International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 14 • Issue 1 pp. 15-21 Sept 2022 - March 2023 www.csjournals.com

Malaria Parasite Detection using Image Processing and CNN Models

M. Vinay Kumar[1], M. Chaitanya[2], K. Sai Vivek[3]

.

[1,2,3]Department of Computer Science and Engineering at Raghu Engineering College, Vishakhapatnam, India.

Abstract:

Malaria parasite is a serious and threatening disease transmitted from mosquitoes, caused by Plasmodium parasites. The conventional diagnostic method involves visual examination of blood smears under a microscope, which is time-consuming and relies on the expertise of the technician. Although automatic image recognition techniques have been explored to diagnose malaria, the performance has been suboptimal. Therefore, our objective is to develop a model that utilizes image processing to identify infected red blood cells rapidly and accurately in thin blood smears on standard microscope slides. Our goal is to improve malaria diagnosis by automating the process with early and effective testing. As no specific results were mentioned in the initial paragraph, we cannot provide any specific results. However, we aim to achieve improved accuracy, efficiency, and speed in the malaria diagnostic process.

Keywords: CNN, Deep Learning, Malaria Detection, VGG19, FROZEN METHOD IN VGG19, FINE TUNING METHOD IN VGG19.

Introduction

Female Anopheles mosquito bites that introduce the Plasmodium parasite into the host are what cause malaria parasite sickness. The most harmful forms of malaria are P. falciparum and P. vivax, with P. falciparum being more common. There were 228 million cases of malaria worldwide in 2019, with 300–500 million cases reported annually, according to a World Health Organization (WHO) assessment. With 405,000 documented fatalities by the WHO in 2019, sub-Saharan Africa is where children die most frequently. The disease significantly affects national economy, especially in developing nations.

To diagnose malaria, a blood sample is examined; thin blood smears are used to determine the type of parasite present, and thick blood smears to determine whether it is present. Observer variability and human skill can both affect how accurately a diagnosis is made. Giemsa staining is a technique used by professionals in pathology labs to find and identify parasites in blood samples. RBCs and Plasmodium parasites are coloured during this staining process for additional investigation. Malaria is detected using a variety of methods, including manual examination, which is laborious and time-consuming, and PCR and RDT, which are quicker but less precise.

OBJECTIVE

The primary goal of this research is to develop a deep learning model for blood sample image analysis that can identify malaria parasites using the VGG19 architecture. The model will undergo training using a collection of blood sample pictures that includes both infected and uninfected samples. To make sure that the model is resistant to changes in sample quality and lighting conditions, the dataset will be pre-processed and enhanced. On a test set, the trained model will be assessed and contrasted with other cutting-edge models for the detection of malaria parasites. The project's end deliverable will be a deep learning model that can precisely identify malaria parasites in images of blood sample, which may be used for malaria early detection and treatment.

CONTRIBUTION

The suggested neural network-based approach for malaria parasite identification intends to enhance malaria diagnosis, especially in rural areas where access to qualified technicians is limited. The system is built on image analysis and supervised learning techniques that begin with pre-processing, move on to blob identification, and then use Google Net to extract features. Using the support vector machine (SVM) for classification enables the separation of infected from uninfected cells. This research makes a contribution by creating an automated approach for diagnosing malaria that is more accurate and reduces human error. Additionally, the use of machine learning techniques allows for more efficient and timely diagnosis, leading to better treatment outcomes. Overall, the proposed system offers a promising approach to combat malaria, which remains a major global health concern.

LITERATURE SURVEY

This section discusses the utilization of different techniques, datasets, and classification algorithms to forecast the presence or absence of malaria parasites in a blood sample.

The goal of Adedeji Olugboja and Zenghui Wang's 2017 study, "Malaria parasite detection using different machine learning classifier," was to develop an automated system for detecting plasmodium parasite in stained blood sample photographs. The authors discuss several diagnostic techniques for determining the presence of malaria, including the traditional microscopy approach, the rapid diagnostic test (RDT) detection technique, and the polymerase chain reaction method. (PCR). They underline how important the light microscope is as the standard technique for identifying malaria. The study describes the method for detecting malaria parasites, including image preprocessing, segmentation, feature extraction, and classification. The authors evaluate a number of classifiers and find that Subspace KNN outperforms Fine Gaussian SVM in terms of overall performance.

To decrease human error and improve accuracy, Shipra Saraswat, Utkarsh Awasthi, and Neetu Faujdar [2] (2017) suggested an automated approach for malaria identification using image processing. This study uses pre-processing, morphological procedures, and HSV segmentation and is based on edge detection approaches. An 80-image dataset of thin blood smears was used by the authors to evaluate the system. This paper mentions the difficulty of the sensitivity of tiny objects to illumination and changes in microscope settings. The authors propose that future research in this area should concentrate on creating new techniques and parameters to enhance microscopic image analysis.

Saqlain Razzaq, Muhammad Adeel Asghar, Muhammad Jamil Khan, IlsaRameen, Ayesha Shahadat, and Mehwish Mehreen. [3] (2021) In "Leveraging Supervised Machine Learning Techniques for Identification of Malaria Cells using Blood Smears," supervised learning methods are used to identify and categorise malaria-infected cells. The study makes use of deep convolutional neural network GoogleNet for feature extraction and thresholding in image processing. The authors use Otsu's thresholding method and emphasise how crucial image processing is for accurately detecting malaria. The suggested approach additionally employs ensemble machine learning and SVM algorithms for classification. Overall, the study makes a significant contribution to medical image analysis for the identification of malaria.

Rahul Mapari, Divyansh Shah, Khushbu Kawale, Masumi Shah, Santosh Randive [4] (2020) By employing patient blood cell pictures and a CNN model with a transfer learning strategy, the paper Malaria Parasite Detection employing Deep Learning by Divyansh Shah et al. (2020) demonstrates how to identify malaria using this parasite. The writers draw attention to the widespread presence of malaria and its serious effects. In addition to highlighting the value of early detection, they also highlight the efficiency of deep learning models for image processing and classification tasks. Three convolutional layers make up the proposed CNN model, which has a detection accuracy of about

95% overall. The authors come to the conclusion that their model offers noticeably good accuracy and is similar to other models with more layers.

PRE-PROCESSING METHODS

Image Processing

In deep learning, image processing is a crucial element, particularly for tasks involving image analysis. To classify or predict, deep learning models employ various techniques to process images and extract relevant features.

Using Convolutional Neural Networks (CNNs) to extract features, pooling layers to preserve key features, activation functions like ReLU to introduce non-linearity, dropout for regularisation to prevent overfitting, transfer learning to fine-tune pre-trained models, object detection to identify and classify objects, semantic segmentation to label image regions, and generative adversarial networks (GANs) are some common methods used in image processing for deep learning.

Overall, image processing is a fundamental component of deep learning for image analysis tasks. It involves using different techniques and architectures to extract relevant features, identify objects, segment images, and synthesize new images, among other tasks. These techniques enable deep learning models to perform complex tasks, achieve high accuracy, and improve the overall performance of image analysis.

Image Segmentation

Image segmentation is a technique used to divide a digital image into multiple segments or regions, each comprising sets of pixels. This process helps in simplifying or altering the image's representation, making it more interpretable and convenient to analyze. The primary aim of segmentation is to partition an image into meaningful regions, based on its properties like intensity, color, or texture, and extract important information from it.

Feature Extraction

Feature extraction involves converting raw data into numerical features that can be analyzed without losing any critical information. Deep Learning is a technique that employs neural networks to interpret data features and relationships by passing data through multiple processing layers. The CNN model is ideal for detecting malaria in patient blood cell images since it takes an image as input and processes it through various layers. For the implementation of the project, we have chosen the Python programming language.

Dataset Description

The Malaria Cells Image Dataset is a collection of microscopic images of blood cells infected with the malaria parasite. The dataset includes images of both the malaria parasite (Plasmodium) and the red blood cells they infect. The images were acquired using brightfield microscopy and are annotated with information about the parasite species and the severity of the infection.

The dataset is commonly used in computer vision research and machine learning to develop algorithms for automatically detecting malaria infection in blood cell images. It can also be used for other tasks such as image segmentation and object detection. The dataset is publicly available and has been used in a variety of research studies, including the development of deep learning models for malaria diagnosis.

Data Pre-processing

Data pre-processing is an essential step in deep learning that involves transforming raw data into a format that is suitable for training a machine learning model. Data pre-processing aims to ensure that the input data is consistent, standardized, and well-structured, which can improve the accuracy and efficiency of the model?

There are several common techniques used in data pre-processing for deep learning:

Data cleaning: This involves removing any errors, missing values, or outliers from the data.

Data normalization: This involves rescaling the data to ensure that all features are on a similar scale.

Data encoding: This involves converting categorical data into numerical data, typically using one-hot encoding.

Data augmentation: This involves generating new training data by applying various transformations such as rotation, scaling, or cropping.

Data splitting: To assess the model's performance, the data are divided into training, validation, and testing sets.

Data pre-processing is a critical step in deep learning that can significantly impact the accuracy and efficiency of a machine learning model. It is important to carefully consider the specific needs of the dataset and the deep learning model when selecting and applying pre-processing techniques.

ALGORITHMS USED

CONVOLUTIONAL NEURAL NETWORKS:

Deep neural networks of the sort known as convolutional neural networks (CNNs) are frequently employed for image and video analysis. Animal visual cortex serves as the biological model for CNNs, which can automatically learn hierarchical representations of visual data.

In-order to extract features like edges, forms, and patterns from the input image, CNNs use convolution operations on many layers of filters. The next step is to run these features through more layers of filters and pooling processes so that the image can gradually learn more sophisticated representations.

CNNs have become popular in computer vision due to their ability to learn features automatically from large datasets, leading to superior performance on a range of image analysis tasks such as object detection, segmentation, and classification.

VGG19:

VGG is a deep convolutional neural network with 19 layers, known for its superior performance on image recognition tasks. It can classify images into 1000 categories, and a pretrained version is available with over a million images from the ImageNet database. VGG19 has three additional convolution layers compared to VGG16, but both networks have the same pooling and fully connected layers. VGG19 has 138 million parameters and an accuracy of 92.7%. It uses small 3x3 filters with a stride of 1, resulting in a more discriminative decision function with three ReLU units.

FROZEN METHOD IN VGG19:

Freezing a layer in neural networks restricts the modification of its weights and is used to accelerate training time without sacrificing accuracy. This technique is commonly implemented to progressively freeze hidden layers and prevent the backward pass to that layer, resulting in a significant speed boost. Freezing is often utilized in transfer learning when using a pre-trained base model on another dataset.

FINE TUNING METHOD IN VGG19:

Using pre-trained architecture and weights, fine-tuning is a versatile and effective technique for building large-scale models. This procedure, also known as transfer learning, can build reliable models with smaller datasets and shorter training times by training only specific layers on fresh data. A pre-trained model is fine-tuned by applying it to a particular job on a fresh dataset. By initialising the weights to those of the tissue-classification network, the existing network can be reused rather than being trained from start. To increase the precision of machine learning predictive models, fine-tuning is a vital step.

RESULT AND ANALYSIS:

We have used Jupyter software to implement the simulation and to display the outputs.

OUTPUT AFTER PRE-PROCESSING AND SCALING OF IMAGES:

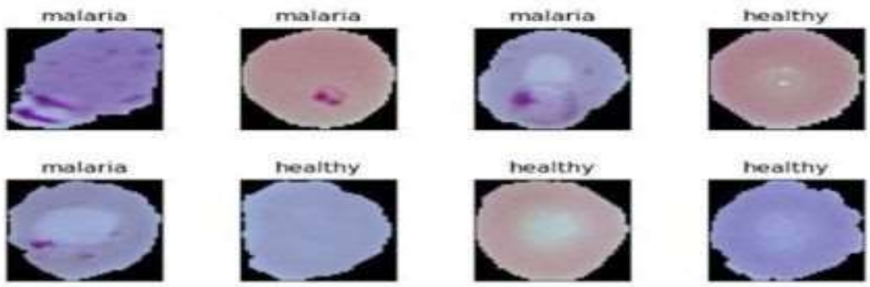

Fig (b): Pre-processing and scaling output

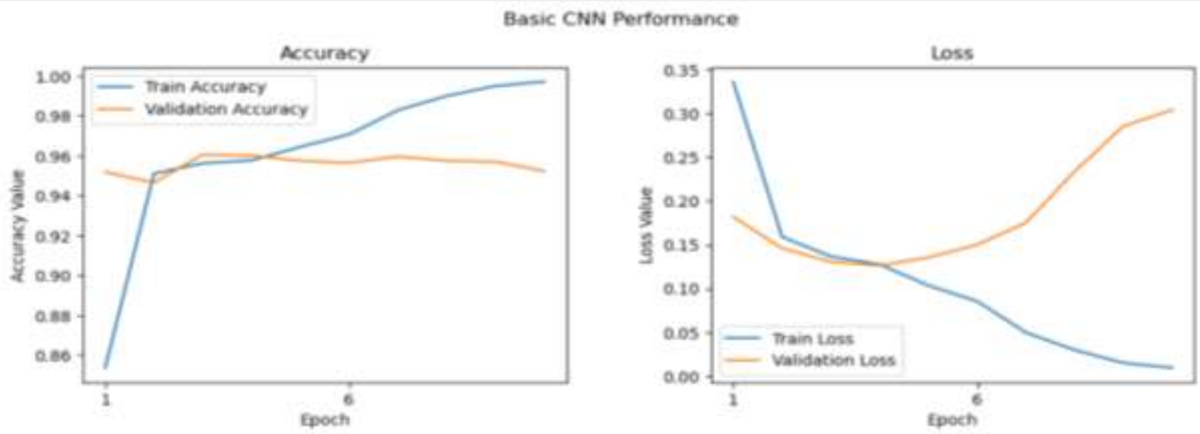

Fig (c): Performance Graph of basic CNN

International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 14 • Issue 1 pp. 15-21 Sept 2022 - March 2023 www.csjournals.com

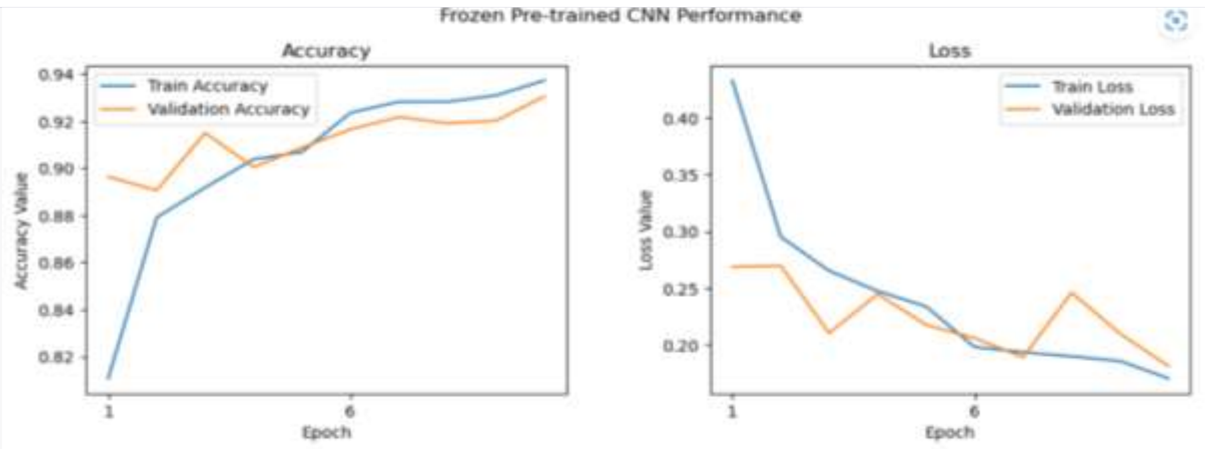

Fig (d): Performance Graph of VGG-19 Frozen

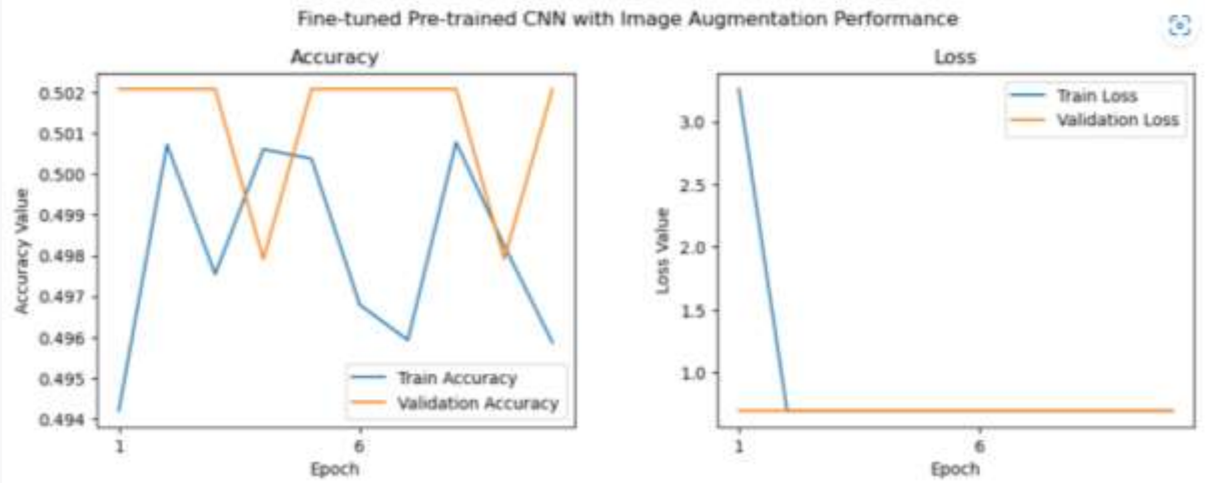

Fig (e): Performance Graph of VGG-19 Fine-tuned

	Accuracy
CNN Model from Scratch	0.9523
VGG-19 Frozen	0.9305
VGG-19 Fine-Tuned	0.5021

Fig (f): Accuracy Comparison Results of three Methods

CONCLUSION:

A system that utilizes image processing and deep learning techniques to reduce the likelihood of human error in detecting Malaria Parasites. The system employs a convolutional neural network and labelled datasets to create an image classification model that achieves an accuracy of approximately 95%. It is a robust system that can withstand external factors that may impact it. The creators believe that with more computing power and an improved model, even better results could be achieved. Additionally, the system's potential for expansion is noted, and there is potential to employ the system for diagnosing other diseases by analysing blood samples. The proposed model's accuracy is impressive, and it utilizesthe model's accuracy is attributed to the successful integration of convolutional and dense layers. Interestingly, existing models with many layers provide accuracy comparable to that obtained by this model. Overall, the proposed model's results are quite promising, and it demonstrates the potential for using deep learning techniques to improve disease diagnosis.

REFERENCES:

[1] Adedeji Olugboja, Zenghui Wang, "Malaria Parasite Detection Using Different Machine Learning Classifier" Proceedings of the 2017 International Conference on Machine Learning and Cybernetics, Ningbo, China, 9-12 July, 2017.

[2] Divyansh Shah, Khushbu Kawale, Masumi Shah, Santosh Randive, Rahul Mapari, "Malaria Parasite Detection Using Deep Learning (Beneficial to humankind)" Proceedings of the International Conference on Intelligent Computing and Control. Systems (ICICCS 2020) IEEE Xplore Part Number: CFP20K74-ART; ISBN: 978- 1-7281-4876-2.

[3] Shipra Saraswat, Utkarsh Awasthi, Neetu Faujdas, "Malarial Parasites Detection in RBC Using Image Processing", International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), Sep 20- 22, 2017, AIIT, Amity University, Uttar Pradesh.

[4] IlsaRameen, Ayesha Shahadat, Mehwish Mehreen, Saqlain Razzaq, Muhammad Adeel Asghar, Muhammad Jamil Khan, "Leveraging Supervised Machine Learning Techniques for Identification of Malaria Cells Using Blood Smears", 2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2).

[5] WorldHealthOrganizationWorldMalariaReporthttps://www.who.int/news-room/factsheets/detail/malaria.

[6] Poostchi M, Silamut K, Maude RJ, Jaeger S,ThomaG(2018) Image analysis and machine learning for detecting malaria. Trans Res: J Lab Clin Med 194:36–55.

[7] A Review of Popular Deep Learning Architectures:ResNet, InceptionV3, and SqueezeNet.https://blog.paperspace.com/populardeeplearningarchitectures-resnetinceptionv3-squeezenet/

[8] https://www.mygreatlearning.com/blog/introduction-to-image-pre-processing/

[9] https://www.tutorialspoint.com/flask/index.html

[10] https://docs.opencv.org/4.x/d0/de3/tutorial_py_intro.html

[11] https://www.geeksforgeeks.org/os-module-python-examples/

[12] https://www.tensorflow.org/learn

[13] https://pillow.readthedocs.io/en/stable/

[14] https://pypi.org/project/Werkzeug/