Volume **1** (1): 2023, page 24-34, https://doi.org/10.59785/tjhest.v1i1.3 https://tamjed.com/

Parasitic Egg detection from Microscopic images using Convolutional Neural Networks

Dr.Savitha V[1]

[1]Department of Computer Science and Engineering, SNS College of Technology, India
Email: savicse@gmail.com; drsavithav98@gmail.com

Dr.Karthick.M[2]

[2]Faculty of Computing and Information Systems, University Technology, Malaysia (UTM). Malaysia
E-mail: waleed.alrahmi1@gmail.com

Dr.T.Kalaikumaran[3]

[3]Professor, Department of Artificial Intelligence and Data Science, VSB College of Engineering Technical Campus, Coimbatore, India
E-mail: drkalaikumaran@gmail.com ; tkalaikumaran@gmail.com

Abstract

The most common test for parasitic infection diagnosis is stool parasitological testing. The use of the Kato-Katz method in the preparation of slides for the development of the image bank discussed here was extremely important. Other authors' studies on the same topic were discussed. Various parasite eggs of various species were created. Then the binary and multiclass classifier architectures were empirically defined, and each model was implemented. The performance of the classifiers was evaluated using metrics recommended in the literature for both empirically defined and transfer learning architectures. Finally, experiments were conducted to improve the system's performance by allowing binary and multiclass models to communicate. This data was used to build a database of 66 parasite egg photos from various species. Using data augmentation techniques, a total of 48,000 photos were collected. The examined measures all reached 99.9%. Despite some species' eggs sharing morphological similarities, the second method correctly classified each egg with a 99.9% hit rate. The problems addressed received a 99.9% rating on the evaluation measures used to assess them. The method can also be applied to a larger number of helminth species and detection technologies using the same procedures.

Keywords: Parasitic infections, helminth species, Machine learning, Image Processing, early diagnosis, CNN.

1. Introduction

Parasites are organisms that need other organisms to survive, which are hosts. To stay alive, the parasites use the host's physiological resources to nourish themselves and be able to reproduce. Currently, it is estimated that there are 4 billion people infected with parasitic diseases worldwide, and in children and immunodeficient people infections can cause greater physical or behavioral disorders and, in the worst cases, lead to death [1]

Among these 4 billion, it is estimated that 700 million people are infected with the species *Diphyllobothrium latum* [2] and that schistosomiasis affects about 220 million people living in tropical and subtropical areas of 78 countries in Africa. , America and Asia [3]. Most of these infected people are asymptomatic, which presents a problem since the infected individual can transmit the disease to healthy people. There are three main classes of parasites that can cause disease in humans: protozoa, helminths, and ectoparasites. In previous studies, it was found that some helminths have a higher prevalence and can be found in positive stool samples, they are: *Fasciolopsis buski, Echinococcus granulosus, Diphyllobothrium latum, Fasciolopsis buski, Strongyloides stercoralis*and *Trichinella spiralis,* not necessarily in that order of prevalence. These same studies reported that the higher prevalence of these parasites is due to the form of infection of each one of them [5]. The diagnosis of a patient with suspected parasitological infection may vary according to the availability of the physician and the resources he has. The physician uses the patient's clinical signs and individual history to assess the need to order a diagnostic test for helminthiasis. In most cases, this diagnosis can be difficult, so it is very common for the professional to request more than one type of exam. Clinical examination is the first step towards diagnosis, with stool parasitological examination being the most common test used for this type of diagnosis. One of the most commonly used techniques in this type of examination is the Kato-Katz thick stool smear (Katz et al., 1972). This method has the advantages of being cheap and easy to obtain qualitative and quantitative results on the presence and parasite burden of the most common intestinal infections by helminths, also intestinal schistosomiasis. In addition, the Kato-Katz technique is a good test to detect people with few eggs in the stool, as it uses more fecal material on the slide compared to other methods. Trained specialists, based on their prior knowledge and with the aid of a microscope, examine the patient's fecal material in search of parasite eggs. It should be noted that this entire process is performed manually by the specialist, in which diagnostic errors are common due to tiredness, fatigue and lack of professional experience [6], resulting in false-negative rates, especially in cases where there is a low number of eggs in the material [8-10]

In addition, in endemic areas, government health systems carry out diagnostic actions on a sample of the population. In these cases, most individuals have a low load of parasite eggs, which makes the work of health agents even more difficult. In situations of this nature, the stool test is commonly used, since it is cheap compared to other tests and, therefore, is quite useful on a large scale. To solve the problem of the lack of trained specialists for correct decision making [11] and to reduce the time needed for diagnosis from the manual parasitological examination of feces, the development of technologies that may be able to automate this process. The automatic classification of parasite eggs in fecal examinations will allow the inspection of a greater number of samples with a high degree of reliability and objectivity. The technology will be useful, mainly, in countries where there is a high rate of people infected through parasitological diseases. In recent years, deep learning studies have been spreading and its applications have become increasingly present in society's daily life. Deep learning can be understood as a family of machine learning methods that are based on artificial neural networks (ANN). This type of learning can be supervised, semi-supervised or unsupervised.

2. Previous Works

Using a computer system that automatically analyses microscopic images, [12] were able to identify and classify intestinal parasites. An ANN and a neuro-fuzzy system are used in conjunction to segment data and train a classifier. Using a circular Hough transformation, the parasite is first identified and then segmented for analysis. The results show that each of the 20 parasite classes has a 100% success rate in classification..

It was discovered that an expert medical system might be used to automatically diagnose 20 different kinds of human intestinal parasitosis. It was built using a decision-making algorithm. Using information gathered from literature and clinicians, a database of parasite-related diseases was created. When a user answers a query, the system responds. Circular Hough transforms and a trained neuro-fuzzy classifier are used to cross-check the data. It was tested with 60 cases of infection and compared to the diagnosis of two experts. A 96.6 percent accuracy rate was achieved with 58 correct diagnoses. It was proposed by [13] that a microscope connected directly to a computer may be used to obtain images for diagnosing intestinal parasites. A contour detection approach based on wavelet transformations is used to detect the parasite. In order to execute parasite image segmentation and extraction, the active contours and the Hough transformation are used. A probabilistic neural network is used to classify the data. Intestinal parasite photos from 15 different species were used to evaluate the created method. The results reveal a 100% success rate in being recognised.

2.1 Problem of the present research work

Currently, parasitology techniques are all manual, and can be influenced by uncontrollable variables like the lab technician's attention and expertise. Parasite eggs are identified microscopically by a trained practitioner based on their shape. A trained graduate student earlier identified a parasitic egg in Figure 1(E). As shown in Figure 1(E), manually finding parasite eggs in faeces is not obvious. The dirt, mushrooms, water bubbles, etc. on the plate make it difficult to find the eggs. Another issue is the high rate of false negatives in this form of examination, which occurs when the specialist misses an egg in the faeces sample and misdiagnoses the patient.

2.2 Kato-Katz Method

The Kato-Katz method is a quantitative methodology used to determine a slide's parasite load, or the number of helminth eggs a person excretes. This method works with fresh or formaldehyde-preserved faeces. The preservative employed in preserved faeces must be eliminated at the time of examination. The Kato-Katz method is the major approach for diagnosing helminth presence and is presently regarded the sole technique utilised in routine examinations in public and private health facilities, as well as research institutes. Hookworm eggs are seen in Figure 1 (A-D) in a microscopic field (100x magnification). (A) Hook warm, *Diphyllobothrium latum* (B), *Fasciolopsis buski* (C) and *Trichinella spiralis* (D).

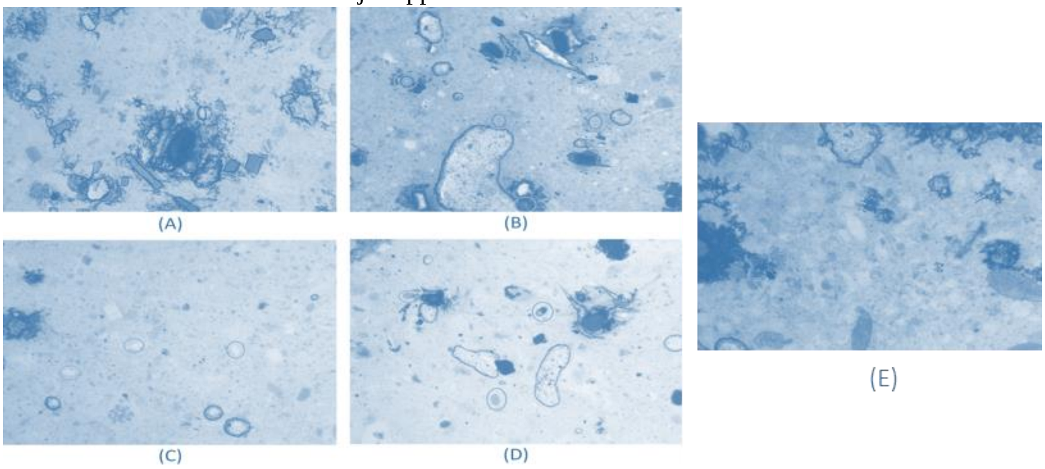


Figure 1 – Example of the use of the Kato-Katz technique in a microscopic field (Ato D) & E-Example of *Strongyloides stercoralis* egg in a microscopic field.

An apron, gloves, and blades are required to use this procedure. To make the slides, the sample must be fresh faeces (or refrigerated for up to 48 hours), not diarrhea. Using the resources provided, the laboratory technician will follow Barbosa et al's conventional protocol [9]. The technician can then use an optical microscope to see the helminth eggs in the faeces.

The procedure has significant limits, and most laboratory professionals' issues stem from the lack of eggs in the excrement. Even patients with the parasitic disease may not have a significant parasite burden due to certain circumstances, resulting in false-negative diagnoses.

2.3 Machine Learning

According to Michie et al. (1994), machine learning is an application of artificial intelligence (AI) that enables computer systems to learn and develop independently from their own experience, eliminating the need for explicit programming. Machine learning is the study of creating computer programs that can learn from data.

Starting with examples, the learning process begins to look for patterns in the data and make better future decisions based on the examples presented. With implicit parameter adjustment, the computer can handle classification or regression problems without human interaction.

Machine learning algorithms can analyze enormous data sets. While it is faster and more accurate to detect profitable opportunities or dangerous dangers, model training may take more time and resources. Machine learning, AI, and cognitive technologies can improve the efficiency of this family of algorithms in processing vast amounts of data.

2.3.1 Classification Problem

A classification task in machine learning involves correctly recognising a sample's class from a training dataset that comprises observations of the same known category. Examples of widely used examples of this type of problem are the classification of emails as spam or not spam.

A classifier is an algorithm that implements classification. A classifier is a mathematical function that transfers input data to a category.

2.3.2 Supervised Learning

Machine learning algorithms can be supervised or unsupervised. A system that offers desired input and output data is referred to as supervised learning. To aid in the classification of unlabeled data, the input and output data are labeled [13]

The goal of a system receiving input and output variables is to learn how they are mapped. The goal is to design a mapping function that allows the model to predict unknown inputs in the future. This is an iterative procedure in which each prediction is adjusted or given feedback until the algorithm performs well [14]

Hecht-Nielsen proposed backpropagation as a mechanism for updating internal parameters in multilayer networks (1992).

Essentially, the method computes the output layer error and compares it to the desired amount. To lower the error gradient, the weights between the output layer and the preceding layer are modified. So on until all weights up to the input layer have been altered. The training data for this algorithm includes examples with similar input subjects and intended outcomes. An AI system might receive labelled photographs of vehicles or trucks in an image-based supervised learning application. The system must be able to recognise and classify unlabeled photos into one of two groups after monitoring and training.

Two common supervised learning applications are classification and regression. A category, such as a car or a truck, is the output value. A regression problem arises when the output is a computed value such as price, weight, temperature, or humidity.

2.4 Deep Learning

Deep learning tries to "teach" robots to act and understand data in a more natural way [15]. Deep learning has several applications, including autonomous vehicles that can recognize and identify the present condition of a traffic light to decide whether to stop or go.

Deep learning is already being used in many fields of study and industry, allowing for previously unattainable results. Until then, manual feature extraction was done by the developer, who chose the best approach. With deep learning, this selection of the greatest resources is automatic, indicating the results.

Using deep learning, a model may learn to correctly categorize images, text, or sound. These models can often exceed humans in terms of accuracy. Models are trained to utilize vast sets of labeled data and complex artificial neural network topologies. According to [14], two premises are required to get accurate and satisfactory results when utilizing deep learning. The first is that deep learning takes a lot of labeled data, i.e. thousands of photos of each animal type to build a model that correctly identifies them. Second, deep learning necessitates tremendous processing capacity, requiring GPUs and other parallel architectures.

Deep neural networks are models that use artificial neural network topologies. In contrast to traditional neural networks, deep neural networks feature hundreds of hidden layers. To learn features directly from labelled data, this depth is required.

2.4 Deep Learning

"Teach" machines to act and understand data in a more natural way using deep learning [15]. Applications of deep learning include autonomous vehicles that can recognize and identify the present condition of a traffic light, making decisions such as stopping or proceeding. Deep learning is now being used in many fields of science and industry because it produces outcomes that were previously unattainable using other techniques. Until then, manual feature extraction was left to the developer's discretion. These days, deep learning does this for you, and the results are immediate.

It can classify photos, text, and voice using deep learning. These models can often exceed humans in terms of accuracy and precision. Data sets with several hidden layers are used to train models. By understanding two key ideas, deep learning may produce accurate and desirable outcomes. On the one hand, deep learning demands a significant amount of labelled data, which means a model must be trained with thousands of photos of each animal type. For deep learning, substantial processing capacity is required, including GPUs and parallel architecture. These models are commonly referred to as deep neural networks. In contrast to conventional neural networks, deep neural networks feature hundreds of hidden layers. The model may learn features directly from the labelled data, eliminating the requirement for human feature extraction.

2.4.1 Transfer Learning

There are currently three approaches to train deep learning models to classify objects: from scratch, resource extraction, and transfer learning. Depending on the amount of data and the learning rate, this form of training can take days or even weeks to complete. The feature extraction method uses a network to learn features from images and then use them in a machine learning model like a Support Vector Machine [11].

According to [16], most modern deep learning models use transfer learning, which entails fine-tuning parameters of a trained model using other data and resources. It starts with an existing network like [12]. Then, using the labeled data, a classifier is trained on the new issue samples. It is feasible to alter the synaptic weights of only specific layers of the architecture while freezing the others during the training stage. The main benefit of this strategy is that it requires less data, reducing the calculation time during the training stage.

The transfer learning procedure involves adequate interface configuration with the pre-trained network parameters to modify and improve these values for the new task. There are now several libraries available to support developers who adopt this method in their models.

2.5 Artificial Neural Networks

ANNs were initially inspired by biological brain circuits, although they now have little in common. Starting with a simple natural neuron-like computational structure, which accepted input data, multiplied it by real values termed synaptic weights, combined the result creating an activation value, which was sent as a parameter to another function to generate output data. As research progressed, artificial neurons were coupled in various network topologies to form ANNs [8].

An ANN is a complex dynamic system, because it is a network of interconnected systems, represented by a weighted and directed graph, with vertices representing the connections between neurons, weights representing synaptic weights, and nodes representing the components of the neuron that make up the combinations of values and generate the outputs [17].

A rudimentary Perceptron model has only one neuron and can only do linear separations with two object groups. It was possible to approximate complex functions and make data predictions by combining many Perceptrons in a topology with layers of neurons [12]. Hidden layer outputs are multiplied by one set of weights, transmitted to the next layer, and then to m neurons in the output layer. The network outputs probability values after applying an activation function.

2.6 Convolutional Neural Networks (CNN)

One of the key reasons why the world has woken up to deep learning is its efficiency in picture recognition. Currently, various research institutes are pushing the boundaries of computer vision, which has applications in autonomous vehicles, robotics, drones, security, medical diagnostics, and blindness therapy.

MobileNet ([7]and DenseNet are examples of pre-trained and consolidated CNN architectures used for image recognition [2]. We chose empirical tests to identify the optimal parameters to describe the network architecture in this work because the challenge is quite particular. Pre-trained architectures can classify objects like humans, cars, planes, animals, and more. But none of these things resemble parasite eggs. The study's goal was to design a CNN architecture that could accurately classify the various egg species.

2.7 Data Augmentation

The basic goal of data augmentation is to make the model more generalizable. With simple geometric transformations such as translations, rotations, changes in scale, shear, horizontal inversions, and others, it is possible to create new data from the original images. Data magnification is a natural and easy-to-apply strategy for image-intensive machine learning tasks.

2.8 Results Evaluation Metrics: When building a machine learning model, it is important to assess its quality in terms of job efficiency. Metrics are mathematical functions that evaluate a model's error and success capabilities. Choosing a decent model to address an issue is critical, but so is selecting a metric to assess the model's performance. This level of review offers dozens of measures, some basic and others sophisticated. Some metrics function better with certain data sets while others work better with others. The proportion of data from each class in the dataset and the classification or prediction purpose (probability, binary, ranking, etc.) must be considered while picking a metric. That's why knowing the metric to utilise can make all the difference when evaluating the model. In all circumstances, none of the measurements is superior. Always evaluate the model's practical application.

2.8.1 Confusion Matrix

A confusion matrix is a table used in machine learning to visualise a model's performance. This table shows anticipated class instances in each row and actual class instances in each column (Stehman, 1997).

A confusion matrix has four values: true positive, true negative, and false positive. This matrix is particularly valuable for evaluating the model because it contains the results of each record's classification and allows you to find other metrics like accuracy, precision, recall, and F1. This table contains four values:

True Positive (TP): When a classifier accurately predicted that a record was positive, the number of records that were correctly predicted is shown in 1-.

True Negative (TN): the input was indeed negative, as indicated by the classifier's answer and the number of records correctly classified as negative.

False positive (FP): There were three instances in which the classifier mistakenly responded to a positive input, despite the input being negative.

False Negative (FN): For example, there were four instances in which the classifier responded wrongly to a positive item, saying that it should have been labelled as negative.

2.8.2 Quality Metrics

With the values obtained through a confusion matrix, it becomes possible to find the other metrics, such as accuracy, precision, recall and F1-score.

3.0 Material and methods

The Material and Methods chapter is a detailed planning of what was done in this research, so that the work can be performed by other researchers, for replicability. All stages of the process were described, from the acquisition of the set of images to the validation of the proposed models through experiments and comparison of results. The methodology applied in this research can be divided into the steps described in the flowchart illustrated in Figure 2.

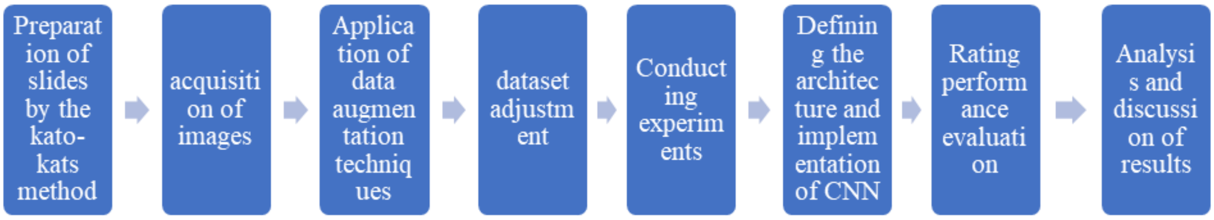


Figure 2 – Flowchart of the proposed system.

3.1 Dataset Creation

There was no public image library of parasite eggs of the species analysed in the literature prior to this investigation. The Institute of Biological Sciences of the Federal University of Minas Gerais created Kato-Katz slides containing faeces samples for examination under the microscope. 66 photos were acquired in RGB with a resolution of 2048 1536 comprising eggs of the following parasite species: *Fasciolopsis buski, Echinococcus granulosus, Diphyllobothrium latum, Strongyloides stercoralis, and Trichinella spiralis*. Images obtained were justified by the samples available at the time of collection, while zoom and resolution were specified by what is frequently employed in this type of laboratory analysis. The parasites *Fasciolopsis buski* and *Echinococcus granulosus* have identical eggs and are so called hookworms. Hookworm eggs do not survive on Kato-Katz slides and their cellular content retracts within hours of preparation. Lesser remains are a thin elliptical eggshell of retracted cellular material. Even after weeks or months of preparation, other helminth eggs commonly observed on Kato-Katz slides have distinct and maintained morphological structures and characteristics. Deep learning requires hundreds of photos to train, according to the literature. Recognizing that deep learning approaches require 66 examples to perform well, we elected to undertake data augmentation operations on these obtained images. Using such processes can greatly increase the number of photographs. To create a representative set of images, a Python script cut all 66 images into smaller images of 200 × 200 (except *Ancylostoma duodenale*, which was cut at 400 400 and afterwards resized to 200 × 200), resulting in around 1000 images for each class. Then each of the 1000 photos was rotated 90º, 180º, and 270º. Only these angles were defined for rotations to keep the image square. Finally, after applying data augmentation methods to collect extra samples, a total of 8000 photos for each class of the numerous parasite species researched in that job were obtained. Initially, it was thought that this number of images would be sufficient to train a CNN model and yield good results in the test set (> 99 percent). The greatest challenge for professionals sweeping faeces for eggs is not correctly identifying helminth species, but mistaking an egg's shape for dirt or impurity in the faeces. To test a binary classifier between a species and the dirt class, 8000 photos were separated that did not contain any eggs of any species. For the dirt class, no rotating techniques were used; these photographs were chosen after clipping the original image of 2048 1536. Figure 4 shows some of the photographs utilised in the work. The original cropped photos are in the first column, labelled X.1. The original photos are rotated 90º in the column with the legend X.2. The first column's image was rotated 180º and 270º for the third and fourth columns. From the fifth through the eighth columns, each image from the previous columns was horizontally inverted.

In Figure 3, the lines depict different parasite egg species, with the first line being Hookworm eggs (*Fasciolopsis buski or Echinococcus granulosus*).

Figure 3 shows eggs of *Diphyllobothrium latum, Fasciolopsis buski, Strongyloides stercoralis, and Trichinella spiralis.* The last line, denoted by the legend F.X, depicts images devoid of parasite eggs, but containing contaminants that may deceive the specialist while analysing the patient's faeces sample.

3.2 Definition of CNN Architecture and Implementation

This study intends to assess a CNN's ability to appropriately classify parasite eggs. The Google Colab development environment was used, which is a free Jupyter notebook environment that runs fully in the cloud and supports the Python programming language. The Keras library was utilised to implement the CNN architecture. CNN Architecture and Implementation In addition to working with TensorFlow, Microsoft Cognitive Toolkit, Theano and PlaidML, it has implemented various open source machine learning methods published in Python. It is a modular and extensible framework that allows rapid construction of deep learning algorithms [18]. We implemented two CNN architectures. The first produces a binary result, indicating if an egg of a certain species is present in the image. The dataset utilised in this design consisted of two classes: photos containing eggs of a single parasite species and images having only dirt and contaminants. This procedure is explained by the difficulties of identifying eggs among dirt and contaminants in a faeces sample. With the second design, the approach can distinguish correctly among all three species studied: *Fasciolopsis buski, Strongyloides stercoralis, and Trichinella spiralis*. Some of the tests for this architecture used transfer learning pre-trained designs such as MobileNet (Howard et al.) The convolution layers were frozen, just the hidden layer weights changing. The number of epochs, batch size of pictures, number of convolution layers, pooling layers, number of feature extractors and their sizes, usage or not of regularisation approaches, such as Batch normalization, Dropout, were varied exhaustively and empirically. Pre-trained architectures like MobileNet [19] were exclusively employed for multiclass classification. Because the problem to be solved is quite specialised, it was chosen to conduct the trials in this fashion.

To discover the optimum network architecture for the problem, many hyperparameter combinations were tested. Feature extractors in each convolutional layer (8, 16, 32, 64 and 128), neurons in fully connected layers (8, 16, 32, 64, 128 and 256) and filter widths ranged from 3 × 3, 5 × 5 and 9 × 9. Also, the model convergence epochs were 1000.

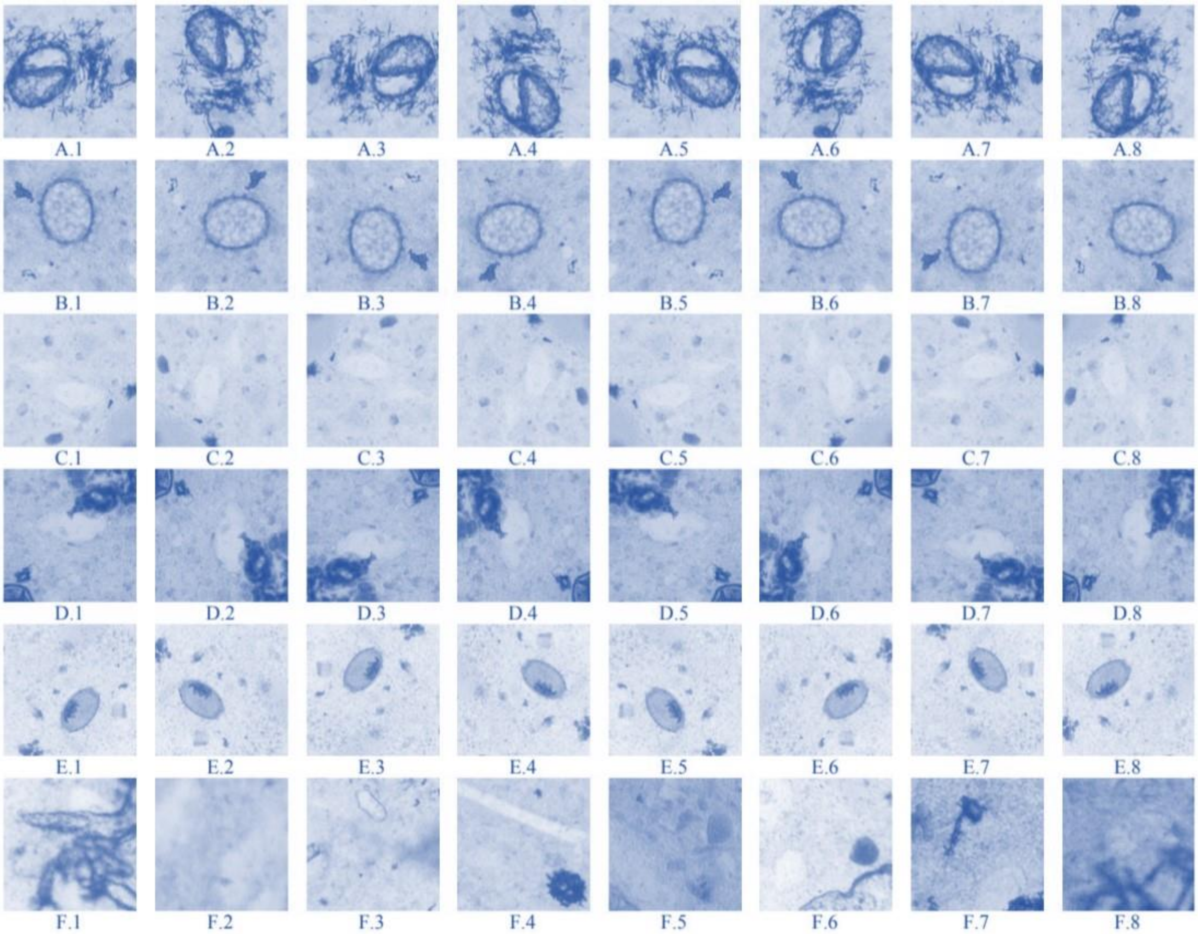


Figure 3: Examples of images contained in the dataset used

3.3 Rating Performance Assessment

It was divided into three sets: training, validation, and testing. The training set had 80% of the photos, whereas the test set contained 20%. The validation set includes 20% of the images drawn for the training set. The sklearn library's train test split method was used for this division, which maintains class balance between training, validation, and test sets. There is no prescribed number of samples in each set, however the values listed are the most common and found in similar studies. For the outcomes analysis, 30 simulations were run for each technique using the assessment metrics already defined: accuracy, precision, recall, and F1-score. Despite the low parasite load, it is preferable for the classifier to err by accusing an egg in an image of absence rather than making a mistake by accusing an egg of absence when in fact there is an egg in the image. How many of the classifications did the model accurately classify? The model's precision is determined by how many right classifications it obtained from all positive classes. The recall is obtained by asking how many of all positive scenarios with expected value are correct. Derived from accuracy and recall, the F1-Score

3.4 Comparison between Models

The doctor determines the need for a helminthiasis diagnostic test based on the patient's clinical signs and history. The patient's history allows the doctor to focus his search for a specific parasite species, which is why the binary classification model was created. The doctor may not have access to the patient's history, which complicates his work because any faeces sample can include eggs of any helminth species.

So, after training the model, it was tested using a test set that solely contained photos of dirt and contaminants, with no eggs of any type. After the multiclass model has classified an image, it may be employed in a class-specific binary classifier that has been trained to distinguish an egg of a certain species from dirt and contaminants in the image.

When the physician is unsure of the patient's parasitic condition, the laboratory technician can send the sample to the multiclass model, which will diagnose it in one of the five parasite classes it was trained for. A binary model for the class designated by the multiclass model is then used to confirm if the patient is indeed infected (as predicted by the multiclass model) or if the sample is merely dirt.

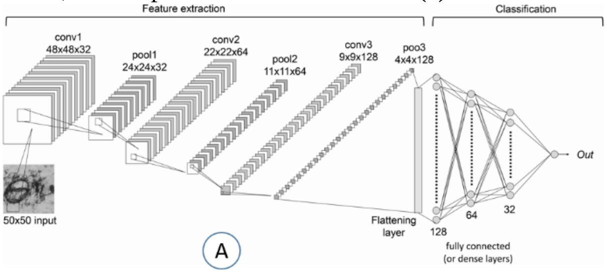
4 Results and Discussion

In the Results and Discussion chapter, the information collected in this research and the analysis of these results are presented. The empirically chosen hyperparameters of the classifiers' architectures and the results obtained using the proposed architectures are reported, in addition to the presentation of the values achieved using transfer learning, with their respective discussions.

4.1 CNN Architecture

It was possible to find a network architecture that achieved sufficient classification performance in the evaluated metrics (> 99 percent) in the test set, both for the binary and multiclass problems, by performing empirical experiments alternating some hyperparameters of the convolutional neural network. The convolutional layers use the ReLU (rectified linear unit) as a function of activation, while the output layer uses the logistic sigmoid. The first convolutional layer includes 32 feature extractors, the second has 64, and the third has 128. They are all 3×3 because the eggs indicate a specific object and small inside the image. Batch Normalization was utilised in the convolution layers to normalise the values in the filters.

The first dense layer consisted of 128 neurons, followed by 64 neurons, and finally 32 neurons. In the dense layers, the Dropout approach was utilised to avoid over-adjusting the network and over-training the model. For the binary problem, we adopted a CNN architecture (a).



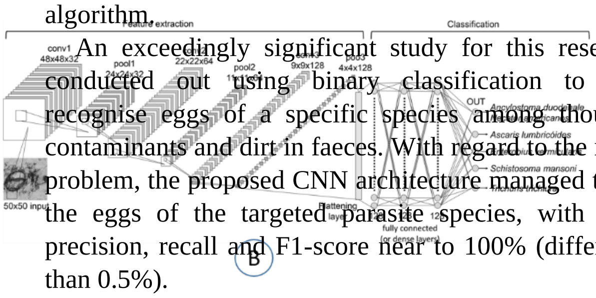


Figure 4 (A) – CNN architecture for the binary classification problem. & (B)- CNN architecture for the multiclass classification problem

The convolution and subsampling layers of the CNN architecture for multiclass classification share hyperparameters with the binary classification architecture. The architectures change in the thick layers, because the multiclass classification issue required 03 (three) dense layers, each with 128 neurons. The Dropout approach is responsible for randomly zeroing 20% of the dense layer neurons. Figure 4B depicts the multiclass CNN architecture. The number of convolution layers, Pooling layers, feature extractors, and their sizes, use or not of regularisation techniques, such as Batch normalization, and Dropout were empirically defined and evaluated by the classification models' evaluation metrics.

4.2 Experiments Using Proposed Architectures

The dataset utilized for the binary classification task consisted of 8,000 photos of one species and 8,000 photographs of dirt. For the multiclass classification task, 40 thousand photos were employed, eight thousand for each class representing a parasite species. Notably, both *Fasciolopsis buski* and *Echinococcus granulosus* are parasites of the Hookworm family, and their eggs have the same shape, placing them in the same class.

[6] ran 30 simulations with random training, validation, and testing sets. Figure 5 shows the mean and standard deviation of each experiment using the architecture provided for the binary problem.

The research listed in the related works section classified eggs of several parasite species with high accuracy (>90%). Most of them used digital image processing to extract egg morphological traits for each species. Figure 5 shows the efficiency (>98%) of utilising convolutional neural networks for the topic in question.

In each experiment, the average recall value for a given parasite egg species was 100% for the Hookworm group, 100% for Diphyllobothrium latum, 99.50% for Fasciolopsis buski, 98.63% for Strongyloides stercoralis, and 100% for Trichinella spiralis eggs.

Because a false negative might cause a patient's disease to progress and even cause death, the recall measure should be given more weight in the evaluation of the suggested algorithm.

An exceedingly significant study for this research was conducted out using binary classification to correctly recognise eggs of a specific species among thousands of contaminants and dirt in faeces. With regard to the multiclass problem, the proposed CNN architecture managed to classify the eggs of the targeted parasite species, with accuracy, precision, recall and F1-score near to 100% (difference less than 0.5%).

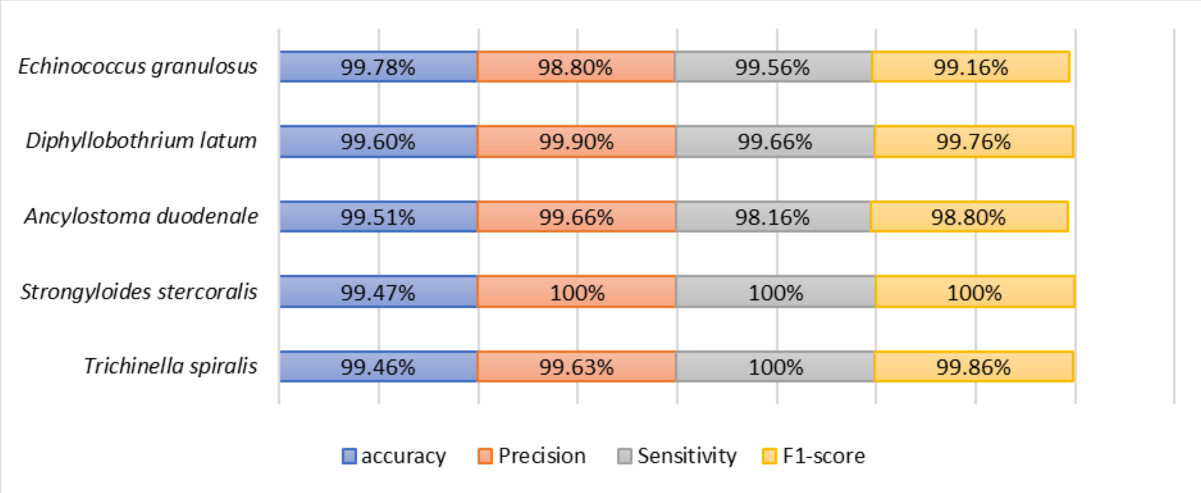


Figure 5 – Results of experiments carried out with eggs of all species studied in this research.

The assessment criteria for all 30 simulations with each model were nearly identical, resulting in a standard deviation close to zero (difference less than 0.05%), validating the proposed architectures. The models were able to accurately classify the parasite eggs in the test photos after training, regardless of the random division of the images in the training, validation, and test sets. An experiment was carried out in which the multiclass model was trained using all photos of the five parasite species addressed in this work, and obtained a value near to 100% (difference less than 0.5%) for all examined metrics. The model was then tested with 4000 photos of dirt and contaminants, but no helminth eggs. The lines indicate the expected output of the model for each image, i.e., each numerical value represents the identification of an image containing just dirt. The numbers range from 0 to 3999 for the 4000 photos submitted. Each image's data are evaluated, and a probability value is generated using the softmax activation function, which represents the network's confidence in classifying this parasite species. A column of colour variation from black to white is visible, as well as values ranging from 0 to 4, one for each parasite species. With this, the model classified each image into a specific species (represented by the white colour, which represents a probability closer to 1 of being in that class), while the other species received a low probability (represented by the colour black). With the heat map, it was evident that most of the time the classification given with greater probability was for class 0, 2, or 3, represented by the eggs of the Hookworm species, *Fasciolopsis buski* and *Ancylostoma duodenale*. The multiclass model may have learned traits from the Hookworm, *Fasciolopsis buski*, and *Strongyloides stercoralis* eggs, as well as from the ground around the eggs. This meant that a particular image of dirt and impurities was more likely to be categorised as one of these three species, as the model may have confused some features present in the dirt image with those retrieved from the egg morphology. After this first classification, the specialist can use the multiclass model to verify if there is a helminth egg of that particular class suggested by the multiclass model, or if it is only dirt. A helminth egg is assumed to be present in the sample because the multiclass model attained a classification accuracy of around 100 percent (difference less than 0.5 percent). So that the specialist can later use the binary model to correctly classify the patient's sample if they don't have the patient's history.

4.3 Experiments Using Transfer Learning

Figure 6 shows the results of multiclass categorization of parasite eggs using pre-trained designs using transfer learning. MobileNet [14]. The convolution layers were frozen, just the hidden layer weights changing. As seen in Figure 6, the results obtained were insufficient (90 percent). MobileNets are built on a simplified architecture that leverages depth separable convolutions to generate deep and lightweight neural networks. The implementation of this design in the solution of the problem presents results in insufficient evaluation (90 percent).

Compared to MobileNet, the pre-trained Xception and DenseNet designs showed an improvement, but not enough to be employable. For this reason, the models could not be employed in a real-world application that relies on medical imagery to diagnose diseases. The results of employing transfer learning with these pre-trained networks demonstrated insufficient evaluation metrics (90%) for the dataset employed, justifying the recommendation of a new network architecture. The lack of success in the evaluation metrics (> 90%) is assumed to be due to the challenge being too particular for feature extractors already trained to identify other objects to generalise to.

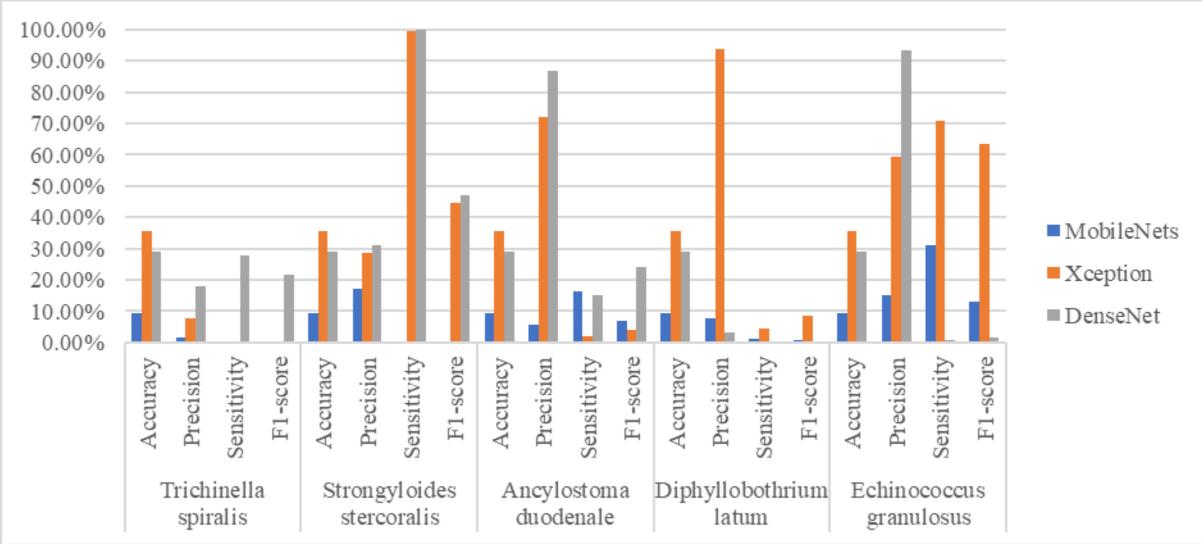


Figure 6 – Results of experiments using transfer learning.

5. Conclusions

Convolutional Neural Networks can be used for other cases that contain biomedical images, such as mammography MRI, microscope or satellite images. Therefore, this work represents a significant contribution to automating the diagnosis of human intestinal diseases and serves as a foundation for the application of CNNs architectures in other problems.

Using the Kato-Katz thick smear method, this work presents an automated methodology for detecting and diagnosing few helminth egg species commonly detected in human faeces: *Fasciolopsis buski*, *Echinococcus granulosus*, *Diphyllobothrium latum*, *Fasciolopsis buski*, *Strongyloides stercoralis*and *Trichinella spiralis* Convolutional Neural Networks were used to classify parasite eggs in optical microscopy images, with the best architecture being determined empirically for the binary and multiclass classification problems.

This work made use of data augmentation operations, which enabled the deployment of deep learning algorithms. Most medical situations, including the one in this study, lack sufficient data to use deep learning algorithms. Using these data augmentation processes becomes critical to achieving good results. CNN architectures: The first successfully distinguished between a species' eggs and contaminants in a faeces slide. All models had 99.9% results in the analysed metrics. Despite certain species' physical similarities, the second managed to classify each egg with 99.9% accuracy. The evaluation metrics for the problems addressed yielded a 99.9% rating. A larger number of helminth species and detection methods can be added to the method.

Convolutional Neural Networks can handle biomedical pictures like mammography, MRI, microscopy, and satellite images. This work automates the diagnosis of human intestinal illnesses and lays the groundwork for additional applications of CNN architectures.

5.1 Future works

The goal of this research was to establish a completely automated computer system for analysing faeces samples that may be used in the Unified Health System (SUS). There are two types of systems. An online method in which the microscope connects to a remote server, which can view the image and detect the eggs. Also, an embedded system entirely connected to an optical microscope might detect eggs automatically. This technique will be used in a future system to detect parasite eggs in faeces for diagnosis. The data received revealed a 99.9% grade in the evaluation metrics for the problems addressed. An actual validation on a population sample, ideally in endemic areas, is required to confirm this performance.

Funding statement: The authors declare that no funding was received.

Data availability statement: The data underlying the results presented in the study are available within the manuscript.

Conflicts of Interest: The authors declare that they have no conflicts of interest regarding the publication of this paper.

Reference

[1]. Peter J. Hotez, Nathan C. Lo, 27 - Neglected Tropical Diseases: Public Health Control Programs and Mass Drug Administration, Editor(s): Edward T. Ryan, David R. Hill, Tom Solomon, Naomi E. Aronson,

Timothy P. Endy, Hunter's Tropical Medicine and Emerging Infectious Diseases (Tenth Edition), Elsevier, 2020, Pages 209-213,

[2]. Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016. Geneva, World Health Organization; 2018.

[3]. Barakat, Rashida & Elmorshedy, Hala & Farghaly, Azza. (2014). Human Schistosomiasis in the Middle East and North Africa Region. 10.1007/978-3-7091-1613-5_2.

[4]. Schaap HB, Den Dulk MO, Polderman AM. Schistosomiasis in the Yemen Arab Republic. Prevalence of Schistosoma mansoni and S. haematobium infection among schoolchildren in the central highlands and their relation to altitude. Trop Geogr Med. 1992 Jan;44(1-2):19-22. PMID: 1496716.

[5]. Barakat, Rashida & Elmorshedy, Hala & Farghaly, Azza. (2014). Human Schistosomiasis in the Middle East and North Africa Region. 10.1007/978-3-7091-1613-5_2.

[6]. McManus, Donald & Dunne, David & Sacko, Moussa & Utzinger, Jürg & Vennervald, Birgitte & Zhou, Xiao-Nong. (2018). Schistosomiasis. Nature Reviews Disease Primers. 4. 10.1038/s41572-018-0013-8.

[7]. Barsoum, Rashad & Esmat, Gamal & El-Baz, Tamer. (2013). Human Schistosomiasis: Clinical Perspective: Review. Journal of Advanced Research. 4. 433–444. 10.1016/j.jare.2013.01.005.

[8]. Guiguet Leal, Diego & Franco, Maura & Neves, Maria & Simões, Luciana & Bastos, Letícia & Allegretti, Silmara & Zanotti-Magalhães, Eliana & Magalhães, Luiz. (2012). Acute Schistosomiasis in Brazilian Traveler: The Importance of Tourism in The Epidemiology of Neglected Parasitic Diseases. Case reports in infectious diseases. 2012. 650929. 10.1155/2012/650929.

[9]. Tallam, Krti & Liu, Zac Yung-Chun & Chamberlin, Andrew & Jones, Isabel & Shome, Pretom & Riveau, Gilles & Ndione, Raphael & Bandagny, Lydie & Jouanard, Nicolas & Eck, Paul & Ngo, Ton & Sokolow, Susanne & De Leo, Giulio. (2021). Identification of Snails and Schistosoma of Medical Importance via Convolutional Neural Networks: A Proof-of-Concept Application for Human Schistosomiasis. Frontiers in Public Health. 9. 642895. 10.3389/fpubh.2021.642895.

[10]. Liu, Zac Yung-Chun & Chamberlin, Andy & Shome, Pretom & Jones, Isabel & Riveau, Gilles & Ndione, Raphael & Bandagny, Lydie & Jouanard, Nicolas & Eck, Paul & Ngo, Ton & Sokolow, Susanne & De Leo, Giulio. (2019). Identification of snails and parasites of medical importance via convolutional neural network: an application for human schistosomiasis. 10.1101/713727.

[11]. Oliveira, Rodrigo & Ferro, Milla & Fernando, Bruno & Bastos-Filho, Carmelo. (2018). Avaliando Técnicas de Aprendizado Profundo para Detecção de Esquistossomose Mansoni em Imagens de Exames Parasitológicos. 1-12. 10.21528/CBIC2017-8.

[12]. Awan, Mazhar & Rahim, Mohd & Salim, Naomie & Rehman, Amjad & Nobanee, Haitham & Shabir, Hassan. (2021). Improved Deep Convolutional Neural Network to Classify Osteoarthritis from Anterior Cruciate Ligament Tear Using Magnetic Resonance Imaging. Journal of Personalized Medicine. 11. 1163. 10.3390/jpm11111163.

[13]. Phung, Son & Bouzerdoum, Abdesselam. (2007). A Pyramidal Neural Network For Visual Pattern Recognition. Neural Networks, IEEE Transactions on. 18. 329 - 343. 10.1109/TNN.2006.884677.

[14]. Soares, Alessandra & Fernandes, Bruno & Bastos-Filho, Carmelo. (2017). Structured Pyramidal Neural Networks. International Journal of Neural Systems. 28. 10.1142/S0129065717500216.

[15]. Delas Peñas, Kristofer & Villacorte, Elena & Rivera, Pilarita & Naval, Prospero. (2020). Automated Detection of Helminth Eggs in Stool Samples Using Convolutional Neural Networks. 750-755. 10.1109/TENCON50793.2020.9293746.

[16]. Si, Minxing & Xiong, Ying & Du, Shan & Du, Ke. (2019). Evaluation and Calibration of a Low-cost Particle Sensor in Ambient Conditions Using Machine Learning Technologies. 10.5194/amt-2019-393.

[17]. Chaganti, Sai & Nanda, Ipseeta & Pandi, Koteswara & Prudhvith, Tavva & Kumar, Niraj. (2020). Image Classification using SVM and CNN. 1-5. 10.1109/ICCSEA49143.2020.9132851.

[18]. Ivašić-Kos, Marina & Pavlic, M. & Matetic, Maja. (2010). Data preparation for semantic image interpretation. 181 - 186.

[19]. Khalaf, O. I., Ajesh, F., Hamad, A. A., Nguyen, G. N., & Le, D. N. (2020). Efficient dual-cooperative bait detection scheme for collaborative attackers on mobile ad-hoc networks. IEEE Access, 8, 227962-227969.