MALARIA CELL DETECTION USING CNN

Rishabh Tiwari[1], Prabhat Kr. Yadav[2], Sagar Saini[3], Ritik Kasaushan[4]

[1] *Student, Institute of Technology and Management, Uttar Pradesh, India*
[2] *Student, Institute of Technology and Management, Uttar Pradesh, India*
[3] *Student, Institute of Technology and Management, Uttar Pradesh, India*
[4] *Student, Institute of Technology and Management, Uttar Pradesh, India*

ABSTRACT

Malaria is a significant global health issue, prompting the use of modern technology like convolutional neural networks (CNNs) for automated parasite detection in blood smear images. CNNs offer efficient feature extraction, aiding in accurate classification. Our study evaluates six pre- trained CNN models and various classifiers, achieving over 94% accuracy in malaria detection, with less complexity than traditional methods. This approach addresses the shortage of diagnostic resources and skilled personnel, offering a scalable solution to combat the disease.

Keyword: *- Malaria, global health, convolutional neural networks (CNNs), automated parasite detection, blood smear images, pre-trained models, classifiers, accuracy, scalability, diagnostic resources.*

1. INTRODUCTION

Malaria, a leading cause of morbidity and mortality worldwide, presents significant diagnostic challenges, particularly in resource-limited settings. Traditional diagnostic methods, primarily involving microscopic examination of blood slides by trained specialists, are time- consuming and subject to variability based on the technician's expertise. This project aims to revolutionize malaria detection by introducing an advanced machine learning solution tailored to overcome these hurdles.

Our model leverages a robust convolutional neural network (CNN) architecture to systematically analyze and identify infected red blood cells in standard microscope slides. Utilizing a dataset of 27,735 single cell images, our CNN has been meticulously trained to detect subtle patterns and variations in cell morphology, achieving high accuracy in predicting viral presence in small blood vessels. This model significantly enhances detection speed and reliability, reducing dependency on high-level expert intervention and complex microscopic equipment.

We address the critical need for rapid, accurate malaria diagnostics in settings with limited healthcare resources, focusing on vulnerable populations such as young children, pregnant women, and travelers from non-endemic areas. By integrating our CNN model with real-time data processing capabilities, we are setting a new standard in malaria diagnostics, aiming for widespread application in medical facilities that previously relied on conventional methods.

This project not only promises to make malaria detection more accessible and efficient but also sets the groundwork for applying similar machine learning approaches to other parasitic diseases, marking a pivotal step towards smarter, technology-driven global health solutions.

2. NEED FOR MALARIA CELL DETECTION

Malaria, a severe and infectious disease transmitted by mosquitoes, is widespread globally, particularly in tropical regions. The severity and fatal nature of the disease often cause initial symptoms similar to the flu, appearing just a few days after a mosquito bite. Without timely and appropriate treatment, the disease can escalate rapidly, potentially leading to death.

The World Health Organization has highlighted several critical statistics about malaria, noting that it endangers the majority of the global population with over 200 million cases annually. Traditional diagnostic methods, heavily reliant on manual labor, are not only labor-intensive but also prone to inaccuracies, increasing the risk of erroneous medical decisions. Our project aims to enhance the healthcare system by delivering quicker and more precise diagnostic results.

Malaria claims the lives of at least 400,000 people each year and is particularly rampant in areas prone to outbreaks, where factors such as extreme poverty, lack of access to quality healthcare, and the presence of the disease carriers— mosquitoes—exacerbate the situation.

Our goal is to improve diagnostic accuracy while also designing a system that is both compact and efficient. This system would utilize computer-assisted technology adaptable for use in remote locations via edge and online devices.

Currently, the primary diagnostic method involves examining blood smears under a microscope to identify infected cells. In this process, a reagent is added to a blood sample to help detect parasites in the red blood cells (RBCs). Doctors then count the infected RBCs, sometimes up to seven thousand, according to WHO guidelines.

Malaria's prevalence, especially in tropical climates, and its symptoms, which are initially similar to flu, underscore the urgency for early detection and treatment. If not addressed promptly, the infection can remain in the body for up to a year, leading to severe complications and death. Early and accurate testing is crucial for saving lives.

3. SYSTEM ANALYSIS AND DESIGN

System Analysis and Design of a malaria cell detection model, utilizing Google Colab and TensorFlow for development. The process begins with loading the database into a Colab notebook, leveraging its capabilities in machine learning software development with libraries like OpenCV, TensorFlow, or Keras. TensorFlow is applied to build and train models, where data preprocessing involves dividing datasets for testing and training and employing data augmentation techniques. Two models are constructed: one from scratch using custom layers, and another using transfer learning. The performance of both models is assessed in terms of accuracy and loss. The chapter also explores key algorithms utilized in the models, including ReLU, Sigmoid, Dropout, and Crossentropy. Furthermore, it provides insights into Convolutional Neural Networks (CNNs) and specifically discusses the Inception v3 model, highlighting its optimizations and advancements in image classification. Through these methodologies and algorithms, the chapter aims to enhance the accuracy and efficiency of malaria cell detection systems, contributing to improved healthcare diagnostics.

4. LITERATURE REVIEW

Author-Amogh Manoj Joshi, Ananta Kumar Das , Subhasish Dhal
Finding- Modern scientific advancements play a pivotal role to combat the disease, along with biomedical research by the medical experts to possibly eradicate this disease from all parts of the world.

Author-Uzair Adamjee, Sayeed Ghani
Finding- When we gave different type of blood smear images to the deep learning model even in that scenario, model is able to identify patterns and learn features with an accuracy up to 94% .

5. EXISTING METHODOLOGIES

Malaria diagnosis primarily involves detecting parasites, identifying parasite types, and assessing parasitemia levels to gauge infection severity and monitor treatment efficacy. Traditional microscopic diagnosis, which remains the gold standard, involves staining and examining blood smears under a microscope to identify malaria parasites. Rapid Diagnostic Tests (RDTs), which can deliver results within minutes using a dipstick or cassette format, offer a viable alternative in settings lacking reliable microscopy. Molecular methods like polymerase chain reaction (PCR) are more sensitive than microscopy and are useful for confirming parasite types after initial diagnosis but are less practical for immediate clinical decision-making due to slower result availability. While microscopy is the most established method, allowing for detailed assessment of parasite types and quantification of parasitemia, it requires significant training and manual labor. RDTs, being less resource-intensive, are increasingly used in areas with limited access to microscopy, though they are not yet a replacement but rather complementary to microscopic tests.

6. ADVANTAGES

- High Accuracy: CNNs can achieve exceptional levels of accuracy in detecting malaria parasites within blood cell images, minimizing false positives and negatives compared to traditional methods.
- Efficiency: Automated detection using CNNs significantly reduces the time required for diagnosis, enabling faster turnaround times for test results and subsequent treatment.
- Reduced Reliance on Manual Examination: CNN- based systems decrease the dependence on skilled personnel for manual examination, mitigating issues related to human error and variability in interpretation.
- Adaptability to Variations: CNNs can be trained on diverse datasets, allowing them to adapt to variations in image quality, staining techniques, and sample characteristics commonly encountered in real-world clinical settings.
- Scalability: These systems can process large volumes of images efficiently, making them suitable for high-throughput screening and population-level surveillance.
- Enhanced Accessibility: Automated malaria detection using CNNs can be deployed in remote or resource-limited areas where access to skilled healthcare professionals may be limited, improving access to timely diagnosis and treatment.
- Improved Patient Outcomes: Timely and accurate diagnosis facilitated by CNN-based systems can lead to better patient outcomes, including prompt initiation of appropriate treatment and reduced risk of disease complications.
- Potential for Integration with Mobile Technology: Integration of CNN-based malaria detection systems with mobile devices could enable point-of- care testing and remote diagnosis, further enhancing accessibility and reach.

7. DISADVANTAGES

- Complexity of Implementation: Implementing CNN-based systems for malaria cell detection requires specialized expertise in machine learning and image processing, which may pose challenges for healthcare facilities lacking in-house technical resources.
- High Initial Setup Costs: The development and deployment of CNN-based systems entail

- significant initial setup costs, including the acquisition of hardware, software, and training data, which may be prohibitive for resource- constrained settings.
- Dependency on Quality of Training Data: The performance of CNNs heavily relies on the quality and diversity of the training data. Biased or insufficient training data can lead to suboptimal performance and reduced accuracy in malaria cell detection.
- Potential for Overfitting: CNNs may exhibit overfitting, wherein the model learns to memorize the training data rather than generalize patterns. This can result in poor performance on unseen data and reduce the reliability of malaria detection results.
- Sensitivity to Variations in Image Conditions: CNN-based systems may be sensitive to variations in image conditions such as lighting, staining techniques, and sample preparation methods, potentially leading to inconsistent performance across different settings.
- Limited Interpretability: The internal workings of CNNs are often considered as "black boxes," making it challenging to interpret how the model arrives at its predictions. This lack of interpretability may hinder trust and acceptance of the system among healthcare providers and patients.
- Ethical and Regulatory Considerations: The deployment of CNN-based systems for medical diagnosis raises ethical and regulatory concerns regarding patient privacy, data security, and liability in the event of misdiagnosis or adverse outcomes.
- Maintenance and Updates: Continuous maintenance and updates are necessary to ensure the optimal performance and reliability of CNN- based systems over time. This requires ongoing investment in resources and infrastructure, which may be unsustainable for some healthcare facilities.

8. ACKNOWLEDGEMENT

I would like to thank my project guide “Mrs. Jyoti”, Assistant Professor, Department of Computer Science & Engineering, Institute of Technology and Management, GIDA, Gorakhpur, U.P. for his valuable guidance and suggestions. I would like to thank my project coordinator “Mr. Nitin Dixit”, Associate Professor, Department of

Computer Science & Engineering, Institute of Technology and Management, GIDA, Gorakhpur, U.P. for his valuable guidance and suggestions. I am thankful for his/her continual encouragement, support, and invaluable suggestions. Without his encouragement and guidance, this project would not have been materialized. Throughout the writing of the project, I have received a great deal of supportand assistance.

I am very thankful to HOD "Mr. Ashutosh Kumar Rao" Department of Computer Science and Engineering, for his kind cooperation. I would also like to thank to the Honorable Director "Dr. N. K. Singh" for his kind help and support. I would also like to thanks to all my friends who continuously supported me.

I want to express my appreciation to every person whocontributed either with inspirational or actual work to this project. Finally, I must express my very profound gratitude to my parents. Thanks to all.

9. CONCLUSION

The malaria cell detection model has been successfully developed using various machine learning techniques and algorithms. The statistical comparisons between the transfer learning model and the scratch model were also analyzed.

The model built from scratch provided 90% accuracy in both the test and verification set.

The model built using transfer learning provided 91% accuracy in the result and approximately 73% in the validation set.

We can adjust the accuracy of the model using a different transmission model or using a different algorithm.

Comparisons with existing methods

Early detection of a malaria cell is performed using Mathematical morphology that is referred as a thesis of local structural experimentation. It is known as morphology as it points to analyze the state of objects. MM is not just atheory, but also a very significant way of analyzing the image.

In addition, there have been models built using TensorFlow but our model not only builds a model from scratch but also uses learning transfer to build another model and measure how it compares to a scratch model. It gives us a statistical analysis of both models.

Our model therefore not only detects malaria but also provides statistical analysis of two different methods of statement of the same problem.

10. REFERENCES

[1] Loddo A., Putzu L., Di Ruberto C., Fenu G. A Computer-Aided System for Differential Count from Peripheral Blood Cell Images; Proceedings of the 2016 12th International Conference on Signal- Image Technology Internet-Based Systems (SITIS); Naples, Italy. 28 November–1 December 2016; pp. 112–118. [Google Scholar]

[2] Di Ruberto C., Loddo A., Putzu L. A leucocytes count system from blood smear images: Segmentation and counting of white blood cells based on learning by sampling. Mach. Vis. Appl. 2016;27:1151– 1160. doi: 10.1007/s00138-016-0812-4. [CrossRef] [Google Scholar]

[3] World Health Organization (WHO) Malaria Fact Sheet December 2016. [(accessed on 6 March 2017)]; Available online: http://www.who.int/mediacentre/factsheets/fs094/en/

[4] Somasekar J. Computer vision for malaria parasite classification in erythrocytes. Int. J. Comput. Sci. Eng. 2011;3:2251–2256. [Google Scholar]

[5] Soille P. Morphological Image Analysis: Principles and Applications. Springer; Berlin, Germany: 2004. p. 392. [Google Scholar]

[6] Giardina C., Dougherty E. Morphological Methods in Image and Signal Processing. Prentice-Hall, Inc.; Upper Saddle River, NJ, USA: 1988. [Google Scholar]