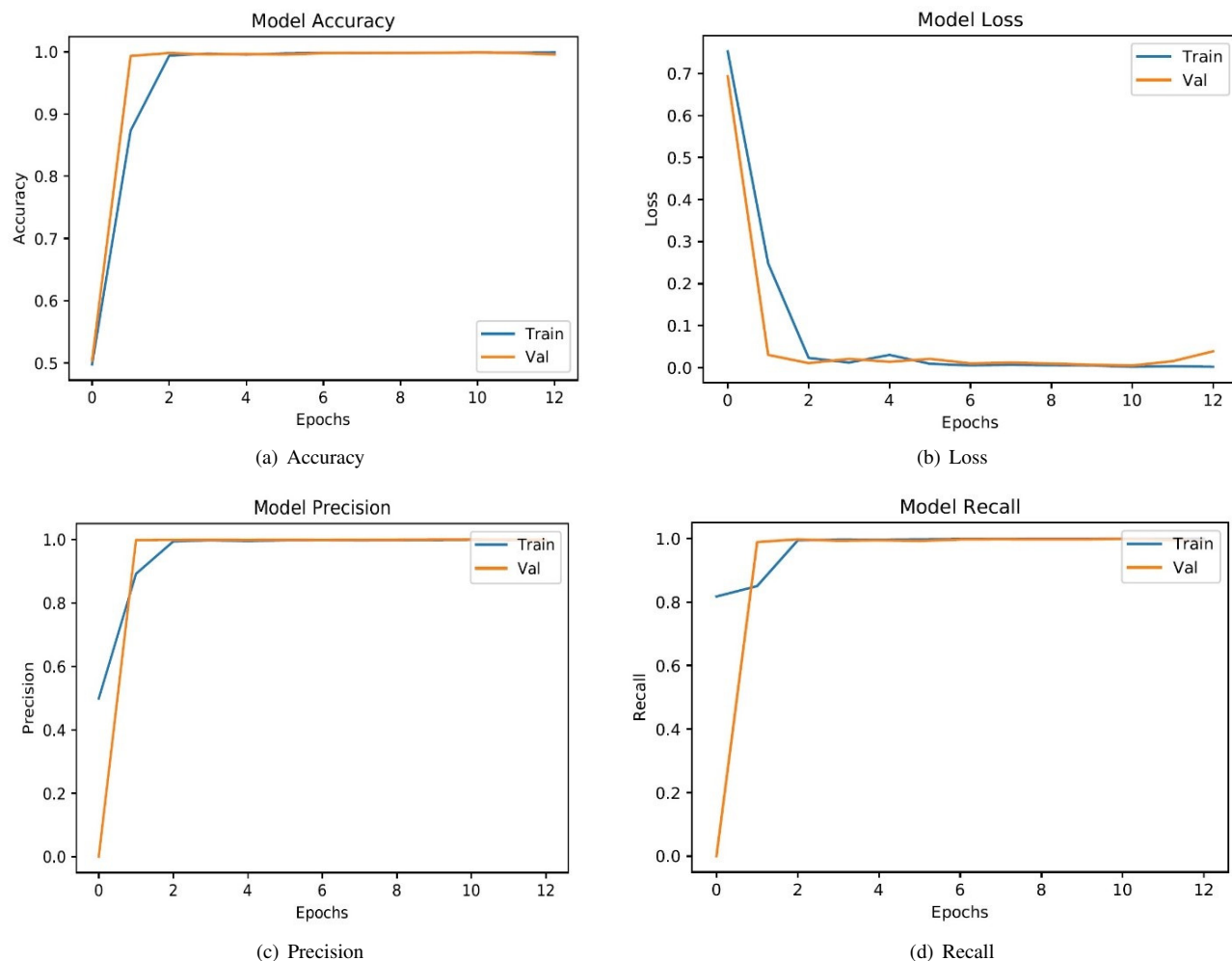


FIGURE 5: Proposed model experimental result evaluation.

Evidently in each fold, our model does not show any variance thus ensure robustness and generality.

TABLE 8: 5-fold cross validation results

<i>K</i> -Folds	Accuracy	Precision	Recall
Fold-1	97.927	100	99.856
Fold-2	99.915	99.976	99.855
Fold-3	99.964	99.976	99.951
Fold-4	99.867	99.734	100
Fold-5	99.988	100	99.976
5-Fold Mean	99.964	100	99.928
Standard Deviation	0.001	0.001	0.001

TABLE 9: Performance Evaluation with best performing models in literature

Model	Estimated Accuracy
ResNET-50 [30]	95.7±0.007
TL-VGG16 [31]	97.0±0.005
Proposed	99.96±0.001

A. IMPORTANCE OF STAIN NORMALIZATION

We also evaluate the performance of our stacked CNN architecture by training it to the blood smear images without applying stain normalization. Applying the CNN model directly to the dataset images gave a poor accuracy value of 49.61% accuracy. It is evident from Table 5 stain normalization remarkably improved the performance of our proposed model by 50% and reaching to 99.96%. Based on the results presented in Table 5 we investigate the importance of stain normalization in our classification task and we found it important step to include in training and evaluation of the proposed model.

VI. CONCLUSION

Traditional machine learning methods have shown limited accuracy for malarial parasite detection. Therefore, this work proposed a stacked CNN model-based on an end-to-end artificial neural network to improve malarial classification from thin blood smear images. The achieved results prove that with varying filter sizes and depth, convolutional layers can extract different abstract level features for classification. This

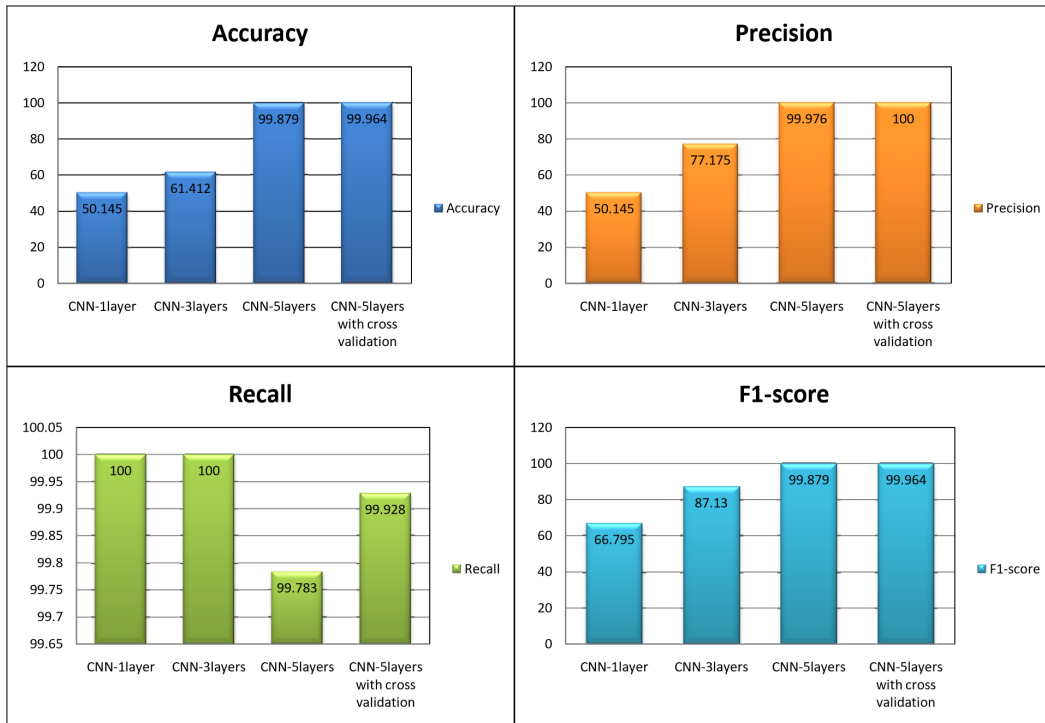


FIGURE 6: Proposed Stacked CNN architecture performance with different number of layers

study proves that features extracted by CNN are better than hand-crafted features. The stacked CNN model using stain normalization outperformed than state-of-the-art deep learning methods. Experimental results of 5-fold cross-validation confirm the superiority of the proposed Stacked CNN model with 99.96% accuracy. Our future direction entails further refine it to improve the classification accuracy within-subject or cross-subject.

REFERENCES

- [1] Hector Caraballo and Kevin King. Emergency department management of mosquito-borne illness: malaria, dengue, and west nile virus. *Emergency medicine practice*, 16:1–23; quiz 23, 05 2014.
- [2] World Health Organization. World health organization. regional for the eastern mediterranean. <https://www.who.int/ith/diseases/malaria/en/>, 2014. Online: accessed 22 September 2019.
- [3] Haidong Wang, Mohsen Naghavi, Christine Allen, Ryan Barber, Zulfiqar Bhutta, Austin Carter, Daniel Casey, Fiona Charlson, Alan Chen, Matthew Coates, Megan Coggeshall, Lalit Dandona, Daniel Dicker, Holly Erskine, Alize Ferrari, Christina Fitzmaurice, Kyle Foreman, Mohammad Forouzanfar, Maya Fraser, and Christopher Murray. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015 - supplementary information 1, 10 2016.
- [4] Karan S. Makhija, Samuel Maloney, and Robert Norton. The utility of serial blood film testing for the diagnosis of malaria. *Pathology*, 47(1):68 – 70, 2015.
- [5] World Health Organization. Malaria microscopy quality assurance manual. <http://apps.who.int/medicinedocs/en/m/abstract/Js19135en/>, 2016. Online: accessed 22 September 2019.
- [6] Mahdiah Poostchi, Kamolrat Silamut, Richard Maude, Stefan Jaeger, and George Thoma. Image analysis and machine learning for detecting malaria. *Translational Research*, 194, 01 2018.
- [7] Kamolrat Silamut, Md Hossain, Ilker Ersoy, Richard Maude, Stefan Jaeger, George Thoma, and Sameer Antani. Understanding the learned behavior of customized convolutional neural networks toward malaria parasite detection in thin blood smear images. *Journal of Medical Imaging*, 5:1, 07 2018.
- [8] M. Ahmad, A. M. Khan, and R. Hussain. Graph-based spatial-spectral feature learning for hyperspectral image classification. *IET Image Processing*, 11(12):1310–1316, 2017.
- [9] Juergen Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61, 04 2014.
- [10] Muhammad Ahmad. A fast 3d cnn for hyperspectral image classification. 2020.
- [11] Muhammad Ahmad, Sidrah Shabbir, Diego Oliva, Manuel Mazzara, and Salvatore Distefano. Spatial-prior generalized fuzziness extreme learning machine autoencoder-based active learning for hyperspectral image classification. *Optik-International Journal for Light and Electron Optics*, 2020.
- [12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton. Imagenet classification with deep convolutional neural networks. *Neural Information Processing Systems*, 25, 01 2012.
- [13] Muhammad Ahmad, Asad Khan, Adil Mehmood Khan, Manuel Mazzara, Salvatore Distefano, Ahmed Sohaib, and Omar Nibouche. Spatial prior fuzziness pool-based interactive classification of hyperspectral images. *Remote Sensing*, 11(9), May. 2019.
- [14] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1–9, 06 2015.

- [15] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv 1409.1556, 09 2014.
- [16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 770–778, 06 2016.
- [17] Muhammad Ahmad, Mohammed A Alqarni, Adil Mehmood Khan, Rasheed Hussain, Manuel Mazzara, and Salvatore Distefano. Segmented and non-segmented stacked denoising autoencoder for hyperspectral band reduction. *Optik - International Journal for Light and Electron Optics*, In Press, Oct 2018.
- [18] Fouad Bousetouane and Brendan Morris. Off-the-shelf cnn features for fine-grained classification of vessels in a maritime environment. In *Advances in Visual Computing*, pages 379–388, 12 2015.
- [19] Kenji Suzuki. Overview of deep learning in medical imaging. *Radiological Physics and Technology*, 10, 07 2017.
- [20] Gloria Díaz, Fabio González, and Eduardo Romero. A semi-automatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images. *Journal of biomedical informatics*, 42:296–307, 02 2009.
- [21] Dev Kumar Das, Madhumala Ghosh, Mallika Pal, Asok Maiti, and Chandan Chakraborty. Machine learning approach for automated screening of malaria parasite using light microscopic images. *Micron (Oxford, England : 1993)*, 45, 11 2012.
- [22] Hongda Shen, W. Pan, Yuhang Dong, and Mohammad Alim. Lossless compression of curated erythrocyte images using deep autoencoders for malaria infection diagnosis. In 2016 Picture Coding Symposium (PCS), pages 1–5, 01 2016.
- [23] F. Tek, Andrew Dempster, and Izzet Kale. Parasite detection and identification for automated thin blood film malaria diagnosis. *Computer Vision and Image Understanding*, 114:21–32, 01 2010.
- [24] Han Park, Matthew Rinehart, Katelyn Walzer, Jen-Tsan Chi, and Adam Wax. Automated detection of *p. falciparum* using machine learning algorithms with quantitative phase images of unstained cells. *PLoS one*, 11:e0163045, 09 2016.
- [25] Wan Mustafa, Raganathan Santiago, Irfan Jamaluddin, Nurul Othman, Wan Khairunizam, and Mohamad Rohani. Comparison of detection method on malaria cell images. In 2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA), pages 1–6, 08 2018.
- [26] Zhaohui Liang, Andrew Powell, Ilker Ersoy, Mahdieh Poostchi, Kamolrat Silamut, K. Palaniappan, Peng Guo, Md Hossain, Sameer Antani, Richard Maude, Jimmy Huang, Stefan Jaeger, and George Thoma. Cnn-based image analysis for malaria diagnosis. In 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pages 493–496, 12 2016.
- [27] Dhanya Bibin, Madhu S. Nair, and P. Punitha. Malaria parasite detection from peripheral blood smear images using deep belief networks. *IEEE Access*, PP:1–1, 05 2017.
- [28] Gopakumar GI, Murali Swetha, Sai Siva Gorthi, and Gorthi Sai Subrahmanyam. Cnn based malaria diagnosis from focus-stack of blood smear images acquired using custom-built slide scanner. *Journal of Biophotonics*, 11:e201700003, 08 2017.
- [29] Couroush Mehanian, Mayoore Jaiswal, Charles Delahunty, Clay Thompson, Matt Horning, Liming Hu, Shawn McGuire, Travis Ostbye, Martha Mehanian, Ben Wilson, Cary Champlin, Earl Long, Stephane Proux, Dionicia Gamboa, Peter Chiodini, Jane Carter, Mehul Dhorda, David Isaboke, Bernhards Ogutu, and David Bell. Computer-automated malaria diagnosis and quantitation using convolutional neural networks. In 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), pages 116–125, 10 2017.
- [30] Sivaramakrishnan Rajaraman, Sameer Antani, Mahdieh Poostchi, Kamolrat Silamut, Md Hossain, Richard Maude, Stefan Jaeger, and George Thoma. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *PeerJ*, 6, 04 2018.
- [31] Aimon Rahman, Hasib Zunair, M Sohel Rahman, Jesia Quader Yuki, Sabyasachi Biswas, Md Ashraf Alam, Nabila Binte Alam, and M. R. C. Mahdy. Improving malaria parasite detection from red blood cell using deep convolutional neural networks. 2019.
- [32] Derek Bradley and Gerhard Roth. Adaptive thresholding using the integral image. *J. Graphics Tools*, 12:13–21, 01 2007.
- [33] Christian Wolf and Jean-Michel Jolion. Extraction and recognition of artificial text in multimedia documents. *Formal Pattern Analysis & Applications*, 6:309–326, 02 2004.
- [34] John Bernsen. Dynamic thresholding of gray-level images. In *Eighth International Conference on Pattern Recognition. Proceedings*, 1986.
- [35] Yiyang Li, Weiqi Jin, Jin Zhu, Zhang xu, and Shuo Li. An adaptive deghosting method in neural network-based infrared detectors nonuniformity correction. *Sensors (Basel, Switzerland)*, 18, 01 2018.
- [36] He Liang and Chai Zhenming. Automatic detection of sister chromatid exchange. *Journal of Histochemistry & Cytochemistry*, 01 1988.
- [37] Santiago Aja-Fernández, Ariel Curiale, and Gonzalo Vegas Sánchez-Ferrero. A local fuzzy thresholding methodology for multiregion image segmentation. *Knowledge-Based Systems*, 83, 04 2015.
- [38] Shiping Wen, Weiwei Liu, Yin Yang, Tingwen Huang, and Zhigang Zeng. Generating realistic videos from keyframes with concatenated gans. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(8):2337–2348, 2018.
- [39] Bo Sun, Yuting Cao, Zhenyuan Guo, Zheng Yan, and Shiping Wen. Synchronization of discrete-time recurrent neural networks with time-varying delays via quantized sliding mode control. *Applied Mathematics and Computation*, 375:125093, 2020.
- [40] Yuting Cao, Shengbo Wang, Zhenyuan Guo, Tingwen Huang, and Shiping Wen. Synchronization of memristive neural networks with leakage delay and parameters mismatch via event-triggered control. *Neural Networks*, 119:178–189, 2019.
- [41] Alexander H. Waibel, Toshiyuki Hanazawa, Geoffrey E. Hinton, Kiyohiro Shikano, and Kevin J. Lang. Phoneme recognition using time-delay neural networks. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 37:328–339, 1989.
- [42] Yann Lecun, Leon Bottou, Y. Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86:2278 – 2324, 12 1998.
- [43] Patrice Y. Simard, David Steinkraus, and John C. Platt. Best practices for convolutional neural networks applied to visual document analysis. *Seventh International Conference on Document Analysis and Recognition*, 2003. *Proceedings.*, pages 958–963, 2003.
- [44] Minghui Dong, Shiping Wen, Zhigang Zeng, Zheng Yan, and Tingwen Huang. Sparse fully convolutional network for face labeling. *Neurocomputing*, 331:465–472, 2019.
- [45] Christophe Monrocq and Yann Lecun. An original approach for the localization of objects in images. *IEE Proceedings - Vision, Image and Signal Processing*, 08 1998.
- [46] Yann LeCun, Y. Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521:436–44, 05 2015.
- [47] Francesco Ciompi, Oscar Geessink, Babak Ehteshami Bejnordi, Gabriel Silva De Souza, Alexi Baidoshvili, Geert Litjens, Bram Van Ginneken, Iris Nagtegaal, and Jeroen Van Der Laak. The importance of stain normalization in colorectal tissue classification with convolutional networks. In 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pages 160–163. *IEEE*, 2017.
- [48] Qing Li, Weidong Cai, Xiaogang Wang, Yun Zhou, David Dagan Feng Feng, and Mei Chen. Medical image classification with convolutional neural network. 2014 13th International Conference on Control Automation Robotics and Vision, ICARCV 2014, pages 844–848, 03 2015.
- [49] Sergio Pereira, Adriano Pinto, Victor Alves, and Carlos Silva. Brain tumor segmentation using convolutional neural networks in mri images. *IEEE Transactions on Medical Imaging*, 35:1–1, 03 2016.
- [50] Yaniv Bar, Idit Diamant, Lior Wolf, Sivan Lieberman, Eli Konen, and Heather Greenspan. Chest pathology detection using deep learning with non-medical training. *Proceedings - International Symposium on Biomedical Imaging*, 2015:294–297, 07 2015.
- [51] Vincent Liu, Mark Clark, Mark Mendoza, Ramin Saket, Marla Gardner, and Benjamin Turk. Automated identification of pneumonia in chest radiograph reports in critically ill patients. *BMC medical informatics and decision making*, 13:90, 08 2013.
- [52] Shruti Kannan, Laura Morgan, McKenzie Cheung, Benjamin Liang, Christopher Lin, Dan Mun, Jean Francis, Vipul Chitalia, Joel Henderson, and Vijaya Kolachalama. Segmentation of glomeruli within trichrome images using deep learning. *Kidney International Reports*, 06 2018.
- [53] Ali Razavian, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson. Cnn features off-the-shelf: an astounding baseline for recognition. *Arxiv*, 03 2014.



MUHAMMAD UMER received his BS degree in Department of Computer Science, Khwaja Fareed University of Engineering & IT(KFUEIT), Pakistan (Oct-2014 to Oct-2018). Since Nov-2018, he got himself enrolled in Master of Computer Science(KFUEIT). He is also serving as Research Assistant at Fareed Computing & Research Center, KFUEIT, Pakistan. His recent research interests are related to Data Mining, mainly working machine learning & Deep Learning based IoT, Text Mining, and Computer Vision tasks.



GYU SANG CHOI received his Ph.D. in Department of Computer Science and Engineering, Pennsylvania State University, University Park, PA, USA in 2005. He was a research staff member at Samsung Advanced Institute of Technology (SAIT) for Samsung Electronics from 2006 to 2009. Since 2009, he has been a faculty member in Department of Information & Communication, Yeungnam University, Korea. His research areas include non-volatile memory and storage systems.



SAIMA SADIQ working as Assistant Professor in Department of Computer Science at Government Degree College for Women. She got herself enrolled in Master of Computer Science at Khwaja Fareed University of Engineering & IT (KFUEIT). Her recent research interests are related to Data Mining, Machine learning & Deep Learning based Text Mining.



ARIF MEHMOOD received his Ph.D. degree in Department of Information & Communication Engineering, Yeungnam University, Korea (Feb-2014 to Nov-2017). Since Nov-2017, he has been a faculty member in Department of Computer Science, KFUEIT, Pakistan. His recent research interests are related to Data Mining, mainly working on AI and Deep Learning based Text Mining, and Data Science Management Technologies.



MUHAMMAD AHMAD is an assistant professor at the Department of Computer Engineering, Khwaja Freed University of Engineering and Information Technology, Pakistan. Muhammad is associated with the research group of Advanced Image Processing Research Lab (AIPRL), a first Hyperspectral imaging lab in Pakistan. Muhammad is also associated with the University of Messina, Messina, Italy, as a research fellow.

Muhammad authored a number of research papers in reputed journals and conferences. He is a regular reviewer for a number of top tier Transactions/journals and conferences. His current research interests include Machine Learning, Computer Vision, Remote Sensing, Hyperspectral Imaging, and Wearable Computing.



SALEEM ULLAH was born in AhmedPur East, Pakistan in 1983. He received his B.Sc. and MIT degrees in Computer Science from Islamia University Bahawalpur and Bahauddin Zakariya University (Multan) in 2003 and 2005 respectively. From 2006 to 2009, he worked as a Network/IT Administrator in different companies. He completed his Doctorate degree from Chongqing University, China in 2012. From August 2012 to Feb 2016, he worked as an Assistant Professor in Islamia University Bahawalpur, Pakistan. Currently, he is working as an Associate Dean in Khwaja Fareed University of Engineering & Information Technology, Rahim Yar Khan since February 2016. He has almost 14 years of Industry experience in field of IT. He is an active researcher in the field of Adhoc Networks, IoTs, Congestion Control, Data Science, and Network Security.