Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606

Volume 07 Issue 01
2024

Research Article
https://doi.org/10.56979/701/2024

Malaria Cell Classification through Exercising Deep Learning Algorithms

Sheharyar Muhammad[1], Muhammad Munwar Iqbal[1*], Saqib Majeed[2], Yasir Saleem[3], and Anees Tariq[1]

[1]Department of Computer Science, University of Engineering and Technology, Taxila, Pakistan.
[2]University institute of Information Technology, PMAS-Arid Agriculture University, Rawalpindi, Pakistan.
[3]Department of Computer Engineering, University of Engineering and Technology, Lahore, Pakistan.
*Corresponding Author: Muhammad Munwar Iqbal. Email: munwar.iq@uettaxila.edu.pk

Received: March 03, 2024 **Accepted:** May 06, 2024 **Published:** June 01, 2024

Abstract: Malaria, a life-threatening disease caused by Plasmodium parasites transmitted through the bites of infected Anopheles mosquitoes, continues to pose a significant global health challenge. This research introduces an advanced automated classification system for malaria detection in blood smear images, integrating cutting-edge machine learning architectures, DenseNet and EfficientNet, to enhance diagnostic accuracy and efficiency. Our approach begins with the original blood smear image, which is processed through Faster R-CNN for detection, followed by ROBOflow for precise segmentation. The classification phase utilizes the strengths of DenseNet and EfficientNet models, coupled with a ReLU activation function and further refinement through ResNet architecture, to achieve superior feature extraction and classification performance. This comprehensive model aims to minimize manual intervention and significantly improve the speed and accuracy of malaria diagnosis. The proposed system not only represents a significant advancement in the field of medical diagnostics but also illustrates the potential of integrating sophisticated machine learning techniques to tackle complex global health challenges.

Keywords: Malaria Cell Classification; Efficient net; Deep Learning; CNN Features Extraction; YoloV5.

1. Introduction

Malaria has long been a dire global health issue, causing significant morbidity and mortality across various climates and geographical regions. The impact of this disease has been particularly severe on children, with a staggering 67% of malaria casualties being children under the age of five as of 2019 [1]. The World Health Organization reported an escalation in malaria cases from 210 million in 2016 to 228 million in 2018, with an estimated death toll reaching approximately 405,000 [2]. These figures underscore the pervasive risk of malaria, which does not discriminate between temperate and tropical zones. In the United States alone, the Centers for Disease Control and Prevention notes around 1700 annual cases [2]. While children are disproportionately affected—constituting 67% of malaria fatalities worldwide in 2018—the disease also poses a severe threat to adults [1]. Regionally, Africa bears the brunt of the burden, with 93% of cases and 94% of deaths [3]. The causative agent of malaria, Plasmodium, is transmitted via mosquito bites. Six Plasmodium species are known to infect humans, with P. Falciparum and P. Vivax identified as the most threatening. Other species, such as P. Curtisi, P. Wallikeri, P. Malariae, and the newly recognized P. Knowlesi, contribute to the disease's prevalence but are less commonly implicated in severe cases [4]. This taxonomy of the malaria pathogen and its human impact are visually summarized in Figure 1.

Anopheles mosquitoes are the primary vectors for malaria, introducing the disease's causative sporozoites into the human bloodstream. These sporozoites navigate through the circulatory system to the liver, where they invade and multiply within the hepatocytes, the liver cells responsible for detoxification and protein synthesis. During this liver stage, the infected host typically does not experience any symptoms [5].

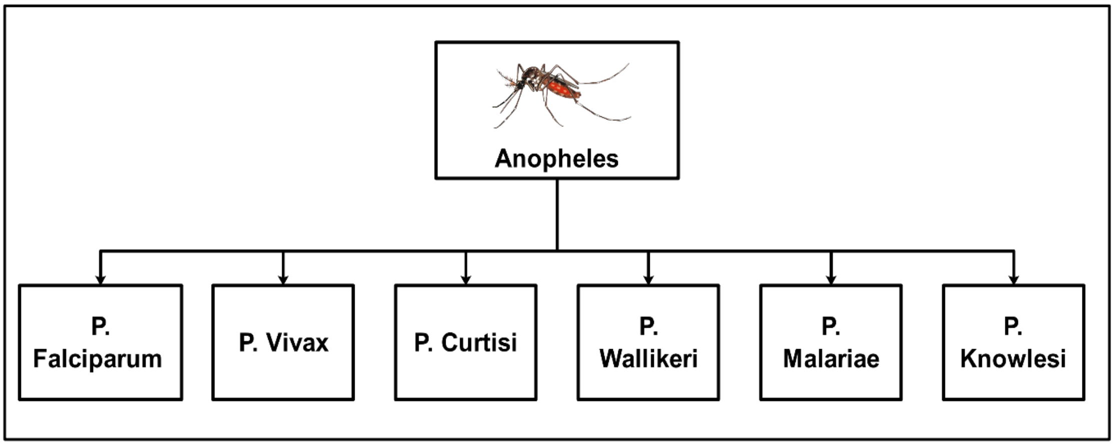

Figure 1. Types of malaria parasites transmitted by the mosquito [4]

Malaria transmission is predominantly linked to the Anopheles mosquito's bite. Globally, the Anopheles genus comprises over 350 distinct species, of which approximately 30 are identified as carriers of malaria. The typical adult Anopheles has a muted coloration that blends into darker environments and features three primary body segments: the head, thorax, and abdomen [7]. Female Anopheles mosquitoes, the gender responsible for blood feeding, lay their eggs in a variety of aquatic habitats. These environments range from clear freshwater to brackish water bodies, and include both vegetated and open areas, shaded or sunlit locales, temporary puddles, water-filled depressions, swamps in forests, and other locales with stagnant water. The resilience of Anopheles eggs allows them to withstand cool temperatures, although they are susceptible to freezing conditions. The global challenge posed by malaria, characterized by the meticulous examination of blood samples under microscopes, underscores the urgency for more sophisticated diagnostic solutions. Traditional microscopy, while widespread, is constrained by its dependence on human expertise, leading to potential delays and inaccuracies in diagnosis. Early detection of malaria is crucial for effective treatment, necessitating a diagnostic system that combines speed, accuracy, and precision. Challenges such as image quality, including issues with illumination and noise, alongside the need for feature optimization to prevent model complexity and inefficiency, further highlight the need for advancement in this area. The primary contributions of this research are:

The research employs Roboflow, a sophisticated tool, for the extraction of significant attributes from the data images. This step is essential for garnering the rich and profound features necessary for effective model training. In an effort to amalgamate and capitalize on the strengths of multiple deep convolutional neural networks, the research adopts a nuanced transfer learning approach.

- This strategy is not merely about combining depth attribute details from various deep models but is focused on synthesizing these details into a compact and efficient feature vector set. This fusion is anticipated to bolster the model's ability to discern and categorize data based on its labels, thereby enhancing the overall performance and accuracy of the automated system.
- This paper presents a novel automated classification model for malaria-infected blood smear images that stands at the confluence of advanced machine learning and deep learning technologies. By integrating DenseNet and EfficientNet, our model not only captures the nuanced features of blood smear images but also improves upon the diagnostic accuracy and efficiency of existing methods. The model pipeline is designed to process the original image through a series of sophisticated algorithms: starting with Faster R-CNN for initial detection, followed by ROBOflow for image segmentation, and culminating in a classification phase that leverages the combined power of DenseNet and EfficientNet models.

2. Literature Review

Worldwide, malaria has caused millions of diseased cases and thousands of deaths over the years in both hot and cold climates, making it a major health concern. Malaria must be diagnosed in its early stages and the parasites that cause it must be identified in order for the patient to have a better chance of recovery. Throughout human history, malaria has posed a serious and intimidating threat to the human health. Over the years, it has been responsible for thousands of deaths in all climates. It is a global health concern, with an estimate of 228 million cases and 400 thousand deaths. When it comes to malaria, kids are more susceptible than adults, as nearly 67 percent of kids will have contracted it by the year 2019. Several preprocessing approaches are used for malaria picture refining in various literature works, notably for contrast enhancement (Masud, Alhumyani et al. 2020), lighting correction, image restoration (Green, Berman et al. 1988), noise removal and region intensification among many others.

In the work of (Jain, Chauhan et al. 2020) images were transformed to the HSV color channel and then stretched in terms of contrast. Adaptive normalization and eigenvalue decomposition were used by Liang, Powell et al. (2016) to enhance brightness and contrast while also reducing the amount of dark detail in the images. Grayscale conversion was conducted by (Halim, Bretschneider et al. 2006) using an automated threshold obtained by studying the brightness for different colours in the image. The resultant image is generated using the Cartesian coordinate system. Grey level histogram and erosion were used by (Singla and Srivastava 2020) to eliminate background noise. The data is then converted to an HSL channel and color segmentation is used to recover the ring and parasitic regions. Preprocessed images were used by (Cinar and Yildirim 2020) which included a Gaussian and a median filter. The deep network process where implemented for feature extraction. Images that have been enhanced with both filters will produce a wide range of results from various networks. When using Gaussian filtered images, DenseNet201 produced the highest classification accuracy at 97.83%.

To segment smear images infected with three types of malarial species, (Jusman, Pusparini et al. 2021) utilized modified thresholding and watershed methods. Thresholding methods from the global and Otsu schools are combined to create a new thresholding model. Morphological opening and closing operations are also used for segmented mask refinement. Segmentation accuracy of 100 percent is achieved by the thresholding model and 90 percent is obtained by watershed method.

Table 1. Literature Review

Year	References	Techniques	Datasets	Results Accuracy
2021	(Montalbo et al.)	Classification is done using KNN Method and linear LDA are applied.	15k blood images are used.	97.00 percent
2021	(Baihaqi, Widiawati, Sabil, & Wati)	Custom CNNs Models such as DRNet and LeNet are applied.	27k parasitized and uninfected blood images are used.	97.00 percent
2021	(Montalbo & Alon)	Grad-CAM interpretation is used for extracting features, Sigmoid classification and EfficientNetB0 are applied.	27k parasitized and uninfected blood images are used.	94.70 percent
2021	33(Swastika, Kristianti, & Widodo)	Custom CNNs Models, ResNet-50 and VGG-16 are applied.	27k parasitized and uninfected blood images are used.	93.10 percent

2021	(Kittichai et al.)	DenseNet-201 and DarkNet-19 is used.	10,548 blood images are used	99.00 percent
2021	(Jusman, Pusparini, Chamim, & Kanafiah)	HSV color spacing, Global and Otsu, watershed segmentation using thresholding.	60 blood images are used	90.00 percent
2021	(Masud et al.)	Custom CNNs Models with normalization, ReLU, and dropout layers with optimization	27k parasitized and uninfected blood images are used.	93.00 percent
2021	(Sinha, Srivastava, Dhiman, Akhilan, & Mishra)	Custom CNNs Models and ResNet.	27k parasitized and uninfected blood images are used..	97.00 percent
2020	(Singla & Srivastava)	Custom CNNs Models with SCM system.	1000 blood images are used	97.00 percent
2020	(Vijayalakshmi)	Vgg-19 and Vgg-16 with SVM classifier.	2550 parasitized and uninfected blood images are used..	93.10 percent
2020	(Sriporn, Tsai, Tsai, & Wang)	Xception, RMSProp, and Nadam optimizers	7k parasitized and uninfected blood images are used.	99.20 percent
2020	(Pattanaik, Mittal, Khan, & Panda)	Deep residual Network (MM-ResNet).	1182 blood images are used.	98.00 percent
2020	(Quan, Wang, & Liu)	ResNet and DenseNet.	27k parasitized and uninfected blood images are used.	97.40 percent
2019	(Rahman et al.)	Custom CCN MODEL	27k parasitized and uninfected blood images are used.	97.70 percent
2019	(Suriya, Chandran, & Sumithra)	DEEP CCN MODEL with custom layers	26k blood images are used.	98.91 percent
2019	(Reddy & Juliet)	Fused features of VGG, Inception, and ResNet	27k parasitized and uninfected blood images are used.	95.40 percent
2018	(Devi, Roy, Singha, Sheikh, & Laskar)	Classification using following methods SVM, KNN, NB.	1300 blood images are used.	98.50 percent

3. Proposed Methodology

The proposed research introduces a fault diagnosis framework for the diagnostic of non-stationary vibratory signals generated by the varying conditions of the bearings.

Figure 2 depicts all the phases of the proposed diagnostic framework. The proposed research comprises steps.

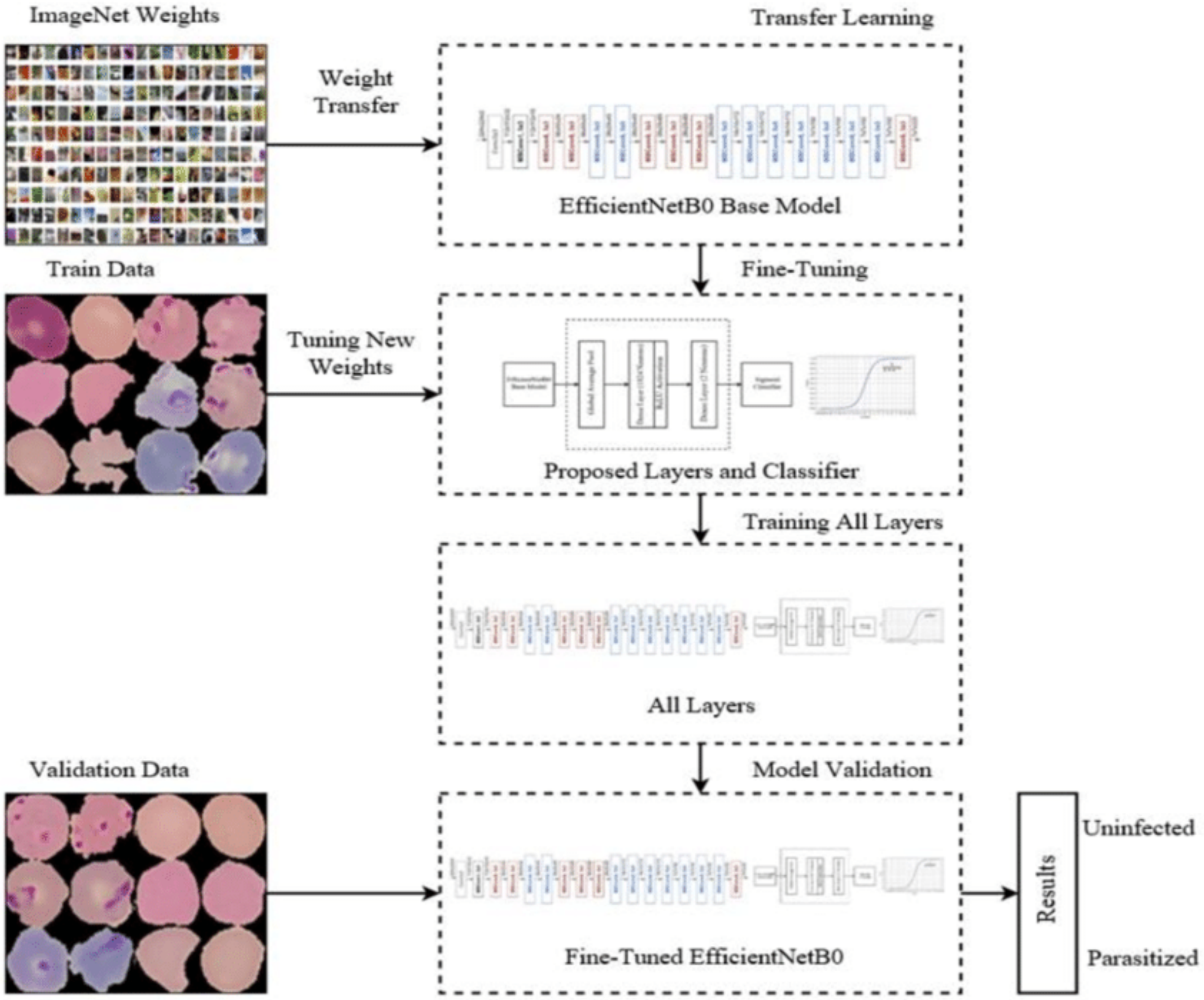

Figure 2. Methodology Proposed

Datasets: The dataset encompasses a total of 27,558 microscopic images of blood cells, neatly divided into two distinct categories: 'parasitized' and 'uninfected.' The 'parasitized' group consists of images marked by the presence of malaria-inducing Plasmodium species, whereas the 'uninfected' group comprises images from healthy individuals. Figure 3 below presents a selection of these images as illustrative examples.

Data Augmentation: Data augmentation plays a pivotal role in enriching the dataset of blood cell images, crucial for the detection of malaria. Techniques such as rotation, scaling, and horizontal flipping are applied to mimic various imaging conditions, thus enhancing the model's ability to generalize from the augmented data.

DenseNet Architecture: Our approach employs DenseNet-201 for its unique architecture, characterized by dense connectivity between layers. This design promotes feature reuse and significantly reduces the vanishing-gradient issue, facilitating effective feature extraction from blood smear images.

Dense Block and Layers: DenseNet-201 consists of dense blocks that contain layers directly connected to every other layer in a feed-forward fashion. Transition layers between these dense blocks compress and consolidate the feature maps, optimizing the flow of information through the network and enhancing the model's ability to accurately classify malaria-infected cells.

Transfer Learning: To expedite the learning process and enhance the model's performance, transfer learning is applied. By leveraging pre-trained models on extensive datasets, our model adapts to the specific task of malaria detection, benefiting from previously learned features.

Training and Evaluation: Use the extracted features and their corresponding labels (parasite or normal) to train a classifier, such as a support vector machine (SVM), random forest, or a deep neural network, to learn the relationship between the features and the class labels. Evaluate the trained classifier on a separate validation or test set to assess its performance.

Interpretation and Analysis: Once the classifier is trained and evaluated, analyze the importance of different features in distinguishing between parasite and normal cells. This analysis can provide insights into the visual characteristics or patterns that contribute to the classification decision.

.

4. Results

The exploration into the classification of malaria-infected cells via deep learning models, specifically using the EfficientNet architecture, has yielded highly accurate outcomes. The ensuing discussion delineates the ramifications of these results and their potential to enhance diagnostic processes. The validation of the automated system is demonstrated by an accuracy of 0.99, with validation accuracy of 0.97 alongside a minimal loss of 0.1206. These figures were derived at the culmination of the model's training, which spanned 15 epochs. The progression of training and validation accuracy over these epochs further reinforces the model's stability and consistency.

Table 2. Precision, Recall and F1-score using EfficientNet

Metrics	Precision	Recall	F1-score	Support
Parasitized	0.97	0.97	0.97	1091
Uninfected	0.97	0.97	0.97	1115
Accuracy			0.97	2206
Macro avg	0.97	0.97	0.97	2206
Weighted avg	0.97	0.97	0.97	2206

The empirical evidence garnered from the application of the EfficientNet model indicates a remarkable precision in classifying cells as parasitized or uninfected. These findings are substantiated by the metrics, with both categories achieving a precision, recall, and f1-score of 0.97, as supported by a considerable number of samples. These statistics affirm the model's robustness and reliability in cell classification tasks.

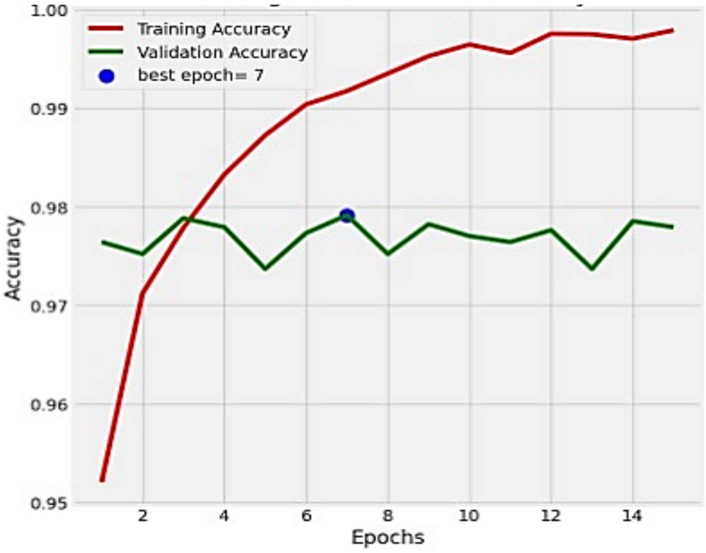

Figure 3. Training and Validation Accuracy

The analysis using the EfficientNet model, the classification of blood cell images into 'parasitized' and 'uninfected' categories demonstrated substantial accuracy. The model discerned between these categories with a notable high true positive rate for both, as indicated by the confusion matrix with 1029 true positive classifications for 'parasitized' and 1094 for 'uninfected' cells, and relatively few false negatives and false positives respectively.

Figure 4. Confusion Matrix Using EfficientNet

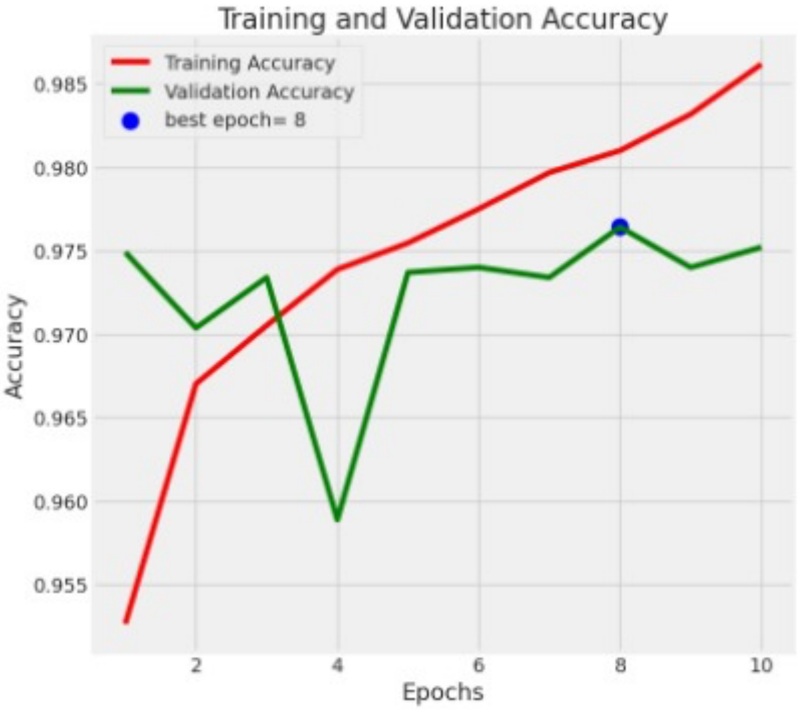

Figure 5. Training and Validation Accuracy

During the training phase, the model's performance improved consistently as seen in the training and validation accuracy graph. It peaked at the eighth epoch, suggesting that the model had effectively learned the distinguishing features of the blood smear images by this stage of the process. The final epoch of the training, the tenth, concluded with the model achieving an impressive accuracy of 98.62% on the training set and 97.52% on the validation set. These results are coupled with a loss of 0.677 for the training set and a slightly higher loss of 0.1402 for the validation set, showcasing the model's ability to generalize well from the training data to unseen data.

5. Conclusion

The study invertigates the deep learning techniques have shown great promise in malaria cell classification and segmentation. The application of EfficientNet within the realm of malaria diagnosis is a testament to the potential of deep learning techniques in medical imaging. The high accuracy 0.99 using EfficientNet and 0.98 using DenseNet, and precision of the model lay the groundwork for its adoption in clinical environments, contributing to the broader effort to combat malaria worldwide. Further investigation into alternative deep learning architectures could enhance diagnostic accuracy and efficiency in malaria detection. Efforts will be made to improve the models' generalizability across different datasets, particularly those representing a wide range of imaging conditions and geographic diversity.

References

1. Baihaqi, W. M., Widiawati, C. R. A., Sabil, D. P., & Wati, A. (2021). Analisis Gambar Sel Darah Berbasis Convolution Neural Network Untuk Mendiagnosis Penyakit Demam Berdarah. CogITo Smart Journal, 7(1), 148-159.
2. Devi, S. S., Roy, A., Singha, J., Sheikh, S. A., & Laskar, R. H. (2018). Malaria infected erythrocyte classification based on a hybrid classifier using microscopic images of thin blood smear. Multimedia Tools and Applications, 77(1), 631-660.
3. Jusman, Y., Pusparini, A., Chamim, A. N. N., & Kanafiah, S. N. A. M. (2021). Comparison of Malaria Parasite Image Segmentation Algorithm Using Thresholding and Watershed Method. Paper presented at the Journal of Physics: Conference Series.
4. Kittichai, V., Kaewthamasorn, M., Thanee, S., Jomtarak, R., Klanboot, K., Naing, K. M., Boonsang, S. (2021). Classification for Avian Malaria Parasite Plasmodium Gallinaceum Blood Stages by Using Deep Convolutional Neural Networks.
5. Masud, M., Alhumyani, H., Alshamrani, S. S., Cheikhrouhou, O., Ibrahim, S., Muhammad, G., Shorfuzzaman, M. (2020). Leveraging deep learning techniques for malaria parasite detection using mobile application. Wireless Communications and Mobile Computing, 2020.
6. Montalbo, F. J. P., & Alon, A. S. (2021). Empirical Analysis of a Fine-Tuned Deep Convolutional Model in Classifying and Detecting Malaria Parasites from Blood Smears. KSII Transactions on Internet and Information Systems (TIIS), 15(1), 147-165.
7. Pattanaik, P., Mittal, M., Khan, M. Z., & Panda, S. (2020). Malaria detection using deep residual networks with mobile microscopy. Journal of King Saud University-Computer and Information Sciences.
8. Quan, Q., Wang, J., & Liu, L. (2020). An Effective Convolutional Neural Network for Classifying Red Blood Cells in Malaria Diseases. Interdisciplinary Sciences: Computational Life Sciences, 12(2).
9. Rahman, A., Zunair, H., Rahman, M. S., Yuki, J. Q., Biswas, S., Alam, M. A., Mahdy, M. (2019). Improving malaria parasite detection from red blood cell using deep convolutional neural networks. ArXiv preprint arXiv: 1907.10418.
10. Reddy, A. S. B., & Juliet, D. S. (2019). Transfer learning with ResNet-50 for malaria cell-image classification. Paper presented at the 2019 International Conference on Communication and Signal Processing (ICCSP).
11. Singla, N., & Srivastava, V. (2020). Deep learning enabled multi-wavelength spatial coherence microscope for the classification of malaria-infected stages with limited labelled data size. Optics & Laser Technology, 130, 106335.
12. Sinha, S., Srivastava, U., Dhiman, V., Akhilan, P., & Mishra, S. (2021). Performance assessment of Deep Learning procedures on Malaria dataset. Journal of Robotics and Control (JRC), 2(1), 12-18.
13. Sriporn, K., Tsai, C.-F., Tsai, C.-E., & Wang, P. (2020). Analyzing Malaria Disease Using Effective Deep Learning Approach. Diagnostics, 10(10), 744.
14. Suriya, M., Chandran, V., & Sumithra, M. (2019). Enhanced deep convolutional neural network for malarial parasite classification. International Journal of Computers and Applications, 1-10.
15. Swastika, W., Kristianti, G., & Widodo, R. (2021). Effective preprocessed thin blood smear images to improve malaria parasite detection using deep learning. Paper presented at the Journal of Physics: Conference Series.
16. Vijayalakshmi, A. (2020). Deep learning approach to detect malaria from microscopic images. Multimedia Tools and Applications, 79(21), 15297-15317.