

Automated Malaria Identification Using a Hybrid CNN-RNN Model on Microscopic Blood Smears

Ramnadh Babu T G ¹, Hashmi Patan ², Anusha Pinneboina ³, Hemalatha Muchumari⁴, Samatha Anthappagudem ⁵, Sharada Adepu ⁶, Dr.Sireesha Moturi⁷

^{1,2,3,4,7}Department of Computer Science and Engineering,

Narasaraopeta Engineering College (Autonomous), Narasaraopet.

⁵Department of AIML, GRIET. ⁵ samatha1774@gmail.com,

⁶Department of CSE, G. Narayanamma Institute of Technology & Science(women). ⁶ sharada@gnits.ac.in,

¹baburamnadh@gmail.com, ²patanhashmi20@gmail.com, ³anushapinneboina888@gmail.com,

⁴hemalatha@gmail.com, ⁷sireeshamoturi@gmail.com

Abstract—Malaria remains a life-threatening disease in many parts of the world, particularly in regions with limited access to expert medical diagnosis. This project presents an automated Deep Learning-based approach for accurate malaria detection using microscopic blood smear images. The automation combines feature extraction using visual CNNs and sequential RNNs for detecting patterns in sequences. Meaningful features were extracted from images using pretrained CNNs such as DenseNet121 and MobileNetV2. These features were then classified using RNNs such as LSTM, GRU, and BiLSTM to construct CNN-RNN hybrid models. After thorough evaluation using a publicly available malaria dataset, the models showed a need for high image quality as well as the rigorously curated clean dataset for robust training. The hybrid DenseNet121-LSTM-GRU model showed the highest accuracy of 97.34%, yielding the best model with precision, recall and F1-score of 0.96 for both the infected and noninfected classes. The results demonstrate how well the hybrid CNN-RNN framework automates the diagnosis of malaria and present a viable method for quick and accurate screening in environments with limited resources.

Keywords: Malaria detection, Deep Learning, CNN, RNN, hybrid model, transfer learning, MobileNetV2, BiLSTM, LSTM, GRU.

I. INTRODUCTION

Malaria is a severe and potentially fatal illness caused by Plasmodium parasites, transmitted to humans through the bites of infected female Anopheles mosquitoes. It remains a major public health issue, particularly in tropical and subtropical regions, where it causes millions of infections and hundreds of thousands of deaths each year. Early and accurate diagnosis is essential for effective treatment and disease management. Traditional diagnostic techniques, such as manual blood smear examination under a microscope, are often time-consuming

and require skilled personnel. These constraints make them less practical in healthcare settings with limited resources. In recent years, deep learning techniques have shown great promise in medical image analysis, especially Convolutional Neural Networks (CNNs)[1][2] due to their ability to identify and learn spatial features from images. However, CNNs are limited in capturing contextual or sequential relationships, which can be important in complex diagnostic tasks. To overcome this limitation, a hybrid deep learning model is proposed that combines the spatial feature extraction capabilities of CNNs with the sequential learning strength of Recurrent Neural Networks (RNNs). Blood smear images are first processed using pretrained CNNs, specifically DenseNet121 and MobileNetV2, to extract high-level image features. These features are then passed to RNN-based classifiers such as Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Bidirectional LSTM (BiLSTM), which help in capturing temporal patterns to enhance classification accuracy. The input dataset underwent extensive preprocessing steps including resizing, normalization, and contrast enhancement to improve the model's generalization and reduce the risk of overfitting. This integrated approach significantly improves the reliability of malaria detection from microscopic blood smear images.

II. RELATED WORK

Antora Dev and Mostafa M.[1] Fouda Created a hybrid deep learning system integrating CNNs and RNNs for precise detection of malaria in microscopic images. Wu, J. and Zhao,[2] Developed a multi-level attention split network for advanced feature extraction to improve malaria cell classification. Alonso-Ramirez et al.[3] Cre-

ated lightweight CNNs for fast, real-time malaria diagnostics appropriate to the resource constraints of certain regions, emphasizing speed and resource efficiency. D. Sukumarran et al.[4] Extended YOLOv4's capabilities to simultaneously identify malaria parasites, infected cells, and vectors within a unified deep learning framework. H. A. H. Chaudhry et al.[5] Developed a deep learning system of low complexity for the automated analysis of blood smears, enhancing efficiency for blood smear examinations. H. M. Asif et al.[6] Enhanced feature propagation for parasite detection utilizing a deep bottleneck residual network, amplifying detection accuracy. A. Anorboev et al.[7] Introduced SIEC, a Selective Intensity Ensemble Classifier implementing a triple-threshold approach for robust classification of malaria images. S. B. Nettur et al. [8] Introduced UltraLightSqueezeNet, a new ultra-compact CNN for mobile uses in malaria detection which cut the number of parameters by 54. Y. Benachour, F. Flitti, and H. M. Khalid [9] Performed a comparative study of different CNN architectures focusing on their accuracy for malaria detection and their detection capabilities. K. V. N. Reddy et al.[10] Demonstrated the use of CNNs for image-based traffic sign detection, illustrating the scope of deep learning technology. S. Moturi et al.[11] Focused on ensemble deep learning for lung cancer detection, applicable in refining models for the diagnosis of malaria. Greeshma, B., Sireesha, M., and Thirumala Rao, S.N.[12] Exploited the use of CNNs for the detection of arrhythmias, showing the growing extent of deep learning in medicine. S. L. Jagannadham et al.[13] Used CNN for the detection of brain tumors and proved their value in the analysis of medical images.

III. METHODOLOGY

The automated malaria detection system starts with collecting images of blood smears under a microscope. Each image goes through a preprocessing sequence consisting of five stages: resizing, denoising, CLAHE-based contrast enhancement, sharpening, and normalization. Image sharpening boosts image quality. Extracting features with pretrained CNNs (MobileNetV2 and DenseNet121) uses the CNNs without their classification heads. These features are utilized by RNN classifiers (LSTM, GRU, and BiLSTM) to capture temporal patterns. To capture both spatial and sequential information, hybrid models are created, like MobileNetV2-LSTM-BiLSTM and DenseNet121-GRU-LSTM. A last binary classifier distinguishes each red blood cell as parasitized or uninfected. Figure 1 illustrates the overall workflow.

A. Dataset Description

The *Cell Images for Detecting Malaria* dataset, provided by the National Institutes of Health (NIH), comprises 27,558 labeled RGB images of human red blood

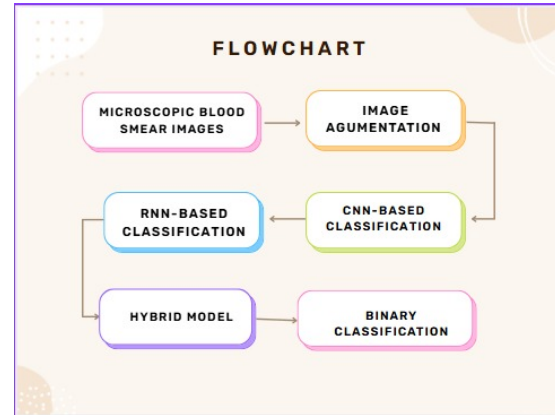


Figure 1: Workflow of the proposed malaria detection system

cells. The dataset is evenly split into two categories: *Parasitized* and *Uninfected*, with approximately 13,779 images in each class.

All images were captured at 100x magnification using a standard light microscope. They vary slightly in resolution and are organized into two separate directories corresponding to their class labels, supporting binary classification tasks.

Due to its balanced distribution, large volume, and public availability, this dataset is widely utilized in malaria detection studies. It serves as a benchmark for evaluating the performance of CNN-based and hybrid CNN-RNN models in automated parasite identification.

To provide a visual understanding of the dataset, Figure 2 presents representative samples of parasitized and uninfected cells. The top row displays parasitized cells, while the bottom row shows uninfected cells.

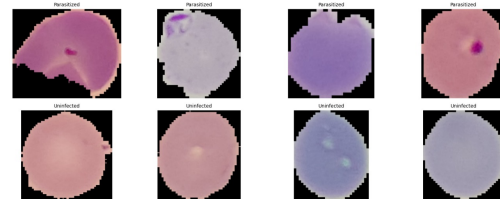


Figure 2: Sample Images In The Dataset

B. Dataset Splitting

After preprocessing, the dataset was divided into three distinct subsets: training, validation, and testing, following an 80:10:10[12],[13] ratio for each class. The original dataset consisted of 13,779 images in both the *Parasitized* and *Uninfected* categories. Based on the defined split, 10,334 images per class were allocated to the training set, 2,066 to the validation set, and 1,379 to the testing set.

This class-balanced partitioning ensured uniform distribution across all subsets, which is essential for effective model training, hyperparameter tuning, and unbiased performance evaluation. A summary of this data distribution is presented in Table I.

TABLE I: Class-wise Distribution of Preprocessed Images into Training, Validation, and Test Sets

Class	Total	Training	Validation	Test
Parasitized	13,779	10,334	2,066	1,379
Uninfected	13,779	10,334	2,066	1,379
Total	27,558	20,668	4,132	2,758

C. Image Preprocessing

The image preprocessing pipeline included several enhancement steps to improve the quality and uniformity of blood smear images. All raw images were first resized to 128×128 pixels for standardized input dimensions. To reduce background noise while preserving cell structures, the Fast Non-Local Means denoising algorithm was applied, followed by contrast enhancement using CLAHE in the LAB color space to emphasize subtle details. A sharpening filter improved edge clarity and boundary features, and pixel intensity values were normalized to $[0, 1]$ for consistent scaling in deep learning models.

Figure 3 illustrates the preprocessing steps for parasitized and uninfected cells. Each transformation—from resizing to normalization—enhanced visual clarity and supported effective feature learning during training.

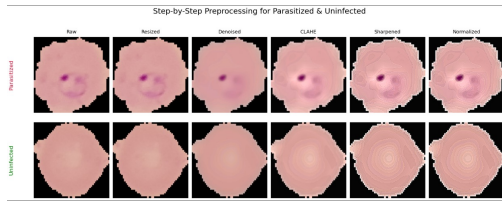


Figure 3: Preprocessing Illustration

D. Model Architecture

The malaria detection system was implemented as a multi-phase deep learning pipeline, with each phase designed to balance computational efficiency and predictive performance. The overall architecture comprises three key stages:

Feature Extraction: Performed using pretrained Convolutional Neural Networks (CNNs), such as MobileNetV2 and DenseNet121, to capture high-level spatial features from the blood smear images. **Sequence Modeling:** Applied using various Recurrent Neural Network (RNN) architectures, including Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and

Bidirectional LSTM (BiLSTM), to model temporal dependencies in the extracted features. **Hybrid CNN–RNN Architecture:** The CNN and RNN modules are combined into an end-to-end hybrid framework capable of learning both spatial and sequential patterns from the data.

The model ultimately performs binary classification, predicting whether a given cell image belongs to the *Parasitized* or *Uninfected* class.

Figure 4 presents the complete architectural design of the hybrid CNN–RNN pipeline used for malaria cell classification.

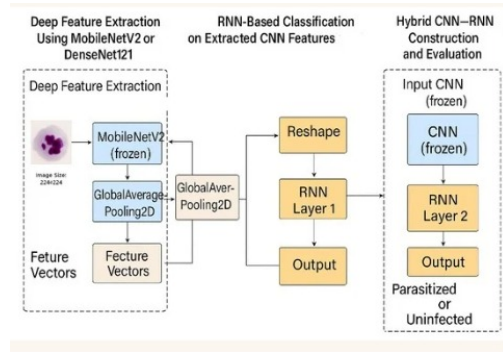


Figure 4: Hybrid CNN–RNN architecture for malaria detection

Phase A: Deep Feature Extraction Using MobileNetV2 and DenseNet121

Using pretrained MobileNetV2 and DenseNet121 for feature extraction, we removed the classification heads and retained ImageNet weights. A GlobalAveragePooling2D layer was used to convert the outputs into fixed-length vectors. The model's base layers were frozen, allowing the model to extract spatial features with distortion and the base layers were not changed. The images were resized and split into training, validation, and test sets, and then processed in batches through the CNNs. The spatial differences were encoded in the feature vectors, stored in NumPy format, and were essential in the RNN classification performed next.

Phase B: RNN-Based Classification on Extracted CNN Features

During this phase, we captured deep temporal dependencies from spatial data by using features extracted from pretrained CNNs like MobileNetV2 or DenseNet121, feeding them into an RNN classifier. For the features from the training, validation, and test sets, we reshaped them into the RNN-compatible format of (samples, timesteps, features), where each sample contained 1 timestep and 1024 features. We also loaded binary class labels: '0' for uninfected and '1' for parasitized.

Each RNN model consisted of two stacked recurrent layers (LSTM \rightarrow LSTM or GRU \rightarrow GRU) and a Dropout layer for regularization. The model also included a Dense layer for abstraction and a binary classification output layer with a sigmoid activation. The models were trained with the Adam optimizer and binary cross-

Phase C: Hybrid CNN–RNN Model Construction and Evaluation

The refinement of the classification of the malaria disease occurred in Phase C of the project by creating hybrid deep learning models which integrate spatial feature extraction CNNs DenseNet121 and MobileNetV2 with RNNs LSTM, GRU, and BiLSTM for temporal sequencing. The spatial features obtained in Phase A were processed and applied with several CNN–RNN hybrids, for example, DenseNet121–GRU–BiLSTM. With each of the models, dropout, dense layers with softmax output, and SGD optimizers with EarlyStopping and ReduceLROnPlateau callbacks were employed. 18 hybrid models were assessed by the Accuracy, Precision, Recall, and F1 score, and confusion matrices. Integration of the CNN and RNNs in the hybrid models was justified by the strong increasing validation and classification accuracy results of the best performing models.

IV. RESULTS AND DISCUSSIONS

The performance of RNN-based classifiers (LSTM, GRU, BiLSTM) and their hybrid configurations is evaluated in Table II. The table compares hybrid CNN–RNN models for malaria cell classification using DenseNet121 and MobileNetV2 as feature extractors. Among all models, the DenseNet121_LSTM_GRU achieved the best results with validation accuracy of 97.21%, test accuracy of 97.34%, precision of 0.9717, recall of 0.9724, and F1-score of 0.9721. Strong performances were also observed in DenseNet121_BiLSTM_LSTM and DenseNet121_LSTM_LSTM, both with test accuracies above 95%. In contrast, MobileNetV2-based models, though more efficient, showed slightly lower performance; the best was MobileNetV2_GRU_BiLSTM with 94.48% accuracy and F1-score of 0.9455. Notably, deeper BiLSTM architectures did not consistently outperform simpler ones, suggesting that balanced architectures yield better results.

Overall, DenseNet121 combined with LSTM or GRU networks provides the highest accuracy and reliability in malaria image classification.

TABLE II: Comparative Evaluation of Hybrid CNN–RNN Models for Malaria Classification Using Microscopic Blood Smear Images

MODEL	VAL ACCURACY	TEST ACCURACY	PRECISION	RECALL	F1 SCORE
DENSENET121_GRU_BiLSTM	0.9521	95.0326	0.9436	0.9579	0.9507
MOBILENETV2_GRU_BiLSTM	0.9492	94.8888	0.9354	0.9558	0.9455
DENSENET121_BiLSTM_LSTM	0.9562	95.1777	0.9456	0.9587	0.9521
MOBILENETV2_BiLSTM_LSTM	0.9451	93.9086	0.9322	0.9471	0.9396
DENSENET121_LSTM_LSTM	0.9538	95.1777	0.9456	0.9587	0.9521
MOBILENETV2_LSTM_LSTM	0.9453	93.6548	0.9288	0.9456	0.9371
DENSENET121_LSTM_GRU	0.9721	97.3406	0.9717	0.9724	0.9721
MOBILENETV2_LSTM_GRU	0.9472	93.9086	0.9320	0.9345	0.9336
DENSENET121_BiLSTM_BiLSTM	0.9430	95.0000	0.9356	0.9234	0.9100
MOBILENETV2_BiLSTM_BiLSTM	0.9400	93.5501	0.9432	0.9212	0.9422
DENSENET121_LSTM_BiLSTM	0.9477	94.7426	0.9452	0.9500	0.9476
MOBILENETV2_LSTM_BiLSTM	0.9468	94.6390	0.9321	0.9212	0.9323
DENSENET121_GRU_LSTM	0.9528	95.2139	0.9515	0.9529	0.9522
MOBILENETV2_GRU_LSTM	0.9526	95.1156	0.9123	0.9324	0.9432
DENSENET121_GRU_GRU	0.9511	95.2139	0.9508	0.9536	0.9522
MOBILENETV2_GRU_GRU	0.9497	94.8564	0.9321	0.9213	0.9432
DENSENET121_BiLSTM_GRU	0.9482	94.8151	0.9414	0.9558	0.9485
MOBILENETV2_BiLSTM_GRU	0.9380	93.1835	0.9281	0.9362	0.9321

A. Training and Validation Accuracy

As shown in Figure 5, the DenseNet121_LSTM_GRU model was trained for 17 epochs on microscopic blood smear images. The left sub-graph illustrates the progression of training and validation accuracy over time. The training accuracy steadily increased and reached approximately 97.8%, while the validation accuracy stabilized around 96.3%, indicating effective learning and good generalization performance.

The right sub-graph presents the training and validation loss trends. The training loss shows a consistent downward trajectory, while the validation loss decreases and plateaus, stabilizing after the 10th epoch. This suggests that the model is converging in a balanced training environment, with no signs of overfitting.

Overall, the performance curves highlight that the DenseNet121_LSTM_GRU model maintains consistent behavior across training and validation, confirming its reliability and robustness.

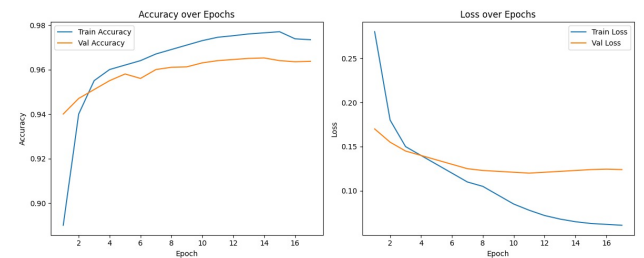


Figure 5: Training and validation accuracy and loss curves

B. Performance Metrics of DenseNet121_LSTM_GRU Model

The of the hybrid model **DenseNet121_LSTM_GRU** applied to the malaria cell classification task is visualized in Figure 6. The figure summarizes the model's key performance metrics—*Precision*, *Recall*, and *F1-Score*.

The model achieved a high precision of **0.9717**, indicating a strong ability to correctly identify true posi-

tive cases while minimizing false positives. The **recall score of 0.9724** reflects the model's high sensitivity, successfully detecting a majority of infected cells (true positives) in the sample population. The resulting **F1-score of 0.9721** demonstrates a balanced performance between precision and recall, underscoring the model's robustness in minimizing both false negatives and false positives.

These high values across all three metrics confirm the model's effectiveness and reliability in accurately diagnosing malaria based on microscopic blood smear images, supporting its practical use in real-world diagnostic scenarios.

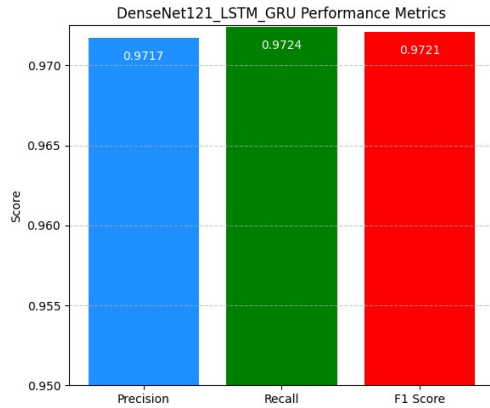


Figure 6: DenseNet121_LSTM_GRU classification performance.

C. Confusion Matrix

The **DenseNet121_LSTM_GRU** model achieved 97.34% accuracy, as shown in Figure ???. The confusion matrix illustrates results for two classes: *Parasitized* and *Uninfected*. The model correctly identified 1,341 uninfected and 1,340 parasitized images, with only 39 and 38 misclassified, respectively. Its near-symmetry indicates balanced sensitivity and specificity, confirming reliability in detecting malaria-infected and healthy cells with minimal error.

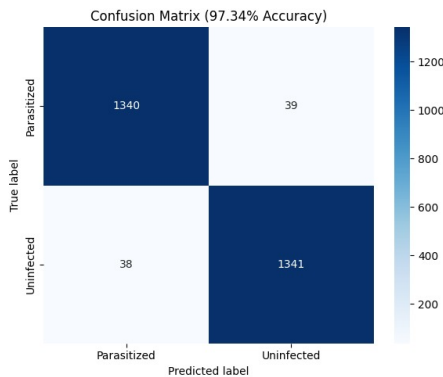


Figure 7: Confusion matrix of the DenseNet121-LSTM-GRU model.

D. Model Output Summary

The output for the best-performing model in malaria cell classification, which is a hybrid model of **DenseNet121_LSTM_GRU**, is provided in Figure 8. Figure 8 shows the predicted class labels of the model, which are *Parasitized* and *Uninfected*, along with the confidence scores in parentheses. Each image is heat-mapped, where the model's relevance for the prediction is depicted by the darker regions.

The overall model confidence scores being quite high and close to 1.00 attest to the model's strong capacity for prediction. It can be concluded that the hybrid DenseNet121-LSTM-GRU model demonstrates high performance in capturing image features in both spatial and sequential patterns, which is essential for accurate classification and for handling difficult shape or stain ambiguities. This reveals that the model is both accurate and interpretable, which adds trust to the predictions made.

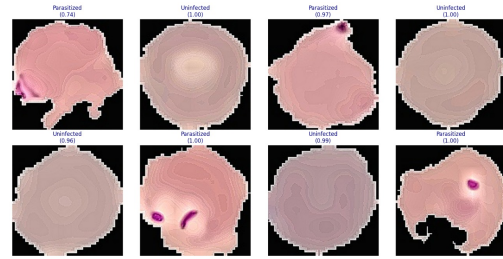


Figure 8: Visualization of Infected and Uninfected Cell Detection by Hybrid Model

E. Explainability and Credibility

To identify the areas of the image that had the greatest influence on the model's predictions, we used Gradient-weighted Class Activation Mapping (Grad-CAM). The model's attention to clinically relevant features is confirmed by the close correspondence between the highlighted areas and parasite-infected regions, as shown in Figure 8. The system's decisions are more trustworthy as a result of this interpretability, which is crucial for medical adoption.

F. Error Analysis

In order to assess the suggested hybrid CNN-RNN model, we processed correctly classified and mislabeled malaria cell images. As Figure 9 illustrates, the majority of cells are properly classified, with certain uninfected cells incorrectly labeled as parasitized because of visual similarity or staining artifacts, demonstrating borderline case problems.

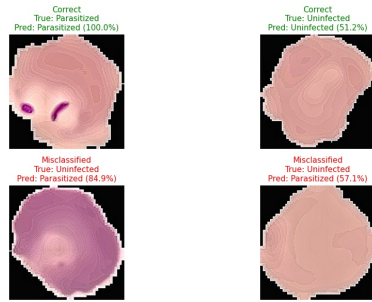


Figure 9: Error analysis of correctly classified and mislabeled malaria cell images.

V. CONCLUSION

As good as the DenseNet121–LSTM–GRU is at identifying infected versus uninfected cells, there remains potential for improvement with the hybrid CNN–RNN model. Scientists can optimize convolutional layers by unfreezing and fine-tuning them to more thoroughly capture domain-specific information. Adding attention mechanisms, spatial and temporal, could further increase feature relevance and model performance.

Transformer models like ViTs and Swin Transformers are particularly good at understanding contextual and spatial relationships and are potential contenders. Their integration onto edge or mobile platforms would allow for real-time malaria diagnosis in disadvantaged areas. Through integration with XAI, the models can offer clear reasons, enhancing clinical confidence.

Extending the model to handle clinical metadata or multi-class classification of *Plasmodium* species would further enhance diagnostic accuracy and functionality.

VI. FUTURE SCOPE AND OPPORTUNITIES

Even though the DenseNet121–LSTM–GRU model attained high accuracy distinguishing between the infected and uninfected cells, there is still possible advancement. Further efforts may consist of better domain adaptation through more granular optimization, focusing on the unfreezing of specific CNN layers and incorporation of attention mechanisms to improve interpretability and focal precision. Performance may be further augmented by consideration of transformer based models such as ViT or Swin Transformers. Device-level deployment could enable real-time screening in remote areas, and trust in the model could be bolstered through the application of XAI techniques. Enhanced diagnostic utility could be attained by expanding to multi-class classification for different *Plasmodium* species and using clinical metadata. Testing multiple sites and populations is critical to the model's proven reliability and widespread applicability. To gain a better understanding of model decision-

making, future research will also incorporate attention-based explainability techniques outside of Grad-CAM.

REFERENCES

- [1] A. Dev and M. M. Fouda, "Developing and deploying a hybrid deep learning system to identify malaria in microscopic pictures," *IEEE Xplore*, 2024.
- [2] J. Wu and X. Zhao, "An innovative method for classifying malaria cells that makes use of a multi-level attention split network," *Data (MDPI)*, 2024.
- [3] J. Alonso-Ramírez, L. F. Mejía, S. Gómez, P. A. Restrepo, and F. Rincón, "Use of lightweight cnns in low-resource settings for real-time malaria detection," *Technologies (MDPI)*, 2024.
- [4] D. Sukumarran, A. Nair, R. P., S. K. Menon, and P. Varma, "Utilizing an improved yolov4 deep learning model to identify parasites, transmission vectors, and cells infected with malaria," *Vectors and Parasites*, 2024.
- [5] H. A. H. Chaudhry, F. Zahra, M. Haseeb, S. A. Raza, and A. Aslam, "Design of an efficient, low-complexity deep learning model for automated analysis of malaria-infected blood smear images," *Applications of Neural Computing*, 2024.
- [6] H. M. Asif, S. Noreen, B. Khan, F. Yousuf, and I. Ahmed, "Detection of malaria parasites using an advanced deep bottleneck residual network approach," *Intelligent and Complex Systems*, 2024.
- [7] A. Anorboev, E. Petrova, S. Saidov, and D. Khamidov, "Selective intensity ensemble classifier (siec): A triple-threshold strategy for microscopic malaria cell image classification," *IEEE Access*, vol. 13, pp. 101 609–101 623, 2025.
- [8] S. B. Nettur, N. V. R., R. Menon, K. Joseph, and A. Mathur, "UltralightsqueezeNet: A deep learning architecture for malaria classification with up to 54× fewer trainable parameters for resource constrained devices," *IEEE Access*, vol. 13, pp. 89 428–89 440, 2025.
- [9] Y. Benachour, F. Flitti, and H. M. Khalid, "Enhancing malaria detection through deep learning: A comparative study of convolutional neural networks," *IEEE Access*, vol. 13, pp. 35 452–35 477, 2025.
- [10] K. V. N. Reddy, Y. Narendra, M. A. N. Reddy, A. Ramu, D. V. Reddy, and S. Moturi, "Automated traffic sign recognition via cnn deep learning," in *2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, Gwalior, India, 2025, pp. 1–6.
- [11] S. Moturi, M. Ainavolu, N. R. Dokku, P. Kasula, N. Yaragani, and V. Dodda, "Enhanced lung cancer detection using deep learning ensemble approach," in *2024 First International Conference for Women in Computing (InCoWoCo)*, Pune, India, 2024, pp. 1–7.
- [12] G. B., S. M., and T. R. S. N., "Detection of arrhythmia using convolutional neural networks," in *Proceedings of Second International Conference on Sustainable Expert Systems*, ser. Lecture Notes in Networks and Systems, S. Shakya, K. L. Du, and W. Haoxiang, Eds. Springer, Singapore, 2022, vol. 351.
- [13] S. L. Jagannadham, K. L. Nadh, and M. Sireesha, "Brain tumour detection using cnn," in *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Palladam, India, 2021, pp. 734–739.