Asim et al. LGURJCSIT 2023

LGU Research Journal of Computer Science & IT

SSN: 2521-0122 (Online)
ISSN: 2519-7991 (Print)

doi: 10.54692/lgurjcsit.2023.0702473

Vol. 7 Issue 2, April – June 2023

Classification of Microscopic Malaria Parasitized Images Using Deep Learning Feature Fusion

Muhammad Asim[1*], Shahid Zaman[2], Abdul Razzaq[2], Salman Qadri[2], Zahid Hussain[3], Zaid Sarfraz[2]
[1]Department of Computer Science, Khawaja Fareed University of Engineering and Technology, Rahim Yar Khan, Pakistan
[1]DD. Research at NCBA&E Sub Campus, Multan, Pakistan
[2] Department of Computer Science, Muhammad Nawaz Sharif University of Agriculture, Multan, Pakistan
[3]NFC Institute of Engineering & Technology,Multan, Pakistan

Email: masimrajwana@ncbae.edu.pk

ABSTRACT:

An infectious disease that causes a chronic and potentially life-threatening infection caused by microorganisms of the Plasmodium class is malaria or malarial disease. It is critical to detect the presence of Malaria parasites as early as possible to ensure that antimalarial treatment is adequate to cure the particular type of Plasmodium. This is to reduce death rates and to focus on various infections in the event of an adverse outcome. This study aimed to develop an artificial intelligence approach capable of separating parasitized erythrocytes from normal basophilic erythrocytes as well as platelets overlying the red blood cells to overcome the high cost of Malaria diagnostic equipment. The tone and texture characteristics of erythrocyte images were extracted using histogram thresholds and watershed methods and then fused with Squeeze Net and ShuffleNet algorithms. The measures included planning, preparing, approving, and testing Deep Convolution Neural Network Segmentation without preparation using a graphic processor unit. A total of 96% accuracy and specificity was obtained for the position of malaria in red blood cells based on the overall results. It has been demonstrated that deep learning techniques can be effective in clinical pathology. This provides new directions for development as well as increasing awareness of researchers in the medical field. The proposed methodology and application of fusion techniques on the images produce good results for the classification and digital identification of the disease.

KEYWORDS: Deep Feature; Fusion Technique; Malaria; Decision Making, Digital Identification

1. INTRODUCTION

Malaria is the world's most common vector-borne illness, with over 2 billion patients diagnosed yearly. Malaria is responsible for 0.6 million fatalities among pregnant women and children each year. The globe is now falling short of the WHO Global Technical Strategy (GTS) for Malaria 2016–2030 objectives for 2020. The goal for 2020 is to cut malaria deaths and illnesses by 40%. Malaria cases have been steadily increasing from 2015 to 2018. Malaria cases worldwide increased by 214000000, 217000000 and 219 - 000000 in 2015, 2016, and 2017, respectively. [17]. Africa, particularly Nigeria, presently accounts for over 90% of worldwide malaria-related infections, costing Africa's economy $12 billion yearly in direct losses and 10.3 percent of annual GDP growth [27]. Malaria is Pakistan's top cause of illness and mortality in different areas of the globe. It is one of the six urgent infectious diseases that threaten millions. Pakistan has now been combined alongside Afghanistan, Somalia, Sudan, and Yemen, accounting for more than 95 percent of the

regional burden of malaria. This is with an estimated 1,000,000 cases and 3 million confirmed cases reported annually. Dengue fever in Pakistan has always followed a seasonal and unpredictable pattern. Heavy transmission in regions or districts along the border area in Sindh and Balochistan, as well as in neighboring regions with Iran and Afghanistan. The top infected provinces/regions, according to the stated: "Annual Parasite Incidence" (API), are FATA, Balochistan, and Khyber Pakhtunkhwa. Rural Sindh districts such as Thatta, Mirpur Khas, Khairpur, and Tharparkar are the province's primary endemic zones. According to the most recent "Stratification," 66 regions and agencies have been classified as being in the "high endemicity stratum," which means their API is greater than 5 per 1000 people. Punjab had the lowest endemicity in the previous two decades, with an API of >1/1000. Punjab has been the lowest endemic province [17]. In 2018, an estimated $2.7 billion was invested and offered for malaria research due to its large global health impact. Technology presents a unique opportunity to develop new methodologies to rebalance the scales. However, technology may redefine what is conceivable and give new possibilities in the fight to handle this fatal infection [1]. The information on Artificial Intelligence, particularly Machine learning in Computer Science, can be applied to Malaria research. If this is accomplished, malaria will be contained effectively within endemic areas [2]. We can identify or separate an infected area on an image using deep learning algorithms. This could assist in identifying diseases in a more straightforward manner than using CNN. CNN utilizes a Neural Network to mimic intelligence [3]. As of late, Deep Learning has picked up consideration since it permits the shirking of the weight of precise devoted division and highlight extraction techniques before the order [7].

As part of deep learning, multilayer Neural Networks (NN) are employed to construct representations of input images to produce specific outcomes, such as division or grouping without manual features. We planned, prepared, approved, and tested two Neural Networks: SqueezeNet and ShuffleNet. The previous was expected to fragment the entire slide PB spreads, and the last to order singular erythrocytes concerning whether they are parasitized. As a result, the principal objective of this work is to plan an NN-based framework for the programmed detection of malaria by linking the two organizations with a transitional advance that makes the yield of the first company immediately applicable to the second. Its Digital Pathology System (DPS) is formulated to assist clinical laboratories in making faster, less abstract choices toward clinical diagnosis. The tail pipeline introduced in this work shows a worldwide division exactness for RBCs of 98.3% and an explicitness for Malarial sickness. When healthcare professionals use microscopes to identify malarial cells, a study such as this will be helpful. To eradicate this endemic disease by 2030, the World Health Organization (WHO) plays an instrumental part. The WHO Global Malaria Program (GMP) is responsible for planning WHO's worldwide activities to control and eliminate malaria. Pakistan is also trying to eliminate this fatal disease through financial and men power resources for this purpose. Here are some costs details that Pakistan has planned and executed since 2013. (These details are mentioned from DMS Pakistan) are listed below:

Table 1: Malaria Funds Details in Pakistan

	Details of Fund Received from the Global Fund					
	SSF Round 10					
	2015-16		2014-2015		2013-2014	
	Local Currency Pakistani Rupees	Equivalent US $	Local Currency Pakistani Rupees	Equivalent US $	Local Currency Pakistani Rupees	Equivalent US $
Receipts						
Grant Received	489,253,826	4,719,081	100,701,807	1,001,012	241,631,023	2,359,938
Grant Received In Kind	294,746,209	2,834,098	534,333,217	5,263,108	121,301,523	1,143,382
Total	784,000,035	7,553,179	635,035,024	6,264,120	362,932,546	3,503,320

This study aims to help doctors find malaria in the world's most afflicted regions. Companies can utilize our created model to produce equipment for detecting cells affected by malaria. Due to the hardware needs for trained models, equipment costs are relatively expensive. However, because our model is trained on straightforward hardware, developing expensive hardware and commercializing it is unnecessary. Instead of classifying malaria into two classes, this paradigm can be expanded.

2. RELATED WORK

Numerous studies have shown that malaria remains a key public health problem worldwide. However, concerted efforts have been made to control malaria spread and transmission within and between societies. As a result of the work carried out by information technology, which continues to assist in the medical field, thousands of inventions are improving health issues in a significant way as it has been able to overcome doctors' errors by automating several manual tasks. People are becoming more aware of the importance of IT in all health fields as time progresses. Many people have contributed to this endemic issue with their capabilities and expertise and produced remarkable results that are still being replicated today. However, there are many gaps to be filled [9]. Models in deep learning apply a progression of successive layers with nonlinear preparation units that can discover various levelled include relations inside the crude picture information. (low-level) features that are preoccupied with high-level features contribute to nonlinear decision-making, learning complexity, and end-to-end extraction and classification [3]. Also, deep learning models show better execution than portion-based calculations; for example, they Support Vector Machines (SVMs) in enormous volumes of information and computational assets, making them incredibly versatile [4]. The specialists attempted to build up a portable application that takes photographs of blood tests to examine intestinal sickness immediately. Mobile phone applications enable us to analyze plasma tests without requiring professional magnifying instruments [5]. In 2017, a man worked with 10,000 photographs to imagine deep learning enactments. We can explore the central focuses by imagining the distinctive features and activations of a fundamental, modified profound learning model. We apply it to the trial of Malaria cell grouping, and accordingly, the model achieves 98.61% order exactness under a lower multifaceted design [7]. Overall, the fusion feature technique on the leaves has been discussed with an accuracy of 93% with better classification and identification results [22]. Virtualization Machines have been used with fuzzy techniques in the inference system [23]. Plant disease classification on digital pictures has been used with better results of CNN architecture[24]. Diabetes identification with fuzzy logic has been discussed with some good results [25]. Natural Languages Processing Techniques and Machine Learning Techniques have been applied in the study [26].

Most analysts employed transfer learning models to reduce planning time at the expense of execution due to their TL models, which may be helpful when larger datasets are not accessible.[8]. They will, by and large, perform incapably on data they are not already configured. In addition, we are also obligated to the extent that the framework plan requires. We can't change the prearranged framework [14]. Another project done in the same year is segmenting images and classifying them. It explains in detail how to segment a cell using limit controls. It also explains how to put the cell in focus through a series of pictures to obtain the preparation dataset. This overcomes image noise by sliding window cutting [10].

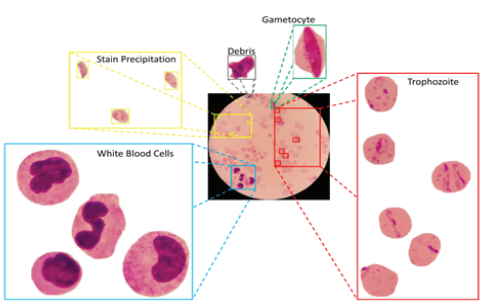

Figure 1: Microscopic Visualized Cell Segmentation of Malarial Infected Cell

By adapting Transfer Learning to Malaria image recognition, we achieve a precision of 94.58%[12]. Specifically, JPEG 2000 decompressed pictures could reach 95% precision after 30 to-1 compression. For grouping results based on the common MNIST dataset, handwritten digits were much easier to arrange utilizing decompressed pictures. Using just one single-bit plane was still possible with about 90%

accuracy [14]. "Approximately 3.2 billion people live in malaria-prone regions." The poorest countries' populations are the most vulnerable" (p.1). The World Health Organization (WHO) confirmed this in 2015. Malaria kills about 800 children per day in the United States. Africa is a continent with diverse children (every 2 minutes, there is one child). Because of the difficulty and poverty, their homes are not adequately constructed. There are places where mosquitoes can get in while the house is being built. The wealthiest individuals live in luxurious homes. Their dwellings are beautifully constructed [19].

3. MATERIAL AND METHOD

The proposed method includes deep learning features, which are used in the classification between microscopic images of normal cells and infected cells, as shown in Figure 2.

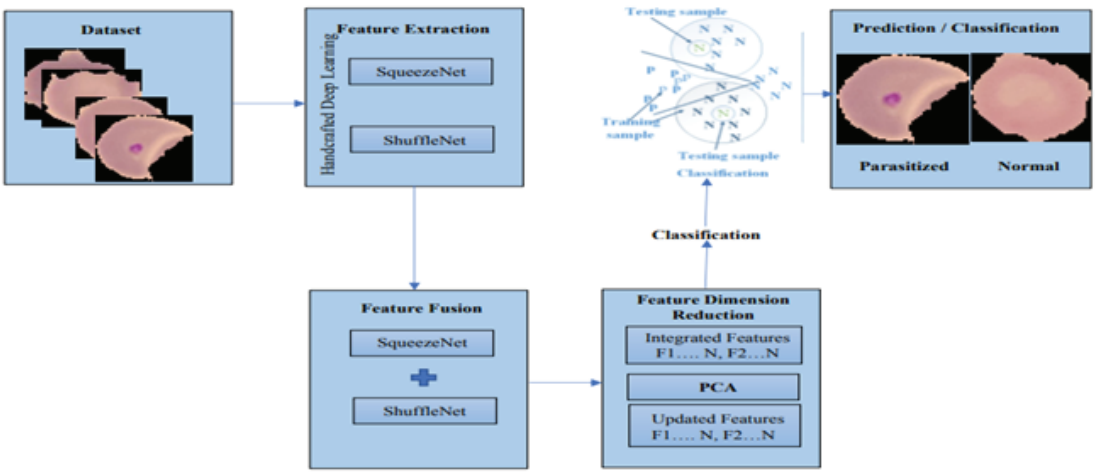

Figure 2: Proposed Model with Deep Learning Features

3.1. Proposed Scheme

The basic fundamental of malarial classification is divided into feature extraction, feature fusion, feature selection and classification. The proposed Scheme has the following steps:

3.1.1. Feature extraction

Extracting the vital information from the image dataset created on various image features like colour, texture, shape and edge features. Deep learning models were used to extract information from dataset images.

3.1.2. Feature fusion

Feature fusion technique to integrate two or more feature vectors that extract from a training dataset without losing data. Feature fusion can use multi-feature of the image

3.1.3. Feature selection

The method is applied to solve the most discriminant and intersecting feature that represents proper information about data from the original feature pool.

3.1.4. Principal Component Analysis (PCA)

reduces the dimensionality of such datasets, enhancing interpretability but minimizing the loss of information simultaneously.

3.1.5. Classification/prediction

calculate the similarity between the query image and training dataset image features constructed on a machine learning algorithm that checked, and the feature closer to the test sample features the matching image of that classification.

3.1.6. Deep Learning Features

In particular, as opposed to just applying group convolution on 3x3 filters, as shown below.

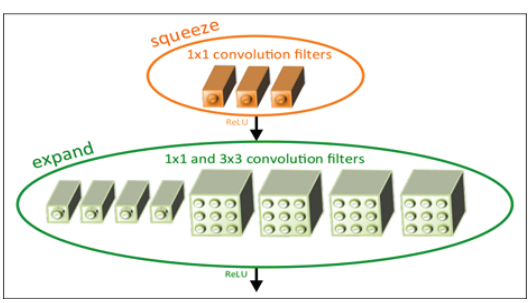

Figure 3: Expanded View of SqueezeNet

The network ShuffleNet and SqueezeNet has the input size of images as 299x299x3 and 224x224x3, consecutively. The deep learning model gives us features after learning the previous features. It provides us with the images' mixed shape, colour and texture features. Table 2 represent the deep feature dimension.

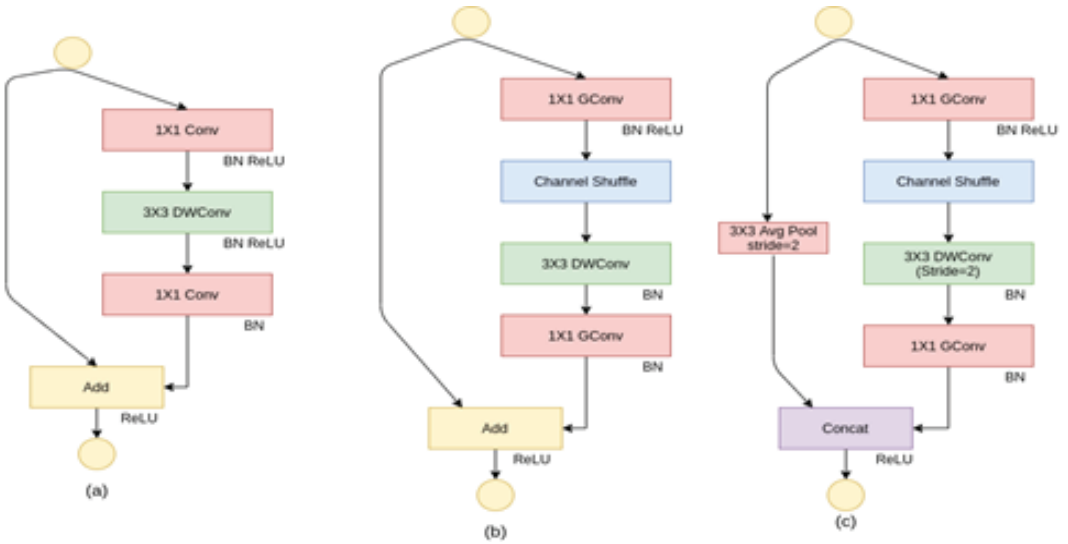

Figure 4: ShuffleNet Units [21]

In the proposed work, ShuffleNet is used, a convolutional neural network with a total of 50 layers consisting of convolutional, pooling, fully connected and SoftMax layers. The first step is to represent the structure of the ShuffleNet network in the table. This is based on an extract of the predictions layer. In the next step, SqueezeNet is used, a convolutional neural network consisting of 18 layers and consisting of convolutional, pooling, fully connected and SoftMax layers. According to the experience feature extracted from layer 'f18'. The dimension of this feature is 1 x 1000 on each image. Features extracted from each augmented image are stored in a database for future use.

Table 2: Represent the Features Dimension

Features Extractor	Deep learning feature	Feature layer	Features	Images			Total features
				Normal	Pneumonia	Total	
ShuffleNet	✓	predictions	1 x 1000	13780	13780	27560	27560 x 1000
SqueezeNet	✓	pool10	1 x 544	13780	13780	27560	27560 x 544

3.2. Performance Evaluation Metrics

These metrics are used in proposed models to evaluate their performance, such as accuracy, sensitivity, specificity, recall, precision and F1score. Formulas to evaluate the performance of the purposed model are listed below.

Accuracy = (TP+TN)/(TP+FN+TN+FP) (1)

Sensitivity = TP/(TP+FN) (2)

Specificity = TN/(TN+FP) (3)

Precision = TP/(TP+FP) (4)

Recall = TP/(TP+FN) (5)

F1score = 2x ((precisionxrecall)/ (precision+recall))

Here we used the cross-validation technique, which is a widely used technique. It is an evaluation technique that can generalize models for modelling and prediction.

3.3. Dataset for Proposed Methodology

Our dataset comprises 27,560 pictures that are organized in 2 envelopes (parasitized, uninfected)—portioned cells from the slender blood smear slide pictures from the Malaria Screener research movement. Giemsa-recoloured flimsy blood smear slides from 150 P. falciparum-contaminated and 50 solid patients were gathered and shot at Chittagong Medical College Hospital, Bangladesh. A specialist slide peruses at the Mahidol-Oxford Tropical Medicine Research U nit in Bangkok, Thailand physically explained the pictures.

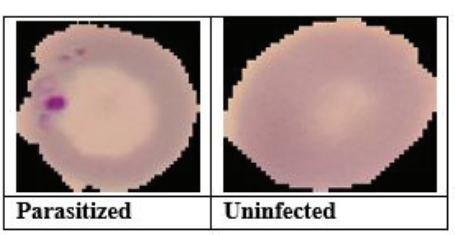

Figure 5: Classes of Dataset

3.4. Experimental Results

Our work evaluates the performance measure of the proposed method, which is the requirement of the algorithms we have implemented. For this purpose, we resized our data according to the target model. We have used two Deep Learning models ShuffleNet and SqueezeNet. While ShuffleNet accepts images of 299x299x3 dimensions, SqueezeNet accepts images of 224x224x3, but in LBP, images are set to 277x277x1. We tested a model by using 27560 front blood cell images. The binary classes contain 13780 malaria and 13780 standard images. We did 3 experiments in table 3 based on our acquired data to get different observations.

Table 3: Accuracy of different experiments

Experiment	Algorithm	Prediction Speed (obs/sec)	Accuracy(%)
1	ShuffleNet V2	~3300	91.3
2	SqueezeNet	~140000	95.2
3	ShuffleNet + SqueezeNe	~81000	96.0

By using the above division of images, we have extracted features using ShuffleNet, and then these features have use to classify malaria parasites. This complete experiment gave 91.3 % results with a prediction speed of 3300 obs/sec.

3.4.2. Experiment 2

In this experiment, we used 10000 Blood cell images which were also divided into two classes, 5000; of them consisted of infected malaria cells, and the second half was uninfected. Using the above division of images, we have extracted features using SqueezeNet, and then these features have been used to classify malaria parasites. This complete experiment gave 95.3 % results with a 140000 obs/sec prediction speed.

3.4.3. Experiment 3

In this experiment, we used 27560 Blood cell images, divided into two classes, 13780 consisting of infected malaria cells, and the second half was uninfected. Using the above division of images, we have extracted features using SqueezeNet and ShuffleNet. Then, by using the linear feature fusion method, both After fusing features, we classified them using Quadratic SVM and got 96% result with minimum time as per what we have seen in Experiment no.1 and Experiment no.2. Actual goal to do this experiment was to minimize downsides of both algorithms and get more authentic results with less hardware cost. We have shown this in the table below.

For more in detail view of results that how the classification model performs on a collection of test data for which the true values are known. Here is the confusion matrix in figure 6.

Table 4: Result of Proposed Method by Performance Evaluation Metrics

Classification Algorithm	Images	Accuracy	Precision	Recall	Specificity	Sensitivity	AUC
Support Vector Machine	27,560	96.0	0.98	0.97	0.9699	0.9481	0.99

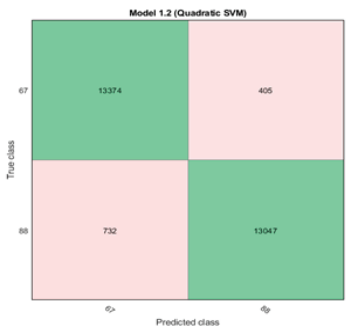

Figure 6: Confusion Matrix with Complete No. of Observations [67=Uninfected, 88=Parasitized]

True Positive: Predicted values that are positive are 13374. False Positive Predicted values are 405, and False Negative values are 732. True Negative: Predicted values predicted an actual negative is 13047. The figure shows verified True Positive and False Negative rates on our Scattered plot.

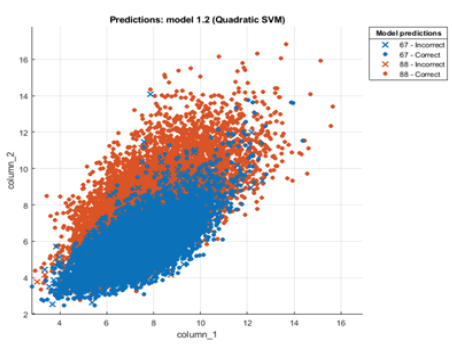

Figure 7: Model Predictions on Complete Data [67=Uninfected, 88=Parasitized]

Table 5: Accuracy comparison of literature techniques

References	Accuracy
[13]	98.61%
[12]	94.58%
[11]	97.06%

Moreover, we see a False positive rate graph (x-axis) versus a true positive rate (y-axis) for a range of candidate threshold values from 0.0 to 1.0 in Figure 8.

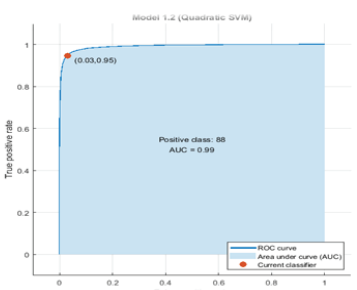

Figure 8: True Positive rate with ROC Curve for Parasitized class

4. CONCLUSION

Among the many underdeveloped nations currently grappling with the deadly disease of malaria, Pakistan is one of them. As a result of this problem, the country's economy eventually suffers. A report by the WHO indicates that the Middle East and Africa are the regions where malaria is most widespread and is still being targeted. There is a high cost associated with malaria treatment and equipment. AI-based systems that use artificial intelligence can reduce the actual cost of diagnosing malaria since they

use artificial intelligence in the process. SqueezeNet and ShuffleNet models were enhanced with feature fusion based on the prior models. They considered precision, testing duration, and accuracy. The proposed model performs better than the previous one with 96% classification accuracy.

A deep learning model called ShuffleNet and SqueezeNet is used in the first step of the analysis to extract the characteristics of cell pictures. Following the reduction of the dimension of the feature using PCA (Principal Component Analysis), we applied a classification model to categorize the feature. Through the development of low-cost tools for detecting malaria parasites, as a result of this study, developing nations will be able to deal with this disease more effectively. One of the advantages of this approach is that it can be used not only to detect malaria, one of the common endemic diseases, but also to classify other forms of the disease and their types that may also be present. As a result of the integration of medical diagnoses into the system, a considerable amount will reduce the workload for doctors, and efficiency and empirical analysis of the data will be improved with good results.

REFERENCES

[1] Pham, N. M., Karlen, W., Beck, H. P., & Delamarche, E. Malaria and the 'last' parasite: how can technology help? Malaria journal, 17(1), 260. https://doi.org/10.1186/s12936-018-2408-0, 2018.

[2] Oluwagbemi, O. O., Adeoye, E., & Fatumo, S. Building a computer-based expert system for malaria environmental diagnosis: An alternative malaria control strategy. Egyptian Computer Science Journal, 33(1), 55-69, 2009.

[3] Usama, M., Ahmad, B., Wan, J., Hossain, M. S., Alhamid, M. F., & Hossain, M. A. Deep feature learning for disease risk assessment based on convolutional neural network with intra-layer recurrent connection by using hospital big data. IEEE Access, 6, 67927-67939, 2018.

[4] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, 15(1), 1929-1958, 2014.

[5] N. E. Ross, C. J. Pritchard, D. M. Rubin, and A. G. Dusé, "Automated image processing method for the diagnosis and classification of malaria on thin blood smears," Medical & Biological Engineering & Computing, vol. 44, no. 5, pp. 427–436, 2006.

[6] M. Poostchi, K. Silamut, R. J. Maude, S. Jaeger, and G. R. Thoma, "Image analysis and machine learning for detecting malaria," Translational Research, vol. 194, pp. 36–55, 2018.

[7] Sivaramakrishnan, R., Antani, S., & Jaeger, S. Visualizing deep learning activations for improved malaria cell classification. In Medical informatics and healthcare (pp. 40-47), October, 2017.

[8] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee, June, 2009.

[9] 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.

[10] Shang, W., Sohn, K., Almeida, D., & Lee, H. Understanding and improving convolutional neural networks via concatenated rectified linear units. In international conference on machine learning (pp. 2217-2225), June,2016.

[11] Gopakumar, G. P., Swetha, M., Sai Siva, G., & Sai Subrahmanyam, G. R. K. Convolutional neural network-based malaria diagnosis from a focus stack of blood smear images acquired using custom-built slide scanner. Journal of biophotonics, 11(3), e201700003, 2018.

[12] Shi, L., Guan, Z., Liang, C., & You, H. Automatic Classification of Plasmodium for Malaria Diagnosis based on Ensemble Neural Network. In Proceedings of the 2020 2nd International Conference on Intelligent Medicine and Image Processing (pp. 80-85), April, 2020.

[13] Y. Dong, Z. Jiang, H. Shen and W. D. Pan, "Classification accuracies of malaria infected cells using deep convolutional neural networks based on decompressed images," SoutheastCon

2017, Charlotte, NC, pp. 1-6, doi: 10.1109/SECON.2017.7925268, 2017.

[14] Kuo, P. C., Cheng, H. Y., Chen, P. F., Liu, Y. L., Kang, M., Kuo, M. C., ... & Liu, D. P. Assessment of expert-level automated detection of Plasmodium falciparum in digitized thin blood smear images. JAMA network open, 3(2), e200206-e200206, 2020.

[15] Kaur, M. and N. Sohi. A novel technique for content based image retrieval using color, texture and edge features. in 2016 international conference on communication and electronics systems (ICCES). IEEE, 2016.

[16] Vehtari, A., A. Gelman, and J. Gabry, Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and computing. 27(5): p. 1413-1432, , 2017.

[17] World Health Organization. World Malaria Report 2019; World Health Organization: Geneva, Switzerland; Licence: CC BY-NC-SA 3.0 IGO, 2019.

[18] Khan, M., Mahmood, H. Z., Noureen, S., Muhmood, K., Husnain, M. I. U., & Hameed Khaliq, I. The health burden of malaria and household choices regarding treatment and prevention in Pakistan. Tropical Biomedicine, 36(3), 664-676, 2019.

[19] Abanyam, N. L. Malaria and Socio-Economic Development in Nigeria. Lead City Journal of the Social Sciences, 81.

[20] Marrelli, M. T., & Brotto, M. The effect of malaria and antimalarial drugs on skeletal and cardiac muscles. Malaria journal, 15(1), 1-6, 2016.

[21] Fan, G., Chen, F., Chen, D., Li, Y., & Dong, Y. (2020). A deep learning model for quick and accurate rock recognition with smartphones. Mobile Information Systems, 2020.

[22] Asim, M., Ullah, S., Razzaq, A., & Qadri, S. Varietal Discrimination of Guava (Psidium Guajava) Leaves Using Multi Features Analysis. International Journal of Food Properties, 26(1), 179-196, 2023.

[23] Alyas, T., Javed, I., Namoun, A., Tufail, A., Alshmrany, S., & Tabassum, N. Live migration of virtual machines using a mamdani fuzzy inference system. Computers, Materials & Continua, 71(2), 3019-3033, 2022.

[24] Saleem, M. A., Aamir, M., Ibrahim, R., Senan, N., & Alyas, T. An Optimized Convolution Neural Network Architecture for Paddy Disease Classification. Computers, Materials & Continua, 71(3), 2022.

[25] Tabassum, N., Rehman, A., Hamid, M., Saleem, M., Malik, S., & Alyas, T. (2021). Intelligent nutrition diet recommender system for diabetic's patients. Intelligent Automation & Soft Computing, 29(3), 319-335, 2021.

[26] Tabassum, N., Namoun, A., Alyas, T., Tufail, A., Taqi, M., & Kim, K. H. Classification of Bugs in Cloud Computing Applications Using Machine Learning Techniques. Applied Sciences, 13(5), 2880, 2023.