

# Empirical Analysis of a Fine-Tuned Deep Convolutional Model in Classifying and Detecting Malaria Parasites from Blood Smears

Francis Jesmar P. Montalbo<sup>1\*</sup> and Alvin S. Alon<sup>2</sup>

<sup>1</sup> College of Informatics and Computing Sciences, Batangas State University  
Batangas City, 4200, Philippines  
[e-mail: francismontalbo@ieee.org]

<sup>2</sup> College of Engineering, Architecture and Fine Arts, Batangas State University  
Batangas City, 4200, Philippines  
[e-mail: alvin.alon@g.batstate-u.edu.ph]

\*Corresponding author: Francis Jesmar P. Montalbo

*Received August 13, 2020; revised November 16, 2020; accepted December 19, 2020;  
published January 31, 2021*

---

## Abstract

In this work, we empirically evaluated the efficiency of the recent EfficientNetB0 model to identify and diagnose malaria parasite infections in blood smears. The dataset used was collected and classified by relevant experts from the Lister Hill National Centre for Biomedical Communications (LHNCBC). We prepared our samples with minimal image transformations as opposed to others, as we focused more on the feature extraction capability of the EfficientNetB0 baseline model. We applied transfer learning to increase the initial feature sets and reduced the training time to train our model. We then fine-tuned it to work with our proposed layers and re-trained the entire model to learn from our prepared dataset. The highest overall accuracy attained from our evaluated results was 94.70% from fifty epochs and followed by 94.68% within just ten. Additional visualization and analysis using the Gradient-weighted Class Activation Mapping (Grad-CAM) algorithm visualized how effectively our fine-tuned EfficientNetB0 detected infections better than other recent state-of-the-art DCNN models. This study, therefore, concludes that when fine-tuned, the recent EfficientNetB0 will generate highly accurate deep learning solutions for the identification of malaria parasites in blood smears without the need for stringent pre-processing, optimization, or data augmentation of images.

---

**Keywords:** Malaria, Convolutional Neural Networks, Deep Learning, Fine-Tuning, Transfer Learning

---

This work was supported by Batangas State University. Without its support, this work would not have been possible.

## 1. Introduction

Malaria had been a devastating threat to human lives for the past decades and infecting about 228 million people worldwide. According to the 2018 report of the World Health Organization (WHO), malaria had claimed 405,000 lives worldwide, with 67% of the total recorded deaths from children ages below five [1]. The infection occurs when an Anopheles mosquito transfers various plasmodium parasites from its body by biting into an unsuspecting victim [2]. Commonly, the most appropriate diagnosis is the use of thick and thin blood smears on glass slides extracted from a potentially infected host. The slides then pass through a light microscope for observation. However, this diagnostic method requires an advanced level of proficiency to attain accurate results [3]. Furthermore, additional complexity during diagnosis and treatment can lead to unreliable and slow outcomes due to insufficient specialized equipment and expertise in developing areas with many cases [4].

In recent years, the difficulty of relying only on experts and expensive medical equipment to perform medical diagnosis had reduced with the help of Deep Learning (DL). Medical experts can now detect and recognize various life-threatening diseases more rapidly, increasing efficiency in decision making [5-7]. In this modern era, the growing fascination of DL researchers had continuously made improvements in medical imaging. With the state-of-the-art DL models, the process of medical diagnosis had never been the same [8-10]. An influential and well-known DL model, the Convolutional Neural Network (CNN), revolutionized the way machines interpret images in the modern era. With CNN, computers can learn patterns from a large set of 2D-array images using striding filters, a backpropagation algorithm, and various combinations of techniques that aim to generate accurate predictions like a real human being [11].

Fig. 1 illustrates the CNN structure and how it learns from features to predict an input image. CNN performs feature extraction with a multi-layer architecture that consists of several methods. The convolution process involves an  $N \times N$  (e.g.,  $3 \times 3$ ) sized filter that captures basic patterns from the original image to generate a stack of feature maps passed to a series of succeeding layers. The pooling layers then downsize the entire feature set with a smaller striding filter (e.g.,  $2 \times 2$ ) for efficiency. The process repeats until it reaches the fully-connected (FC) layers that merge the patches of high to low-level features to generate a highly detailed representation. Each FC layer and convolution contains neurons that are weighted based on the importance of features learned. The activation layer or classifier then yields the probability of what the unseen image represents based on the accumulated scores from the network's earlier levels [12].

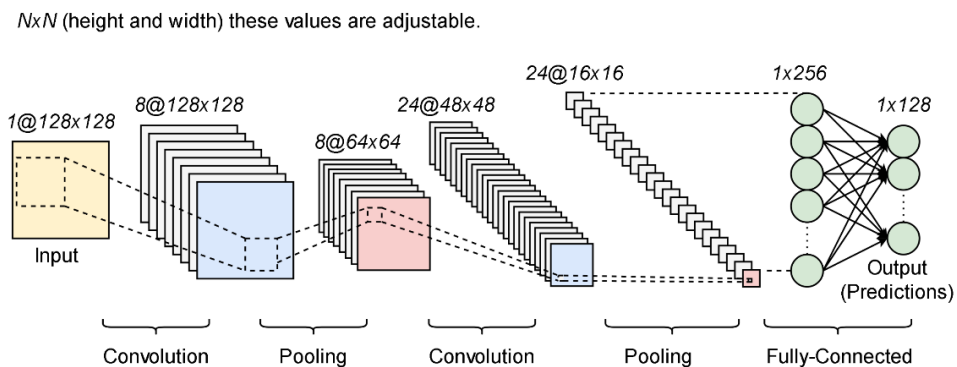


Fig. 1. The architecture of a standard convolutional neural network

CNN has recently achieved success in image classification, recognition, localization, and detection challenges on numerous occasions to solve real-world problems. The rise of more extensive, more profound, and more complex CNNs became a trend in the field of DL and computer vision. State-of-the-art models like AlexNet [13], VGG [14], GoogleNet [15], and ResNet [16] made significant improvements since the original CNN. These changes include adding layers, increasing the network width with separable convolutions, and eliminating the results' saturation from the increased model volume [17, 18]. The previous CNN became more complex than ever, turning it into a Deep Convolutional Neural Network (DCNN). DCNNs had proven its capability through the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC). A competition that put computer vision solutions to the test; this test includes 1000 images with a million samples to be classified simultaneously. However, replicating DCNNs for specific purposes requires a tremendous amount of computing power due to its architectural complexity and data requirement [19]. With this dilemma, a new method emerged called transfer learning. This approach solved the re-usability of large-scale DCNNs by transferring its pre-trained weights from ImageNet that contain essential features for image recognition to train other models. With fine-tuning methods, DCNNs can now perform specific tasks beyond its original purpose without the need for tremendous computing resources that can even assist in the field of medical image analysis [20]. Currently, several works have explored such methods to perform the automated task of detecting malaria parasites from blood cell images. DCNNs through the years have kept getting better to cope up with various demands not just in classification accuracy but also in overall efficiency and scalability. However, not much study had employed transfer learning and fine-tuning to recently released DCNN models to perform the task of malaria parasite detection. Therefore, in our work, we focused on employing these methods to recent DCNNs for malaria parasite detection in blood smears to yield new findings and conclusions that may establish a new perspective for future researchers that may also tackle such a difficult task.

## 2. Related Works

In this section, we reviewed several related works that employed CNN and similar solutions in detecting and classifying infections of malaria parasites from blood smears.

To assist in the growing cases of malaria, the work of Devi et al. devised a computer-assisted solution by combining several Machine Learning (ML) models to develop a hybrid classifier. Through microscopic samples of malaria-infected and non-infected blood cells, their work applied SVM, KNN, Naïve Bayes, and Artificial Neural Networks (ANN), which attained a maximum 96.32% accuracy with ANN in detecting parasitic infections. The model trained using images with a morphological segmentation, which extracted the cell pixels from the background using thresholds and watershed techniques [21].

With the growing popularity of DL and CNN, Liang et al. proposed their version of a 16-layer CNN. At that time, they indicated that CNN is a robust algorithm that can learn visual patterns effectively compared to a classical ML. Their intuition behind the added layers enhances the existing CNN's learning capability and generates more feature maps to exceed its previous performance. They managed to end up with a successful 97.37% accuracy, using randomly generated weights, concluding that their modified CNN can outperform even the current state-of-the-art models at the time in terms of malaria classification [22].

Soon after, another popular DL model was proposed for the same tasks by Bibin et al. Their work is the first to apply a Deep Belief Network (DBN) to perform a binary class classification of malaria blood smears. DBN is a multi-stacked layer of Restricted Boltzmann Machines

(RBM), where each node pass and receives inputs multiplied by its corresponding weights. Their DBN had a depth of 484 visible units and an output layer of two, containing four hidden layers that had 600 nodes each. The results attained an accuracy of 96.21%, indicating DBN's effectiveness in classifying malaria-infected and non-infected blood smears [23].

With the ease of fusing new ideas into a CNN model, Gopakumar et al. customized a CNN model for a glass slide scanner to detect malaria parasite infections. The custom CNN model trained with focus stacked samples managed to resolve cells' counting dilemma through a two-level segmentation approach. Their results became successful as it reached an accuracy of 98.47%. Their study provided a low-cost, quick, and easy way to diagnose malaria compared to existing solutions on the market [24].

Gezahegn et al. performed other methods combining ML and image processing. Their work compared handcrafted feature extractors' performances like Scale Invariant Feature Transform (SIFT) and traditional extraction methods for classifying malaria infections. The classifier used was the SVM algorithm. Their accuracy achieved 78.89% with a sensitivity of 80% and 76.67% specificity with their trained model. The study showed that feature extraction with the conventional SVM for detecting and diagnosing malaria is still improvable if a robust feature set existed. However, with the limited number and difficulty of data collection, fewer features are generated, making it difficult for the selected algorithms to learn further [25].

With the work of Rajaraman et al., instead of traditional feature extraction methods, they used DCNNs to produce a higher number of features from a balanced dataset of 27,558 cell images from blood smears, infected and non-infected by malaria parasites. Their selected models that performed the classification consist of AlexNet, VGG16, ResNet50, Xception, DenseNet-121, and their proposed model. The following are considered state-of-the-art employed with an additional number of layers and sophisticated modifications compared to conventional CNN. They generated exceptional classification results from training the classifiers with features extracted from optimal layers of each model to produce accuracy scores of 94.4%, 95.9%, 95.9%, 91.5%, 95.2%, and 92.7%, respectively with the mentioned models. Even with similar accuracies, ResNet50 attained higher specificity and performance than VGG16. Their work also included optimization measures like hyper-parameter optimization and regularization techniques to increase model performance [26].

Vijayalakshmi et al. then introduced a hybrid DCNN combined with traditional ML. The architecture is composed of a pre-trained VGG feature extractor with an SVM classifier. Their VGG-SVM applied transfer learning to train the upper layers while maintaining pre-trained parameters from ImageNet. The results generated an accuracy of 89.21% with VGG16-SVM, while 93.13% from the VGG19-SVM [27].

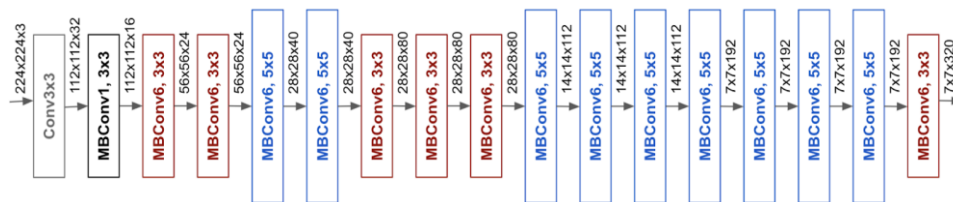
Due to the limited data produced for such a complicated task, Pattanaik et al. decided to use unsupervised algorithms rather than data augmentation methods. Their work proposed a novel multi-layered CAD scheme with various node sizes incorporated with a Functional Link ANN (FLANN) and a Stacked Sparse Auto Encoder (SSAE) to increase the size of features during training. Through their approach, they applied the 10-folds cross-validation and achieved an impressive accuracy of 89.10% from a limited dataset of 1182 images [28].

Another work from Pattanaik et al. employed a novel Multi-Magnification ResNet (MM-ResNet). Their novel MM-ResNet had built several concatenations of input and output layers to prevent the vanishing gradient problem and attain improved performance compared to other state-of-the-art. With that said, their model preserves the small to large feature samples across the network without drastic saturation and achieves better information handling than the baseline model. Together with the parallelism technique to train multiple inputs, they trained their model with 1000 epochs and achieved a remarkable 98.08% accuracy [29].

The growing importance of mobile applications made Fuhad et al. develop a mobile-based application for classifying malaria infection using SVM and KNN. Their method used CNN as the feature extractor applied with knowledge distillation, augmentation, and autoencoders. Through their process reached an accuracy of 99.23% with the help of autoencoders. A generative DL model helped create synthetic versions of the original input using an encoder and decoder. Even with such a robust approach, their work still delivered a lightweight model that can work efficiently with mobile devices [30].

The following works discussed made significant contributions for diagnosing malaria infection in blood smears. However, as we mentioned in our introduction, other forms of DCNN models came out recently that may also have a significant impact on malaria diagnosis. Upon our investigation, a recent study made by Marques et al., [31] presented that a recent DCNN like the EfficientNet model could perform remarkably in terms of several medical image classifications like chest x-rays. They conducted experiments with the EfficientNet model and tackled the task of chest x-ray classification from samples with and without COVID-19 infections. Justified by their results, their fine-tuned EfficientNet outperformed some of the previously released well-known DCNNs like VGG [14], ResNet [16], MobileNet [32] with a 99.62% accuracy. With their work, it has shown to us that EfficientNet has the potential to emit promising results in other fields of medical imaging. In the best of our knowledge, no existing study had made use of the said model in terms of malaria parasite detection and classification from blood smears.

**Fig. 2** presents the recent EfficientNetB0 baseline model architecture. We propose to use the EfficientNetB0 baseline model as our entry point that takes in an input image with a 224x224x3 dimension. The model then extracts features throughout the layers by using multiple convolutional (Conv) layers using a 3x3 receptive field and the mobile inverted bottleneck Conv (MBConv). Our intuition to employ the EfficientNetB0 is due to its balanced depth, width, and resolution that can produce a scalable yet accurate and easily deployable model. Compared to other DCNNs, EfficientNetB0 scales each dimension using a fixed set of scaling coefficients. This approach surpassed other state-of-the-art models that trained on the ImageNet dataset. Even with transfer learning, EfficientNet still achieved exceptional results, indicating its effectiveness beyond the usual ImageNet dataset. In its release, the model had scales of 0 to 7, showing an increase of parameter size and even accuracy. With the recent EfficientNet, users and developers can access and provide improved ubiquitous computing imbued with DL capabilities in several platforms for various needs [33].



**Fig. 2.** EfficientNetB0 baseline model architecture [33]

It is worth to mention that this work focuses on the empirical analysis of the recent DCNNs and EfficientNetB0 towards the classification of malaria parasitized and uninfected blood smears.

Our work mainly contributes to the use of the EfficientNetB0 model with modified ending layers incorporated through layer freezing via fine-tuning and trained to address the problematic classification and detection of malaria parasites from blood smears. With our

proposed method, we aim to produce a model that consumes a very minimal disk space that makes it easily deployable, transferrable, and reproducible as needed, whether in the local network or through the internet without sacrificing performance. Our work also compares and analyzes the performance of EfficientNetB0 to other recent state-of-the-art models that have not yet performed the classification of malaria-infected blood smears. Considering this inclusion can emanate additional perspectives for other researchers in the future. Lastly, our work provides transparency by providing the detection capability of the modified EfficientNetB0 as opposed with other state-of-the-art models in terms of malaria parasite localization through the Gradient-Weighted Class Activation Mapping (GRAD-CAM) algorithm that the other studies did not include or performed.

### 3. Materials and Methods

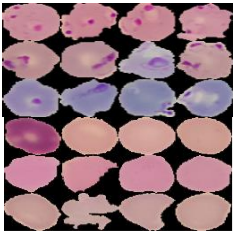
#### 3.1 Malaria Dataset

The dataset used for this work came from the work of Rajaraman et al. [34]. **Table 1** presents the dataset specification with a total of 27,558 blood smears separated into two different classes, namely, malaria parasitized and uninfected blood cells, each having 13,779 copies and has a three-channel format of Red, Green, Blue (RGB), making it suitable for training.

In our work, we allocated the dataset equally for both classes to prevent class superiority over the other [35]. To do this, we parted the dataset into two, having 80% (22046) for training and 20% (5512) for validation. The images are partitioned in a stochastic fashion to prevent the increase of bias that may affect the model [36, 37]. Furthermore, we did not include any forms of augmentation or extensive pre-processing to let the models train on the actual images as it is, similarly to how it would be in most real-life situations. This reason is to highlight the strength of recent DCNN using limited training data but with a reasonable amount of validation data.

However, the collection of images had no fixed dimension. Therefore, we normalized each image sample using an automated resizing script with Keras [38] that automatically resized all inputs into a 224x224 dimension.

**Table 1.** Specification of malaria parasitized and uninfected dataset

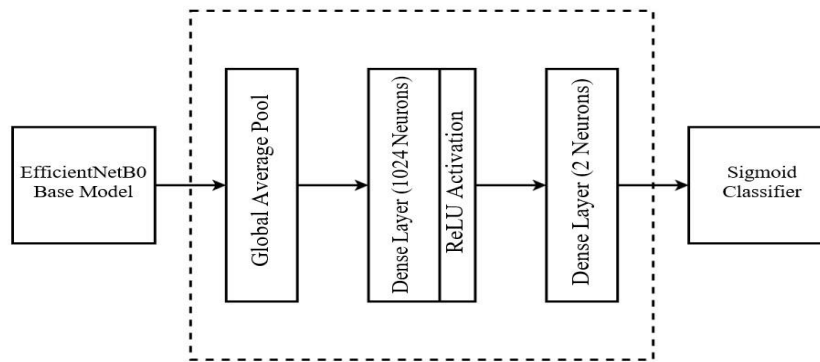
Samples	Label	Train (80%)	Validation (20%)
	Parasitized	11023	2756
	Uninfected	11023	2756
Total: 27,558		22046	5512

#### 3.2 Proposed Layers

For us to make use of the pre-trained EfficientNetB0 to learn from our prepared data, we proposed to replace the final layers of the original EfficientNetB0 baseline model and add our set of layers to activate new weights from it.

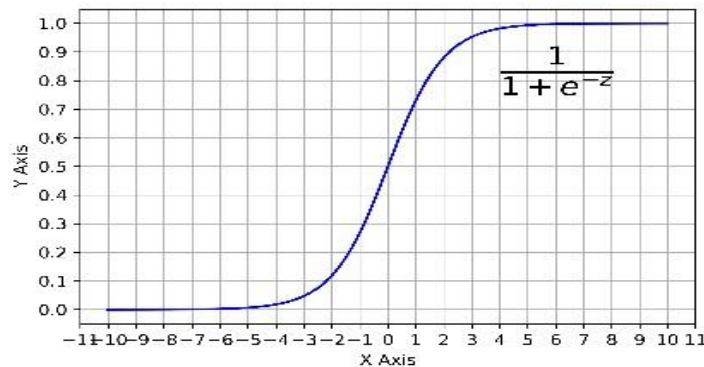
**Fig. 3** illustrates our proposed final layers for the EfficientNetB0 composed of a Global Average Pool (GAP), two FC dense layers, and a sigmoid classifier. To prevent instances of severe overfitting from the sophisticated feature handling, we added a pooling layer. The GAP layer further reduced the number of parameters by rescaling the height, width, and depth of the incoming tensor from the base model into a 1x1x3 dimension, respectively. We controlled the tremendous burst of features passing to the dense layer that can overwhelm the classifier by doing this method. In the said process, the GAP did not entirely reduce the portions of the feature maps. Instead, it averaged the entire spatial features and maintained the most intricate patterns required to recognize the image [39]. The GAP was also proven to work reliably with DCNNs in handling medical image classifications [40].

Before predicting results, we directed the feature sets from the previous GAP to a dense layer consisting of 1024 hidden units connected to another dense layer with two neurons representing our two given labels. This approach provided a new set of weights and biases to each feature map through a linear fashion that produced a probability. Also, we applied a Rectified Linear Unit (ReLU) activation on the hidden layer of the 1024 dense layer to provide non-linearity and speed up the training process, as ReLU can immediately adjust all negative input values of the previous layer to zero [41].



**Fig. 3.** Proposed final layers for fine-tuning the EfficientNetB0 model

**Fig. 4** presents our selected classifier. The Sigmoid activation function served as a replacement for the multi-class classifier, Softmax. Similarly, Sigmoid is also a logistic function that specifically performs a binary class classification [42]. The S-shaped non-linear function binds values to a 0 or 1, indicating either a parasitized or an uninfected blood cell, respectively. The two neurons of the last dense layer represent these classes.

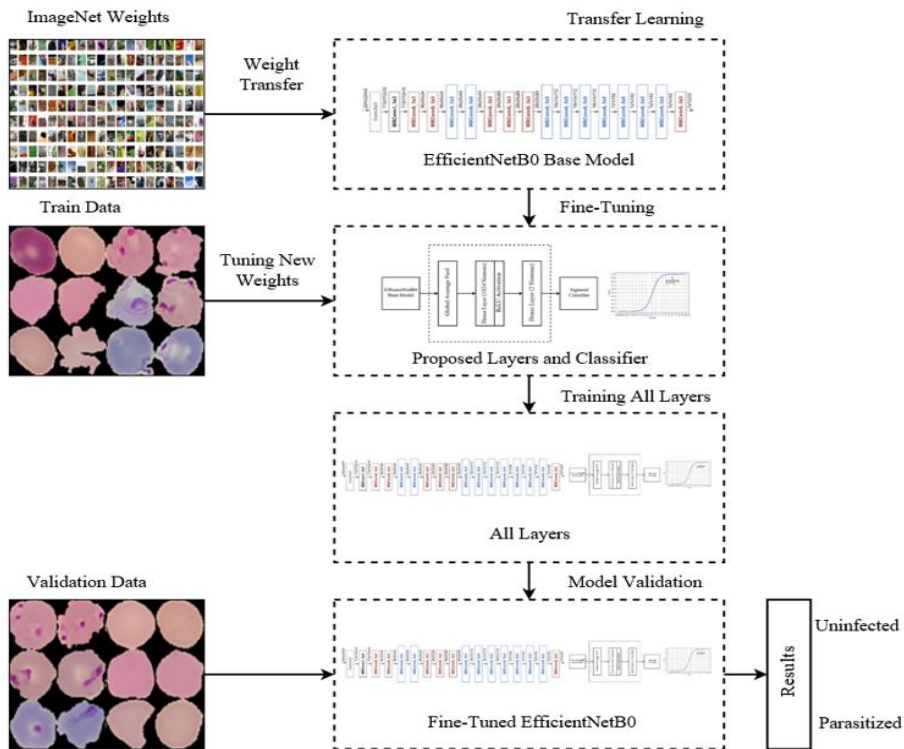


**Fig. 4.** Graph of the sigmoid activation function

### 3.3 Transfer Learning and Fine-Tuning

**Fig. 5** shows how we trained and finalized our model. First, with transfer learning, we enlarged the size of the training parameters pre-emptively. With the pre-initialized weights from ImageNet, the base model instantly used its features and improved image recognition. ImageNet weights contain features that can help detect shapes, edges, and other vital components needed for an image classification task [19]. This method accelerated the process with reduced efforts compared to randomly initialized weights [13].

As shown in the figure, our base model pre-trained with the ImageNet data consists of 1000 different classes, with over 14 million images [43]. With this matter, fine-tuning is imperative, as the current weights and structure of the EfficientNetB0 cannot immediately work for our selected task [44]. Therefore, we froze the beginning layers on the base model, then trained our proposed ending layers with the malaria train data through fine-tuning. We managed to preserve the ImageNet features within the extraction layers with this approach and prevented it from getting overwritten during training updates. Subsequently, after training both the extractor and our proposed layers, we re-trained the entire network with the malaria dataset and the ImageNet weights and produced our final model. We then validated the final model with the use of our validation data.



**Fig. 5.** Transfer learning and fine-tuning process of the EfficientNetB0 malaria parasite classifier

### 3.4 Hyper-Parameters and Loss Function

This section presents the selected appropriate hyper-parameter values and loss function for the task to yield efficient results.

Identifying the performance of a DL model does not solely rely on accuracy but also in terms of loss [45]. A DL model's main objective is to attain its lowest possible rate of errors



as a model with a less calculated loss indicates better efficiency [46]. In this work, we selected the cross-entropy (CE) loss function to calculate the average measure in the distinction between the expected and predicted value. Equation (1) shows the measurement of loss for binary classification, where  $y$  represents the binary values of 0 or 1, and  $p$  is the probability [47].

$$CE = -(y \log(p) + (1 - y) \log(1 - p)) \quad (1)$$

To ensure the optimal decrease of loss during training, we directly selected Adam as our optimizer. This optimization algorithm operates as an adaptive gradient descent function that helps the weights decline faster towards the local minima [48]. We primarily chose Adam over other optimizers due to its ease of implementation, efficient memory consumption, and faster learning phase, compared to others like Stochastic Gradient Descent (SGD) [49] or RMSProp [50]. It is also worth mentioning that Adam recently had successful DL implementations that trained models for assisting in medical imaging analysis [51].

**Table 2** presents our hyper-parameter settings, where we have set a small learning rate (LR) to work well with the other selected hyper-parameters. Adam worked effectively and achieved rapid convergence in a short period than SGD [52]. The batch size of 32 provided a decent load to pass information through the network without consuming our entire computing memory. Moreover, we selected several durations to train each model increasingly to see how it will behave over time. With 10, 25, 50, and 100 epochs, the following are considered shorter than the other works. However, we primarily decided to use smaller numbers as we applied a faster optimizer.

**Table 2.** Selected hyper-parameters for training

Hyper-parameters	Value
Optimizer	Adam
LR	0.0001
Batch Size	32
Epochs	10, 25, 50, 100

### 3.5 Evaluation Metrics

In ML or DL, the Confusion Matrix (CM) is a standard tool that visualizes how accurate a trained model can predict from a respective validation dataset. The CM has corresponding rows and columns representing the actual class and the ground truth labels, consisting of a parasitized, uninfected blood cells. Simultaneously, the predicted values indicate the number of correct and incorrect predictions or classifications made for each validation sample. The True Positive (TP) denotes the number of correctly classified positive samples as positive, while True Negatives (TN) corresponds to the correctly predicted negatives as negatives. False Positives (FP) are predictions where the image was classified as positive but is not. In contrast, False Negatives (FN) are negative results but are positive [53].

With the following values, we computed for the overall accuracy, precision ( $PR$ ), sensitivity ( $SE$ ), specificity ( $SP$ ), and the F1-score of each model. In this work,  $SE$  signifies the ratio of correctly predicted parasitized cells or TPs to all TPs and FNs, while  $SP$  refers to the other way around.  $PR$  is the frequency of how often the model makes a correct prediction of an actual class. The accuracy is the total of all accurate predictions out of all the given samples. At the same time, F1-score is the weighted average between  $PR$  and  $SE$  [53].

The following performance metrics are calculated based on the given equations below.

$$\text{Sensitivity} = TP / (TP + FN) \quad (2)$$

$$\text{Specificity} = TN / (TN + FP) \quad (3)$$

$$\text{Precision} = TP / (TP + FP) \quad (4)$$

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (5)$$

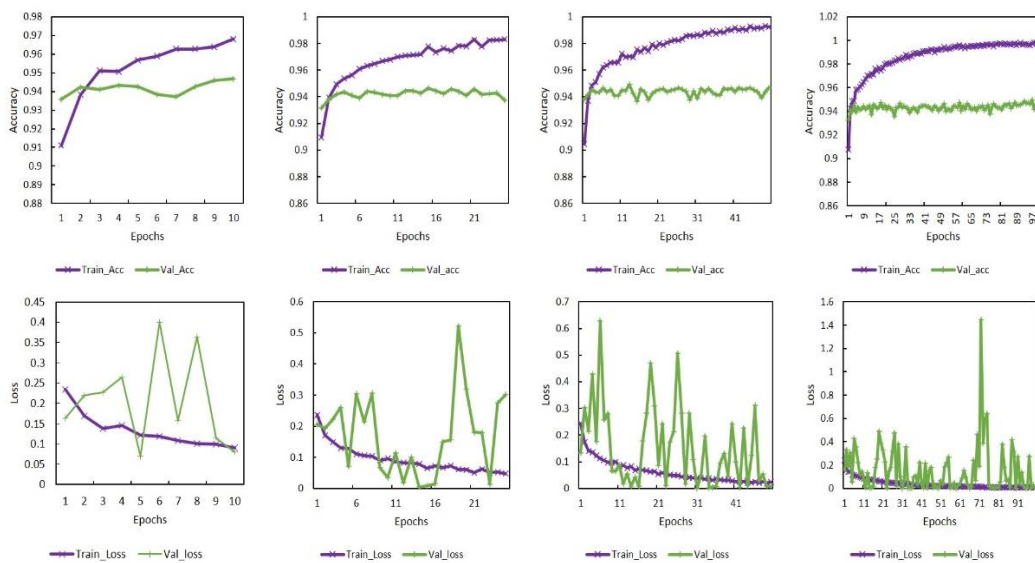
$$F1 - \text{Score} = 2TP / (2TP + FP + FN) \quad (6)$$

## 4. Experimental Results and Discussion

This section discussed the results generated during training and validation using the prepared dataset. We identified the number of correctly classified and misclassified samples using a CM and calculated the performance through the mentioned metrics. To further evaluate the efficiency, we included the produced weight sizes and the Grad-CAM algorithm's activation maps.

### 4.1 Accuracy and Loss

**Fig. 6** presents the accuracy and loss graphs against the various epochs calculated with the CE loss. The proposed model's training and validation accuracy had shown a rapid increase within a short period with the given hyper-parameter values. However, the results of the validation accuracy stopped increasing at around 93% to 94%. In contrast, the validation loss had a bit of trouble and generated some unstable oscillations on the graph. Nonetheless, at the end of each phase, we evaluated that all models still had an exceedingly small error gap, except for the model that validated with 25 epochs that had a severe overfitting problem at the end.



**Fig. 6.** Train and validation trends of the proposed model using various epochs

## 4.2 Classification Performance

To determine the classification performance, we utilized the CM for a visual understanding of how well the model classified each sample. We then calculated its performance using our selected performance metrics from section 3.5.

**Table 3** presents the EfficientNet model results that trained with various epochs using a validation set of 5512 samples of malaria parasitized (2756) and uninfected (2756) blood cells.

**Table 3.** Confusion matrix

Epochs	Classes	Actual Values	
		<i>Parasitized</i>	<i>Uninfected</i>
10	<i>Parasitized</i>	2573	183
	<i>Uninfected</i>	110	2646
25	<i>Parasitized</i>	2612	144
	<i>Uninfected</i>	202	2554
50	<i>Parasitized</i>	2587	169
	<i>Uninfected</i>	123	2633
100	<i>Parasitized</i>	2559	197
	<i>Uninfected</i>	109	2647

In **Table 4**, using the CM matrix, we calculated the overall performance of the EfficientNetB0 according to its accuracy, *PR*, *SE*, *SP*, and F1-Score. Additionally, we compared our results with other recent state-of-the-art DCNNs trained and validated using the same approach. The number that follows the name of the model corresponds to the number of epochs.

EfficientNetB0-50 attained the highest accuracy of 94.70% from our evaluated results, followed by EfficientNetB0-10 with 94.68%. On the other hand, InceptionResNetV2-50 managed to reach the highest *PR* of 95.10%, followed by EfficientNetB0-25 with 94.78%. In terms of *SE*, ResNet152V2-100 performed best with 97.25% and ResNet152V2-100, next to it with 96.18%. While in *SP*, InceptionResNetV2-50 achieved the highest score of 94.98% and having EfficientNetB0-25 next to it with a 94.66%.

Therefore, EfficientNetB0-50, with the highest F1-score of 94.66%, states overall dominance towards the task than the other state-of-the-art trained in this work. Furthermore, we observed that training up to 100 epochs with our proposed approach did not show much performance improvement. This reason mainly occurred due to the limited train data and large validation data ratio.

**Table 4.** Comparison of performance with other recent state-of-the-art models

Model	ACC (%)	PR (%)	SE (%)	SP (%)	F1 (%)
EfficientNetB0-10	94.68	93.36	95.90	93.53	94.61
EfficientNetB0-25	93.72	94.78	92.82	94.66	93.79
EfficientNetB0-50	<b>94.70</b>	93.87	95.46	93.97	<b>94.66</b>
EfficientNetB0-100	94.45	92.85	95.91	93.07	94.36
ResNet152V2-10	93.85	92.63	94.94	92.81	93.77
ResNet152V2-25	94.03	92.02	95.88	92.33	93.91
ResNet152V2-50	94.32	92.31	96.18	92.61	94.20
ResNet152V2-100	94.29	91.15	<b>97.25</b>	91.67	94.10
NASNetMobile-10	93.29	92.96	93.57	93.01	93.27

NASNetMobile-25	93.67	93.21	94.07	93.28	93.64
NASNetMobile-50	93.56	93.00	94.06	93.07	93.52
NASNetMobile-100	93.21	92.34	93.98	92.48	93.16
InceptionV3-10	93.31	93.58	93.07	93.54	93.32
InceptionV3-25	93.63	93.76	93.52	93.74	93.64
InceptionV3-50	93.32	93.80	92.92	93.74	93.36
InceptionV3-100	93.02	90.46	95.33	90.92	92.83
InceptionResNetV2-10	92.56	89.04	95.78	89.76	92.29
InceptionResNetV2-25	94.19	92.49	95.76	92.74	94.09
InceptionResNetV2-50	93.90	<b>95.10</b>	92.88	<b>94.98</b>	93.98
InceptionResNetV2-100	94.03	91.87	96.02	92.21	93.90

### 4.3 Saliency Maps

We employed the Grad-CAM algorithm to present a visualized localization of salient features on the infected blood cell to add transparency to our results. The Grad-CAM is a generic algorithm that makes use of the final activation layer of a CNN model. From there, a set of high-level features generate in the form of a heatmap.

According to Selvaraju et al., the algorithm calculates the classes' gradient score to its corresponding feature maps from the convolutional layers to generate the visual heatmap. The returning gradients then enter the GAP to capture the importance of the individual weights. Subsequently, the combined weighted feature maps are activated with a Rectified Linear Unit (ReLU) function to achieve the Grad-CAM heatmap [54].

In Fig. 7(b), the generated features in the form of a heatmap indicate the area where the model detects an infection. To make use of (b) effectively, we pre-processed the input image (a) and overlapped it with our produced heatmap (b) to see how the model detected an infection. The output (c) from this approach had shown us how the model created its decision based on its interpretation. Furthermore, the results of EfficientNetB0 in Table 4 consistently shown that even with the change in accuracy and epoch, there is no observable difference seen on its Grad-CAMs. However, we observed a significant change and inconsistency in attention based on heatmaps' dissipation across the infection's surrounding areas with the other models.

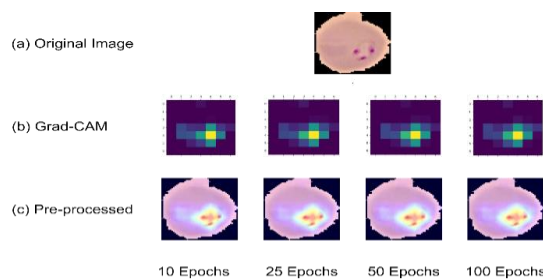
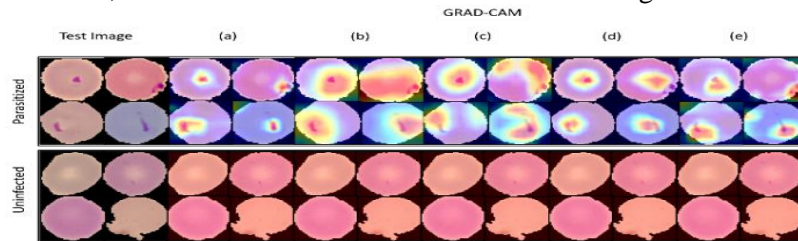


Fig. 7. Process and application of the Grad-CAM algorithm to the original image

In Fig. 8, we present the comparison of parasite detection using the Grad-CAM algorithm were (a) to (e) showed four parasitized and four uninfected test samples for each model. From our observations, the heatmap of (b) and (c) dissipated widely, indicating a dynamic shift of attention towards the other sections and away from the object of interest. While (d) and (e) somehow managed to keep its attention on the crucial areas. Unlike the rest, (a) presents the most compelling interpretation and precise detection of parasites as the heated area maintained the infection's better localization. The other samples with a more dispersed heatmap may have a higher chance of misclassification due to the increased detection of irrelevant features [55].

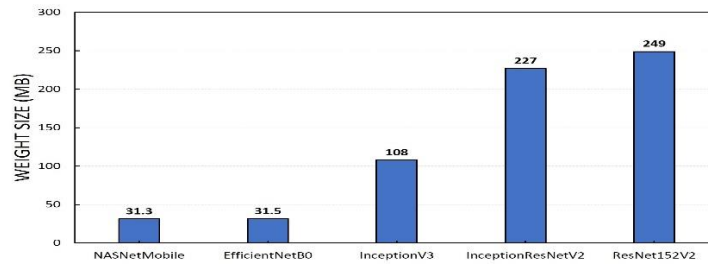
As for the uninfected samples, all models produced a clear vision. This outcome tells us that the models managed to identify an uninfected blood sample much more comfortable than a parasitized blood cell, and EfficientNetB0 had the most outstanding results.



**Fig. 8.** Comparison of Grad-CAM samples of (a) EfficientNetB0 with (b) InceptionResNetV2, (c) InceptionV3, (d) NASNetMobile, and (e) ResNet152V2

#### 4.4 Weight Size

**Fig. 9** presents our generated weight sizes regardless of the epochs as it did not contribute to the change of size. The following shows that NASNetMobile produced the lowest with only 31.3MB, considering its design solely made for mobile devices. Next to it, EfficientNetB0 had a size of 31.5MB, making it only a bit larger than NASNetMobile. For the other trained models, ResNet152V2 had the largest with 249MB, InceptionResNetV2 with 227MB, and InceptionV3 with 108MB. The reason for the enlarged capacity requirement came from the number of parameters, model depth, and complexity.



**Fig. 9.** Comparison of weight sizes with other recent state-of-the-art models

#### 4.5 Discussion

According to our evaluated results, our pre-trained and fine-tuned EfficientNetB0 model did not require stringent pre-processing, optimization, data augmentation, and even long training epochs to attain a highly accurate performance. Even with a less rigorous selection of hyper-parameter values and training only with a minimum of ten epochs, the fine-tuned EfficientB0 combined with our proposed layers achieved an accuracy of 94.68% from our validation data of 5512 images. It then increased further up to 94.70% with fifty epochs. However, training the model further than fifty epochs did not contribute any improvements in terms of accuracy. Furthermore, to show our proposed method's improvements compared to the EfficientNetB0 trained without it, we present a set of results in **Table 5**.

As presented, we observed a significant difference in performance between an EfficientNetB0 trained with and without our method. The EfficientNetB0 trained end-to-end without our proposed freezing approach, and replacement layers attained its highest validation accuracy of only 92.80% from the similar validation set of 5512 images with 100 epochs. Unlike the model with our proposed method, the highest accuracy achieved was 94.70% accuracy trained with only 50 epochs, making it a better approach. From these results, we

identified that the model trained with the replacement layers compared to the conventional structure had better classification prowess in terms of malaria-infected and uninfected blood cells due to the additional features produced mainly for it. With that said, our procedure to incorporate the proposed replacement layers with the freezing method attained a beneficial accelerated convergence and better overall performance for the EfficientNetB0.

**Table 5.** Results of the proposed EfficientNetB0 model with and without the proposed method

Method	Epochs	ACC (%)	PR (%)	SE (%)	SP (%)	F1 (%)
With our proposed method	10	94.68	93.36	95.90	93.53	94.61
	25	93.72	94.78	92.82	94.66	93.79
	50	94.70	93.87	95.46	93.97	94.66
	100	94.45	92.85	95.91	93.07	94.36
Without our proposed method	10	91.15	90.53	91.66	90.65	91.09
	25	91.96	92.42	91.59	92.35	92.00
	50	92.34	91.76	92.84	91.86	92.30
	100	92.80	91.47	93.96	91.69	92.70

For further comparison, **Table 6** presents a performance summary of this work and other existing studies that used CNN-based solutions for malaria parasite detection and blood smears classification. It is worth to mention that this work does not directly compare the following works due to the differences in data preparation, training and validation methods, and computing resources used. However, other researchers can still attain additional perspectives with the current solutions produced for the said topic through the given summary and discussion.

**Table 6** shows that our work did not achieve the highest accuracy due to the mentioned differences in approach. However, this work still has much room to scale and improve as our model relied only on the base EfficientNetB0. The recently released model became the advantage of this work as it can train in a shorter span and still achieve remarkable results than others that required hundreds to thousands of epochs, making it easier to reproduce and improve as needed. Even with a non-specialized standard GTX 1070 with only 8GB RAM, our work attained such a result. Our produced model also consumed a minimal disk space that conveys the ease of deployment and transfers on most platforms, whether local or the internet, considering that EfficientNet models existed to contribute to ubiquitous computing compared to previous state-of-the-art DCNNs that came before it like like the base ResNet50, and even the conventional CNNs [56]. With that said, future users, researchers, and the likes can benefit from this work that may require its purpose, particularly in developing countries with low-end resources or poor internet connectivity.

Nonetheless, this work still lean towards an empirical analysis rather than a developmental approach about how recent DCNN models, specifically EfficientNet, in malaria parasite detection and classification from blood smears performs.

**Table 6.** Comparison of performance with other similar studies that used CNN-based models

Method	ACC (%)	SE (%)	SP (%)
EfficientNetB0 (Our work)	94.70	95.46	93.97
16-Layer CNN [22]	97.37	96.99	97.75
DBN [23]	96.21	97.60	95.92
Custom-CNN [24]	98.47	97.06	98.50
ResNet50 [26]	95.9	94.7	97.2
MM-ResNet [28]	98.08	95.38	98.30
Novel CAD Scheme [29]	89.10	93.90	83.10

## 5. Conclusion

Due to the difficulty of having an accessible, low-cost, rapid, and accurate malaria diagnosis in most developing countries, many people still suffer from early fatalities. With that said, the mortality rate caused by malaria kept increasing. Therefore, we had the initiative to conduct an empirical analysis of recent DCNNs in classifying and detecting malaria parasite infections from blood smears to contribute to solving this problem. In our work, a recent state-of-the-art model like EfficientNetB0 trained with the use of pre-learned weights and fine-tuned to classify between a parasitized and uninfected blood cells from blood smears. With the help of open-source data from NIH, we trained the model using 22046 images of the mentioned blood cells. The data then had minimal pre-processing to normalize the input data to a 224x224 dimension for added efficiency with our selected model. Upon our evaluation with 5512 images, the EfficientNetB0 outmatched recent state-of-the-art DCNNs like NASNetMobile, InceptionV3, InceptionResNetV2, and ResNetV2-152. The highest attained accuracy rate of 94.70% came from the EfficientNetB0 trained for only 50 epochs, while the 10-epoch variant also had a remarkable 94.68%. Such minimal difference in performance regarding the training length justifies that the 10-epoch variant is better for most cases.

As concluded from our results and discussion, the fine-tuned EfficientNetB0 based on transfer learning could efficiently classify and detect parasitized and uninfected blood samples when fine-tuned properly. Even with the absence of intricate image pre-processing, augmentation, and cumbersome optimization methods, EfficientNetB0 can attain exceptional results. Hence, the lightweight and easily reproducible model can help diagnose malaria infections in areas that require such a solution even without access to high-end computing resources. However, we do not guarantee that this work will have an entirely similar performance to a real-life scenario. During and after our experiments, we hypothesized a particular caveat on training DCNN models, that it might not generalize well with external samples compared to the internal validation set used in most studies including this work due to the difference of capture devices and the morphological complexity in blood cell samples worldwide. Therefore, we recommend performing further studies that include the collection of additional data from other countries with the use of various capture devices, and potential deployments in diverse platforms to analyze this hypothesis and generate future work improvements to guarantee worldwide acceptability.

## References

- [1] World Health Organization, "The World malaria report 2019' at a glance," 2019. [Article \(CrossRef Link\)](#)
- [2] K. J. Arrow, C. Panosian, and H. Gelband, "Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance," The National Academies Press, 2004. [Article \(CrossRef Link\)](#)
- [3] A. Martín-Díaz, J. M. Rubio, J. M. Herrero-Martínez, M. Lizasoain, J. M. Ruiz-Giardin, J. Jaqueti, J. Cuadros, G. Rojo-Marcos, P. Martín-Rabadán, M. Calderón, C. Campelo, M. Velasco, and A. Pérez-Ayala, "Study of the diagnostic accuracy of microbiological techniques in the diagnosis of malaria in the immigrant population in Madrid," *Malaria Journal*, vol. 17, 2018. [Article \(CrossRef Link\)](#)
- [4] J. Osei-Yeboah, G. K. Norgbe, S. Y. Lokpo, M. K. Kinansua, L. Nettey, and E. A. Allotey, "Comparative Performance Evaluation of Routine Malaria Diagnosis at Ho Municipal Hospital," *Journal of Parasitology Research*, vol. 2016, pp. 1-7, 2016. [Article \(CrossRef Link\)](#)
- [5] S. H. Kassani and P. H. Kassani, "A comparative study of deep learning architectures on melanoma detection," *Tissue and Cell*, vol. 58, pp. 76-83, 2019. [Article \(CrossRef Link\)](#)

- [6] A. Kumar, J. Kim, D. Lyndon, M. Fulham, and D. Feng, "An ensemble of fine-tuned convolutional neural networks for medical image classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 31-40, Jan. 2016. [Article \(CrossRef Link\)](#)
- [7] T. Tan, Z. Li, H. Liu, F. G Zanjani, Q. Quyang, Y. Tang, Z. Hu, and Q. Lu, "Optimize transfer learning for lung diseases in bronchoscopy using a new concept: Sequential fine-tuning," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 6, pp. 1-8, 2018. [Article \(CrossRef Link\)](#)
- [8] H. Brody, "Medical imaging," *Nature*, vol. 502, 2013. [Article \(CrossRef Link\)](#)
- [9] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85-117, 2015. [Article \(CrossRef Link\)](#)
- [10] L. Lu, Y. Zheng, G. Carneiro, and L. Yang, "Deep Learning and Convolutional Neural Networks for Medical Image Computing. Cham," *Switzerland: Springer*, 2017. [Article \(CrossRef Link\)](#)
- [11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436-444, May 2015. [Article \(CrossRef Link\)](#)
- [12] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," in *Proc. of IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998. [Article \(CrossRef Link\)](#)
- [13] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 89-40, 2017. [Article \(CrossRef Link\)](#)
- [14] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv:1409.1556*, 2015. [Article \(CorssRef Link\)](#)
- [15] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, and V. Vanhoucke, "Going deeper with convolutions," in *Proc. of 2015 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, pp. 1-9, 2015. [Article \(CrossRef Link\)](#)
- [16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proc. of 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, pp. 770-778, 2016. [Article \(CrossRef Link\)](#)
- [17] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *arXiv:1901.06032*, 2019. [Article \(CrossRef Link\)](#)
- [18] Q. Liu, N. Zhang, W. Yang, S. Wang, Z. Cui, X. Chen, and L. Chen, "A Review of Image Recognition with Deep Convolutional Neural Network," in *Proc. of International Conference on Intelligent Computing*, pp. 69-80, 2017. [Article \(CrossRef Link\)](#)
- [19] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, pp. 211-252, 2015. [Article \(CrossRef Link\)](#)
- [20] G. Liang and L. Zheng, "A transfer learning method with deep residual network for pediatric pneumonia diagnosis," *Computer Methods and Programs in Biomedicine*, vol. 187, 2019. [Article \(CrossRef Link\)](#)
- [21] S. S. Devi, S. A. Sheikh, A. Talukdar, and R. H. Laskar, "Malaria infected erythrocyte classification based on the histogram features using microscopic images of thin blood smear," *Indian Journal of Science and Technology*, vol. 9, no. 45, pp. 1-10, 2016. [Article \(CrossRef Link\)](#)
- [22] Z. Liang, A. Powell, I. Ersoy, M. Poostchi, K. Silamut, K. Palaniappan, P. Guo, M. A. Hossain, A. Sameer, R. J. Maude, J. X. Huang, S. Jaeger, and G. Thoma, "CNN-based image analysis for malaria diagnosis," in *Proc. of 2016 IEEE International Conference on Bioinformatics and Biomedicine(BIBM)*, pp. 493-496, 2016. [Article \(CrossRef Link\)](#)
- [23] D. Bibin, M. S. Nair, and P. Punitha, "Malaria parasite detection from peripheral blood smear images using deep belief networks," *IEEE Access*, vol. 5, pp. 9099-9108, 2017. [Article \(CrossRef Link\)](#)
- [24] G. P. Gopakumar, M. Swetha, G. S. Siva, and G. R. K. Sai Subrahmanyam, "Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner," *Journal of Biophotonics*, vol. 11, no. 3, Mar. 2018. [Article \(CrossRef Link\)](#)



- [25] Y. G. Gezahegn, Y. H. G. Medhin, E. A. Etsub, and G. N. G. Tekele, "Malaria Detection and Classification Using Machine Learning Algorithms," in *Proc. of International Conference on Information and Communication Technology for Development for Africa*, pp. 24-33, 2018. [Article \(CrossRef Link\)](#)
- [26] 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, Apr. 2018. [Article \(CrossRef Link\)](#)
- [27] A. Vijayalasshmi and K. Rajesh, "Deep learning approach to detect malaria from microscopic images," *Multimedia Tools and Applications*, vol. 79, pp. 15297-15317, 2019. [Article \(CrossRef Link\)](#)
- [28] P. A. Pattanaik, M. Mittal, and M. Z. Khan, "Unsupervised Deep Learning CAD Scheme for the Detection of Malaria in Blood Smear Microscopic Images," *IEEE Access*, vol. 8, pp. 94936-94946, 2020. [Article \(CrossRef Link\)](#)
- [29] P. Pattanaik, M. Mittal, M. Khan, and S. Panda, "Malaria detection using deep residual networks with mobile microscopy," *Journal of King Saud University - Computer and Information Sciences*, 2020. [Article \(CrossRef Link\)](#)
- [30] K. M. F. Fuhad, J. F. Tuba, M. R. A. Sarker, S. Momen, N. Mohammed, and T. Rahman, "Deep Learning Based Automatic Malaria Parasite Detection from Blood Smear and Its Smartphone Based Application," *Diagnostics*, vol. 10, no. 5, p. 329, 2020. [Article \(CrossRef Link\)](#)
- [31] G. Marques, D. Agarwal and I. de la Torre Díez, "Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network", *Applied Soft Computing*, vol. 96, p. 106691, 2020. [Article \(CrossRef Link\)](#)
- [32] C. Zhang, P. Patras, and H. Haddadi, "Deep Learning in Mobile and Wireless Networking: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2224-2287, 2019. [Article \(CrossRef Link\)](#)
- [33] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. of International Conference on Machine Learning*, 2019, pp. 6105-6114. [Article \(CrossRef Link\)](#)
- [34] S. Rajaraman, S. Jaeger, and S. K. Antani, "Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images," *PeerJ*, 2019. [Article \(CrossRef Link\)](#)
- [35] Q. Wei and R. L. Dunbrack, "The role of balanced training and testing data sets for binary classifiers in bioinformatics," *PLoS ONE*, vol. 8, no. 7, p. 67863, 2013. [Article \(CrossRef Link\)](#)
- [36] A. Torralba, and A. Efros. "Unbiased look at dataset bias," in *Proc. of Conference on Computer Vision and Pattern Recognition*, pp. 1521-1528, 2011. [Article \(CrossRef Link\)](#)
- [37] J. A. Nichols, H. W. Herbert Chan, and M. A. B. Baker, "Machine learning: Applications of artificial intelligence to imaging and diagnosis," *Biophysical Reviews*, vol. 11, no. 1, pp. 111-118, Feb. 2019. [Article \(CrossRef Link\)](#)
- [38] F. Chollet, Keras, 2015. [Online]. Available: <https://github.com/keras-team/keras>
- [39] M. Lin, Q. Chen, and S. Yan, "Network in network," *arXiv:1312.4400*, 2013. [Article \(CrossRef Link\)](#)
- [40] N. Islam, U. Saeed, R. Naz, J. Tanveer, K. Kumar, and A. A. Shaikh, "DeepDR: An image guideddiabetic retinopathy detection technique using attention-based deep learning scheme," in *Proc. of 2019 2<sup>nd</sup> International Conference on new Trends in Computing Sciences(ICTCS)*, pp. 1-6, 2019. [Article \(CrossRef Link\)](#)
- [41] ReLu, Deep AI, 2019. [Online]. Available: <https://deepai.org/machine-learning-glossary-and-terms/relu>
- [42] H. C. Shin, M. Orton, D. J. Collins, and M. Leach, "Oragn detection using deep learing," *Academic Press*, pp. 123-153, 2016. [Article \(CrossRef Link\)](#)
- [43] M. A. E. Muhammed, A. A. Ahmed, and T. A. Khalid, "Benchmark analysis of popular imagenet classification deep cnn architectures," in *Proc. of 2017 International Conference on Smart Technologies for Smart Nations*, pp. 902-907, 2017. [Article \(CrossRef Link\)](#)

- [44] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, J. Liang, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1299-1312, May 2016. [Article \(CrossRef Link\)](#)
- [45] E. M. Dogo, O. J. Afolabi, N. I. Nwulu, B. Twala, and C. O. Aigbavboa, "A Comparative Analysis of Gradient Descent-Based Optimization Algorithms on Convolutional Neural Networks," in *Proc. of 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems(CTEMS)*, pp. 92-99, 2018. [Article \(CrossRef Link\)](#)
- [46] S. Sun, Z. Cao, H. Zhu, and J. Zhao, "A Survey of Optimization Methods From a Machine Learning Perspective," *IEEE Transactions on Cybernetics*, vol. 50, no. 8, pp. 3668-3681, 2020. [Article \(CrossRef Link\)](#)
- [47] Loss Functions, Loss Functions: ML Glossary documentation, 2017. [Article \(CrossRef Link\)](#)
- [48] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. of International Conference for Learning Representations*, 2015. [Article \(CrossRef Link\)](#)
- [49] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, "On the importance of initialization and momentum in deep learning," in *Proc. of the 30<sup>th</sup> International Conference on Machine Learning*, vol. 28, no. 3, pp. 1139-1147, 2013. [Article \(CrossRef Link\)](#)
- [50] T. Tieleman and G. Hinton, "Lecture 6a. Overview of mini-batch gradient descent," *Neural Networks for Machine Learning*, 2012. [Article \(CrossRef Link\)](#)
- [51] P. Lakhani, D. L. Gray, C. R. Pett, P. Nagy, and G. Shih, "Hello world deep learning in medical imaging," *Journal of Digital Imaging*, vol. 31, pp. 283-289, 2018. [Article \(CrossRef Link\)](#)
- [52] F. J. P. Montalbo and A. A. Hernandez, "An Optimized Classification Model for Coffea Liberica Disease using Deep Convolutional Neural Networks," in *Proc. of 2020 16<sup>th</sup> IEEE International Colloquium on Signal Processing & Its Applications(CSPA)*, pp. 213-218, 2020. [Article \(CrossRef Link\)](#)
- [53] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *International Journal of Data Mining & Knowledge Management Process*, vol. 5, no. 2, p. 1, 2015. [Article \(CrossRef Link\)](#)
- [54] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," in *Proc. of 2017 IEEE International Conference on Computer Vision(ICCV)*, pp. 618-626, 2017. [Article \(CrossRef Link\)](#)
- [55] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning Deep Features for Discriminative Localization," in *Proc. of 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, pp. 2921-2929, 2016. [Article \(CrossRef Link\)](#)
- [56] J. Ferrando, J. L. Domingues, J. Torres, R. Garcia, D. Garcia, D. Garrido, and J. Cortada, "Improving Accuracy and Speeding Up Document Image Classification Through Parallel Systems," in *Proc. of International Conference on Computational Science*, pp. 387-400, 2020. [Article \(CrossRef Link\)](#)



**Francis Jesmar P. Montalbo** acquired both his BSc and MSc degrees in Information Technology from Batangas State University where he also serves as a full-time faculty researcher. He is currently a candidate for the degree of Doctor in Information Technology at the Technological Institute of the Philippines and a member of IEEE-CIS under the Philippine section. His research works revolve around deep learning and computer vision systems.



**Alvin S. Alon** received his Doctor of Engineering degree from the Technological Institute of the Philippines. He is currently an assistant professor and Research Head of the Digital Transformation Center of Science, Technology, Engineering, and Environmental Research (STEER) Hub in Batangas State University. His current research interests include Artificial Intelligence, Remote Sensing, and Data Analytics.