

Classification and Analysis of Malaria Infected Blood Cells

[1] Khaja Tabarak Uddin, [2] Khaja Bilal Ahmed, [3] Saif Ul Islam, [4] Dr. Mohammed Jameel Hashmi

[1] BE Student – Dept. Of CSE, ISL Engineering College, TS, India

[2] BE Student – Dept. Of CSE, ISL Engineering College, TS, India

[3] BE Student – Dept. Of CSE, ISL Engineering College, TS, India

[4] Professor – Dept. Of CSE, ISL Engineering College, TS, India

Article Info

Page Number: 1745-1752

Publication Issue:

Vol. 72 No. 1 (2023)

Abstract

A contagious illness An viral disease called malaria kills more than 500,000 people annually throughout the world. The majority of these fatalities are brought on by a delayed or inaccurate diagnosis. The manual microscope is now thought to be the best technology for diagnosing malaria. On the other hand, it takes time and is subject to human mistake. It is crucial that the evaluation procedure be automated because it is such a significant issue for world health. This article's goal is to promote automation of the diagnosis procedure in order to do away with the necessity for human involvement. The strategy is founded on known erythrocyte and Plasmodium parasite intensity characteristics to be erratic. The same dataset is used to train the transfer learning models CNN, densnet201, Nasnet large, InceptionNet, Xception, Hybrid (CNN + DenseNet201 + NasNet Large + InceptionNet + Xception), KNN, SVM, Mobilenet, VGG16, Resnet50, InceptionV3, Densenet169, Resnet101, Lenet, and efficientnet V2S. Both transfer learning and fine-tuning strategies are used, and the results are contrasted. Experimental evidence demonstrates that the suggested models operate well and generate reliable outcomes.

Keywords: Malaria infection detection, deep learning, machine learning.

Article History

Article Received: 15 October 2022

Revised: 24 November 2022

Accepted: 18 December 2022

1. Introduction

Malaria is spread through the bites of female Anopheles mosquitoes infected with Plasmodium protozoan parasites, which infect red blood cells and cause them to swell and swell up. Every year, 3.2 billion people worldwide are at high risk of developing malaria, according to the World Health Organization. According to a survey conducted by the World Health Organization [1], 91 countries recorded 216 million cases of malaria. The World Health Organization is a non-governmental organization that promotes health worldwide. Global malaria cases were primarily concentrated in the African Region, which was then followed by the Southeast Asia Region and the Eastern Mediterranean Region. The symptoms of malaria are often associated with fever, tiredness, headaches, and, in extreme cases, seizures and coma, all of which can be fatal if not treated promptly. Malaria is a preventable disease that can be controlled with adequate treatment. There is, however, no effective immunization available at this time. Once infected, it is a disease that progresses at a rapid pace. Malaria is a significant load on our healthcare system, and it is the top cause of death in many developing and developing-country populations. It is endemic in many parts of the world, which means that the disease is met on a

regular basis in those areas of the world. As a result, early detection and treatment of malaria are essential in order to save lives. Because of this, we are motivated to increase the effectiveness and timeliness of malaria diagnostics in the future. Specialized technology is required in order to resolve this problem. As a result, it is vital to obtain a prompt diagnosis. The most important task in diagnosing malaria is to determine whether or not parasites are present. The most common method of diagnosing malaria is by the use of a blood sample. In the United States alone, millions of blood samples are tested for malaria each year, with a trained pathologist painstakingly counting parasites and infected red blood cells in each sample. According to the World Health Organization regulation [2], the blood smear should be inspected under a microscope at a magnification of 100x. Diagnostic treatments such as light microscopy and rapid diagnostic tests are two of the most often performed (RDT). The use of these two tests is typical in situations where high-quality microscopy services are not readily available. However, there are several disadvantages to using these procedures, including the fact that the diagnosis is primarily dependent on the pathologist's knowledge and skill, the possibility of false-positive and false-negative diagnoses, which can result in the development of other illnesses, and the fact that they are time-consuming, to name a few.

2. Literature Survey :

Fast analysis in routine histological samples was achieved by utilizing heterogeneous image segmentation algorithms. Most algorithms are rarely dedicated to classification stage and largely to analysis which led to poor detection sensitivities and early detection of parasite life stage. Fast analysis in routine histological samples was achieved by utilizing heterogeneous image segmentation algorithms. Most algorithms are rarely dedicated to classification stage and largely to analysis which led to poor detection sensitivities and early detection of parasite life stage [4]. Cecilia Di Ruberto et al. [5] utilized RGB images and segmented erythrocytes restricted by size obtained from grey scale granulometry analysis using watershed algorithm. Harsdorf distance and Morphological operations were utilized for schizonts classification.

Ross et al. [6] proposed similar technique except for enhancement of early life stage detection using local threshold. A two-stage machine learning technique was used in this work for differentiating parasite species through classification. Positive and negatives were classified in the first stage utilizing the image features like color and texture. The second stage involved assignment of infected erythrocytes based on parasite feature like shape, size and color. Malaria evaluation utilized mathematical morphology. Segmentation of erythrocytes with application of threshold function is achieved by applying edge filter to color of image. Erythrocyte infection was decided based on association factor between edges and sequential application of binary morphological operations.

Detection of erythrocytes using template matching technique was proposed by Halim et al. [7] where original images were employed to construct grey scale template and associating it with predefined binary template. A 3x3 matrix is constructed using the eight-level color quantized image. Parasite classification using pixels is done based on threshold of cooccurrence matrix. Machine Learning is the used to detect the parasite infection.

Later pixels are clustered into objects as parasite by utilizing many traditional mathematical morphology operators and k-nearest neighbor classifier is employed to determine non parasites.

3. Proposed System:

This article's goal is to promote automation of the diagnosis procedure in order to do away with the necessity for human involvement. The strategy is founded on the intensity characteristics of erythrocytes and Plasmodium parasites, both of which are known to be changeable. The same dataset is used to train the transfer learning models CNN, densnet201, Nasnet large, InceptionNet, Xception, Hybrid (CNN + DenseNet201 + NasNet Large + InceptionNet + Xception), KNN, SVM, Mobilenet, VGG16, Resnet50, InceptionV3, Densenet169, Resnet101, Lenet, and efficientnet V2S. Both transfer learning and fine-tuning strategies are used, and the results are contrasted. Experimental evidence demonstrates that the suggested models operate well and yield reliable results.

The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.

The data flow diagram (DFD) is one of the most important modeling tools. It is used to model the system components. These components are the system process, the data used by the process, an external entity that interacts with the system and the information flows in the system.

DFD shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.

DFD is also known as bubble chart. A DFD may be used to represent a system at any level of abstraction. DFD may be partitioned into levels that represent increasing information flow and functional detail.

3.1 Modules:

3.1.1 Tensorflow

TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks. It is used for both research and production at Google.

TensorFlow was developed by the Google Brain team for internal Google use. It was released under the Apache 2.0 open-source license on November 9, 2015.

3.1.2 Numpy

Numpy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays.

It is the fundamental package for scientific computing with Python. It contains various features including these important ones:

A powerful N-dimensional array object

Sophisticated (broadcasting) functions

Tools for integrating C/C++ and Fortran code

Useful linear algebra, Fourier transform, and random number capabilities

Besides its obvious scientific uses, Numpy can also be used as an efficient multi-dimensional container of generic data. Arbitrary data-types can be defined using Numpy which allows Numpy to seamlessly and speedily integrate with a wide variety of databases.

3.1.3 Pandas

Pandas is an open-source Python Library providing high-performance data manipulation and analysis tool using its powerful data structures. Python was majorly used for data munging and preparation. It had very little contribution towards data analysis. Pandas solved this problem. Using Pandas, we can accomplish five typical steps in the processing and analysis of data, regardless of the origin of data load, prepare, manipulate, model, and analyze. Python with Pandas is used in a wide range of fields including academic and commercial domains including finance, economics, Statistics, analytics, etc.

3.1.4 Matplotlib

Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. Matplotlib can be used in Python scripts, the Python and IPython shells, the Jupyter Notebook, web application servers, and four graphical user interface toolkits. Matplotlib tries to make easy things easy and hard things possible. You can generate plots, histograms, power spectra, bar charts, error charts, scatter plots, etc., with just a few lines of code. For examples, see the sample plots and thumbnail gallery.

For simple plotting the pyplot module provides a MATLAB-like interface, particularly when combined with IPython. For the power user, you have full control of line styles, font properties, axes properties, etc, via an object oriented interface or via a set of functions familiar to MATLAB users.

3.1.5 Scikit – learn

Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python. It is licensed under a permissive simplified BSD license and is distributed under many Linux distributions, encouraging academic and commercial use.

5. Implementation

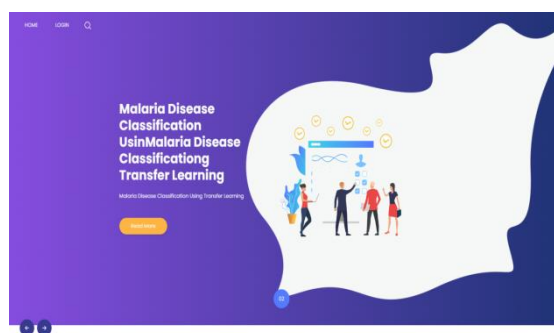


Fig 1: Landing page

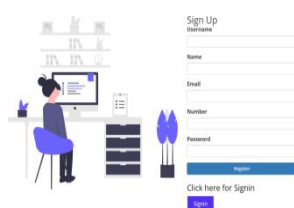


Fig 2: Form

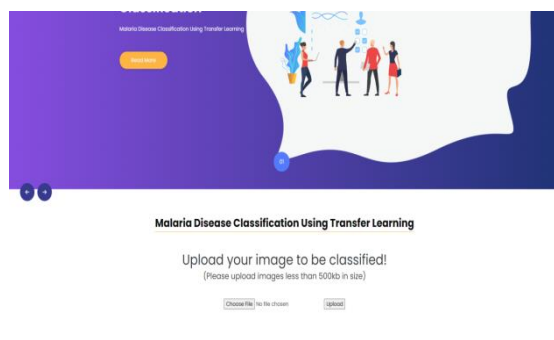


Fig 3 : Website Screenshot

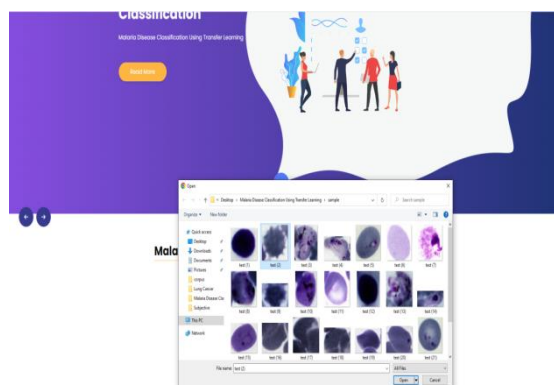
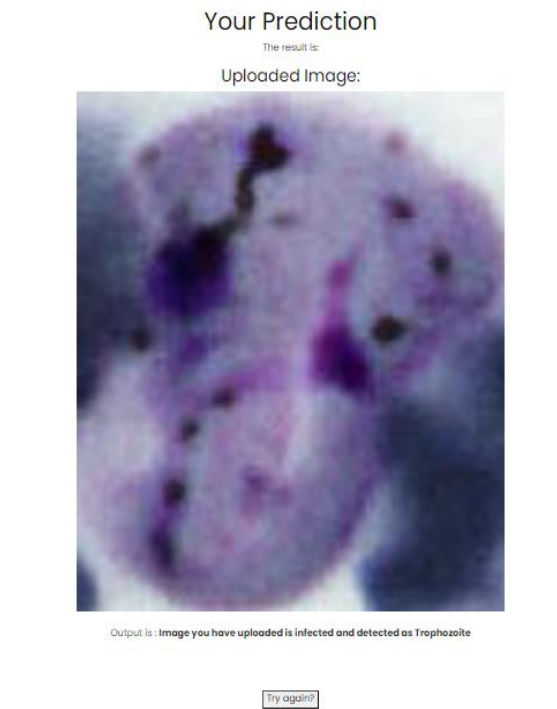


Fig 4: Working Screenshot

Malaria Disease Classification Using Transfer Learning**Fig 5: Result****4. Conclusion and Future Work:**

To increase the performance of malaria diagnosis categorization in this study, we applied end-to-end deep-learning neural networks from start to finish. Deep learning helps computers find meaningful links in a large amount of data and make sense of unstructured data. Transfer learning, on the other hand, is a machine learning research problem that puts an emphasis on loading knowledge gained while trying to solve one problem and implementing it to a different but related problem. Based on our Fastai experience, we believe that using a layered API in deep learning can provide significant research benefits to the community. Compared to custom-built CNN models, it is, however, far more compatible with predefined architectures (for example, ResNet, Inception, and so on). Based on the simulation findings, it was proved that these deep-learning algorithms were capable of reaching extraordinarily high accuracy in pattern recognition. Based on our experimental findings, we conclude that the pretrained convolutional neural network model VGG-19 performs significantly better than ResNet-50, ResNet-34, and VGG-16 for the classification of blood smears. We employed transfer learning and fine-tuning to increase the performance of these pretrained models, and the results were promising. Their performance is influenced by the architecture, the training framework, and the volume of training data that they are given. In order to avoid information loss from images, the dropout approach was not employed in this study. We developed a web-based interface to make it easier for the end-user to use this model and categorize blood smear photos, which is now under development. This has the potential to reduce stress on medical field workers while simultaneously boosting the speed with which diagnoses are made. In the future, we want to concentrate on improving the performance of the CNN models by optimizing their architecture,

which will result in a significant rise in the accuracy of malaria detection. Mobile devices and cloud-based implementation are also options for extending the end-user application's functionality.

REFERENCES:

1. WHO, "Fact sheet: World Malaria Report 2016," in World Health Organization, world Health Organization, 2016. [Online]. Available: <http://www.who.int/malaria/media/world-malaria-report-2016/en>. Accessed: Jan. 25, 2017
2. Bloland PB (2001) Drug resistance in Malaria, WHO/CDS/ CSR/DRS/ 2001.4. World Health Organization, Switzerland, 2001
3. D Ghatge, C. Jadhav, and N. U. Rani, "AUTOMATIC DETECTION OF MALARIA PARASITE FROM BLOOD IMAGES,". [Online]. Available: <http://ijact.org/volume4issue1/IJ0410050.pdf>. Accessed: Jan. 25, 2017
4. Ahirwar A., Pattnaik S., Acharya B., Advanced Image Analysis Based System for Automatic Detection and Classification of Malarial Parasite in Blood Images. International Journal of Information Technology and Knowledge Management. Jan-June 2012, Volume 5, No 1, pp. 59-64
5. Amit Kumar, Prof. Arun Choudhary, Prof. P. U. Tembhare, Prof. C. R. Pote. Enhanced Identification of Malarial Infected Objects using Otsu Algorithm from Thin Smear Digital Images. International Journal of Latest Research in Science and Technology Vol.1, Issue 2: Page No. 159-163, July-August (2012).
6. Syed Shehriyar Ali, Mohammed Sarfaraz
7. Shaikh, Syed Safi Uddin, Dr. Mohammed Abdul Bari, "Saas Product Comparison and Reviews Using Nlp", Journal of Engineering Science (JES), ISSN NO:0377-9254, Vol 13, Issue 05, MAY/2022
8. Hafsa Fatima, Shayesta Nazneen, Maryam Banu, Dr. Mohammed Abdul Bar," Tensorflow-Based Automatic Personality Recognition Used in Asynchronous Video Interviews", Journal of Engineering Science (JES), ISSN NO:0377-9254, Vol 13, Issue 05, MAY/2022
9. Mohammed Shueb, Mohammed Akram Ali, Mohammed Shadeel, Dr. Mohammed Abdul Bari, "Self-Driving Car: Using Opencv2 and Machine Learning", The International journal of analytical and experimental modal analysis (IJAEMA), ISSN NO: 0886-9367, Volume XIV, Issue V, May/2022
10. Mr. Pathan Ahmed Khan, Dr. M.A Bari,: Impact Of Emergence With Robotics At Educational Institution And Emerging Challenges", International Journal of Multidisciplinary Engineering in Current Research(IJMEC), ISSN: 2456-4265, Volume 6, Issue 12, December 2021,Page 43-46
11. Mohammed Abdul Bari, Shahanawaj Ahamad, Mohammed Rahmat Ali," Smartphone Security and Protection Practices", International Journal of Engineering and Applied Computer Science (IJEACS) ; ISBN: 9798799755577Volume: 03, Issue: 01, December 2021

12. Rohokale, M. S., Dhabliya, D., Sathish, T., Vijayan, V., & Senthilkumar, N. (2021). A novel two-step co-precipitation approach of CuS/NiMn₂O₄ heterostructured nanocatalyst for enhanced visible light driven photocatalytic activity via efficient photo-induced charge separation properties. *Physica B: Condensed Matter*, 610 doi:10.1016/j.physb.2021.412902
13. Sai Pandiraju, T. K., Samal, S., Saravanakumar, R., Yaseen, S. M., Nandal, R., & Dhabliya, D. (2022). Advanced metering infrastructure for low voltage distribution system in smart grid based monitoring applications. *Sustainable Computing: Informatics and Systems*, 35 doi:10.1016/j.suscom.2022.100691
14. Sharma, R., & Dhabliya, D. (2019). A review of automatic irrigation system through IoT. *International Journal of Control and Automation*, 12(6 Special Issue), 24-29. Retrieved from www.scopus.com
15. Sharma, R., & Dhabliya, D. (2019). Attacks on transport layer and multi-layer attacks on manet. *International Journal of Control and Automation*, 12(6 Special Issue), 5-11. Retrieved from www.scopus.com
16. Shahanawaj Ahamad, Mohammed Abdul Bari, Big Data Processing Model for Smart City Design: A Systematic Review “, VOL 2021: ISSUE 08 IS SN : 0011-9342 ;Design Engineering (Toronto) Elsevier SCI Oct
17. Mohammed Abdul Bari, Shahanawaj Ahamad, Mohammed Rahmat Ali,” Smartphone Security and Protection Practices”, *International Journal of Engineering and Applied Computer Science (IJEACS)* ; ISBN: 9798799755577 Volume: 03, Issue: 01, December 2021 (International Journal,U K) Pages 1-6
18. Dr. M.A.Bari, ”EffectiveIDS To Mitigate The Packet Dropping Nodes From Manet “, *JACE*, Vol -6,Issue -6,June 2019
19. M.A.Bari & Shahanawaj Ahamad,” Process of Reverse Engineering of Enterprise InformationSystem Architecture” in *International Journal of Computer Science Issues (IJCSI)*, Vol 8, Issue 5, ISSN: 1694-0814, pp:359-365, Mahebourg ,Republic of Mauritius , September 2011
20. M.A.Bari & Shahanawaj Ahamad, “Code Cloning: The Analysis, Detection and Removal”, in *International Journal of Computer Applications(IJCA)*,ISSN:0975-887, ISBN:978-93-80749-18-3,Vol:20,No:7,pp:34 38, NewYork,U.S.A.,April 2011