International Journal of System Dynamics Applications
Volume 11 • Issue 1

Primary Mobile Image Analysis of Human Intestinal Worm Detection

Justice Kwame Appati, University of Ghana, Ghana*
https://orcid.org/0000-0003-2798-4524

Winfred Yaokumah, University of Ghana, Ghana
https://orcid.org/0000-0001-7756-1832

Ebenezer Owusu, University of Ghana, Ghana
https://orcid.org/0000-0002-4670-1342

Paul Nii Tackie Ammah, University of Ghana, Ghana

ABSTRACT

One among a lot of public health concerns in rural and tropical areas is the human intestinal parasite. Traditionally, diagnosis of these parasites is by visual analysis of stool specimens, which is usually tedious and time-consuming. In this study, the authors combine techniques in the Laplacian pyramid, Gabor filter, and wavelet to build a feature vector for the discrimination of intestinal worm in a low-resolution image captured with mobile devices. The dimension of the feature vector is reduced using principal component analysis, and the resultant vector is considered as input to the SVM classifier. The proposed framework was applied to the Makerere intestinal dataset. At its preliminary stage, the results demonstrate satisfactory classification with an accuracy rate of 65.22% with possible extension in future work.

KEYWORDS

Classification, Feature Extraction, Gabor Filter, Intestinal Worm, Laplacian Pyramid, Wavelet, Worm Detection

1. INTRODUCTION

The human intestinal worm is one of the most common infectious diseases in human, which is predominant in developing countries possing economic threat to public health (Haque, 2007; Harhay, Horton, & Olliaro, 2010). The early detection of these worms among children is vital for the diagnostic purpose; nonetheless, most health facilities are less equipped with a sophisticated tool and expect skills to manage their outbreak (Garcia, et al., 2017). Conventionally, medical excepts in this field use visual analysis based on experience to carry out their diagnostic task (Crowley, Naus, Stewart, & Friedman, 2003). Though this approach help with the diagnostic process, it is time-consuming and extremely tedious to the extent of possing other health-related hazards (Momcilovic, Cantacessi,

DOI: 10.4018/IJSDA.302631 *Corresponding Author

Arsic-Arsenijevic, Otranto, & Tasic-Otasevic, 2019; Moody & Chiodini, 2001; Tavares, et al., 2011). With the advent of computational tools, several studies are being carried to search for a sustainable and cost-effective solution to the diagnostic problem (Weatherall, Greenwood, Chee, & Wasi, 2006). In the field of parasitology, concepts of pattern recognitions are employed to diagnose medically relevant parasites (Daugschies, Imarom, & Bollwahn, 1999; Sommer, 1996). The concept takes three forms, thus, image pre-processing, feature extractions (Gupta & Shanker, 2021; Aggarwal, Mittal, & Bali, 2021), and classification (Lim, 1990; Jahne, 2005). Techniques adapting these fundamentals concepts in the diagnosis of intestinal worm depending on the characteristic of the dataset are artificial neural networks (Yang, Park, Kim, Choi, & Chai, 2001; Goundar, Prakash, Sadal, & Bhardwaj, 2020), adaptive network-based fuzzy inference system (Dogantekin, Yilmaza, Dogantekin, Avcic, & Sengurc, 2008), MultiClass Support Vector Machine classifier (Avci & Varol, 2009; Panda, 2019; Goundar, Sam; Bhardwaj, Akashdeep, 2021), active contours (Gupta, Bharadwaj, & Rastogi, 2021) and Bayesian classification system (Castañon, Fraga, Fernandez, Gruber, & Costa, 2007). Though computational tools are proven to be usual, the story is slightly different with the African context given the type of dataset captured from the community. It is well observed that, these types of parasites are common in neglected communities where lifestyle and access to health care is poor. Despite these conditions, quite a good number of natives have access to smartphones but with poor resolution which can be leveraged on to enhance health care. The problem then arises with how to use this low resolution images and still provide to some level of precision good health care. This study therefore seeks to propose a framework for the automatic diagnosis of human intestinal worms captured with a smartphone which poses unique challenges.

2. RELATED WORK

In the past few decades, the automation of an intestinal disease diagnostic tool is being researched extensively with a frequent target of perfection. Generally, diagnostic tools are focused on the detection of cyst, eggs, and trophozoite of the intestinal parasite (Nkamgang, Tchiotsop, Tchinda, & Fotsin, 2018). In the study of Avci & Varol (2009), sixteen eggs of protozoa and helminths were classified using multi-class support vector machine. Their proposed framework consists of noise reduction, thresholding, contrast enhancement, morphological and logical processes. Image features were extracted using the invariant moment. The focus of their work was on the detection of human parasite eggs, which yielded a classification rate of 97.70%. The study of Suzuki, Gomes, Falcao, Papa, & Hoshino Shimizu (2013) used genetic programming to fuse optimum-path forest classifier with a multiple object descriptor. This led to the proposal of an automatic classifier to discriminate fifteen protozoan species. An ellipse matching with image foresting transform is used for the image segmentation step while the proposed method was weighted with number of parasites per species. The model demonstrated an efficiency of 98.19%, specificity of 98.32%, and sensitivity of 90.38%. Yang, Park, Kim, Choi, & Chai (2001) also extracted four morphometric features as descriptors to classify human helminths based on the morphological characteristics of the eggs. An algorithm based on ANN and image processing techniques was then constructed for the classification purpose. With these features, two ANN were merged to build a classification system. The first network does background removal while the second network classifiers eggs by their species with an average classification rate of 86.1% and 90.3% respectively. In Ghazali Kamarul, Hadi Raafat, & Mohamed (2013), Trichiris trichura ova and Ascaris lumbricoides ova were detected and classified as parasites from human faecal. In their study, they employed contrast enhancement techniques, noise reduction and segmentation methods, and other morphological processes for the preprocessing phase. Features such as shape, shell, smoothness, and size are extracted as input for the classifier to resulting in 94% and 93% success rate for Trichuris trichiura and Ascaris lumbricoides respectively. The method of probabilistic neural networks was used by Saha, Tchiotsop, Tchinda, Wolf, & Noubom (2015) to recognize the cyst of nine human intestinal

helminth parasites. Their study recorded a 100% correct classification rate after training their network with an image pixel feature. Dogantekin, Yilmaza, Dogantekin, Avcic, & Sengurc (2008), on the other hand also proposed an Adaptive Neuro-Fuzzy Inference System which is moment invariant to recognize the eggs of protozoa and helminths. The moment invariant features extracted is fed into the network resulting in an average recognition rate of 95%. Finally, in the review of Alva, et al. (2017), the morphologic characteristics of parasite eggs stained in a microscopic image was extracted. A logistic regression model was used to build four distinct algorithms for the recognition of the four eggs in the faecal smears. The results showed 99.10% and 98.29% for sensitivity and specificity respectively on Taenia sp. Diphylobotrium latum was recognised with 100% and 98.38% while they got 99.15% and 98.18% for Fasciola hepatica. The last worm thus Trichuris trichuria recorded 100% and 98.13% in their study. Dataset usually used in this study are captured under well-defined condition and precision. However, in the rural setup with less economic benefit such conditions are not defined. Nonetheless, with the high penetration of smart phones, they serve as a good tool to leverage on them as a capturing device. This in effect introduce challenges such as resolution, orientation, precision and among others making recognition harder and an opportunity for this study to explore at its primary stage. The paper is organized as follows: in Section 3, we present the materials and the methods. Section 4 presents some results as well as discussions of the proposed system, and Section 5 is the conclusion.

3. MATERIALS AND METHODS

3.1 Dataset Description

The dataset used for this study is a secondary data by AI Research (Online, Accessed 02 june 2020) unit of Makerere University, Uganda, which is publicly available for use. The dataset contains 1217 stool images with bounding boxes of 162 parasite eggs (hookworm, taenia and hymenolepsis nana).

3.2 The Gaussian and Laplacian Pyramid

Pyramids are multi-scale signal representation where a signal is subject to repeated smoothing and subsampling. Two types of pyramids exist that is Gaussian and Laplacian pyramid, also known as lowpass and bandpass pyramid respectively (Singh, Sinha, & Gupta, 2000). However, this study places more emphasis on the Laplacian pyramid. The Laplacian pyramid, as applied to a digital image, is constructed by taking the difference between adjacent layers in the Gaussian pyramid. In practice, these layers are of varying size, and the smaller sized layers are expanded using interpolating techniques (Singh, Sinha, & Gupta, 2000). Mathematically, this concept is set up as follows. Given a 2D image, let $u(x, y, t)$ represent a Gaussian pyramid then the difference of levels can be stated as $u(x, y, t_2) - u(x, y, t_1)$. With the assumption that Gaussian pyramid function is a solution to the heat

Figure 1. Sample images from the dataset: (left) Hookworm; (center) Taenia; (right) Normal Image

diffusion equation (Singh, Sinha, & Gupta, 2000), the Laplacian pyramid function *V(x, y, t)* is formed with the relation in Eqn. 1:

$$V\left(x,y,t\right)=\lim_{t_2\to t_2}\frac{u\left(x,y,t_2\right)-u\left(x,y,t_1\right)}{t_2-t_1}=\frac{\partial u\left(x,y,t\right)}{\partial t} \tag{1}$$

With the property of convolution, Eqn. 1 can be rewritten as Eqn. 2:

$$V\left(x,y,t\right)=\Delta G * f=G * \Delta f \tag{2}$$

where * is the convolution operator and *f* is the image function.

The original image *f(x,y)* can be reconstructed from *V(x, y, t)* using Eqn. 3:

$$f\left(x,y\right)=-\int_0^T V\left(x,y,t\right)dt+u\left(x,y,T\right)\ 0\le t\le T \tag{3}$$

Algorithmically, the Laplacian pyramid as first proposed by Burt and Adelson (Burt & Adelson, 1993) is outlined in the steps below:

1. Convolve the original and initial image f_0 with a Gaussian filter *G* and subsample it by two to create a reduced Gaussian version of the image f_1.
2. The resultant image is upsampled using well defined interpolating techniques by convolving it with *G* to obtain the enlarged version of the Gaussian image f'_1 and then subtracted pixel by pixel from the original. This gives the detailed image V_0 as defined by Eqn. 4:

$$V_0=f_0-f_1^{'} \tag{4}$$

3. Finally, if an image of size $2^N \times 2^N$ is given, can be further obtained by recursively performing step 1 and 2 on the Gaussian and subsampled image f_1 a maximum of *N* number of times.

To generate the original image f_0 given *N* detailed images say $V_0, V_1, \ldots, V_N$ and a Gaussian image $f_{N,}$ the inverse transformation is required (Pradham, Younan, King, & Stathaki, 2008) as outlined below:

1. f_N is upsampled where the missing values are first interpolated and subsequently convolved with the filter *G* to obtain the image f'_N.
2. The approximation image at the next resolution level in Eqn. 5 is obtained from the sum of image f'_N and the lowest detailed image V_N:

$$f_{N-1}=V_N-f_N^{'} \tag{5}$$

3. The original image is finally obtained by repeating Steps 1 and 2 on the detail images $V_0, V_1, \ldots, V_{N-1}$.

3.3 Feature Extraction

The study made use of the manual feature engineering approach by leveraging on the two-dimensional discrete wavelet transform and Gabor filters. Conventionally, signal transformation into frequency domain is the dominant practice in signal processing as it gives a more effective representation of signals compared to the time domain processing. The subsections below provide a detailed description.

3.3.1 Multiresolution Analysis and Discrete Wavelet Transform

In the concept of multiresolution analysis, a subspace containing parts of a whole function is constructed through the decomposition of an original function *x(n)* at different resolution. These decompositions at different scales are defined as multiresolution (Suter, 1998). The wavelet transform is a projection of a given signal into a set of basis function known as wavelets. This is a generalization of the short-time-Fouriertransform (STFT) where the basis functions are localized in the frequency domain. The discrete wavelet transforms of a signal or function *x[n]* is defined on the approximation coefficient $W_{\phi}\left[j_0,k\right]$ and detailed coefficient $W_{\psi}\left[j,k\right]$ as explained as follows:

$$W_{\phi}\left[j_0,k\right]=\frac{1}{\sqrt{M}}\sum_{n}x\left[n\right]\phi_{j_0,k}\left[n\right] \tag{6}$$

$$W_{\psi}\left[j,k\right]=\frac{1}{\sqrt{M}}\sum_{n}x\left[n\right]\psi_{j,k}\left[n\right]\quad for\ \ j\geq j_0 \tag{7}$$

and the inverse discrete wavelet transform is defined as:

$$x\left[n\right]=\frac{1}{\sqrt{M}}\sum_{k}W_{\phi}\left[j_0,k\right]\phi_{j_0,k}\left[n\right]+\frac{1}{\sqrt{M}}\sum_{j=j_0}^{J}\sum_{k}W_{\psi}\left[j,k\right]\psi_{j,k}\left[n\right] \tag{8}$$

where *M* denotes the number of transformation samples:

$$n=0,1,2,\cdots,M-1,\ j=0,1,2,\cdots,J-1,\ \text{and}\ k=0,1,2,\cdots,2^{j}-1$$

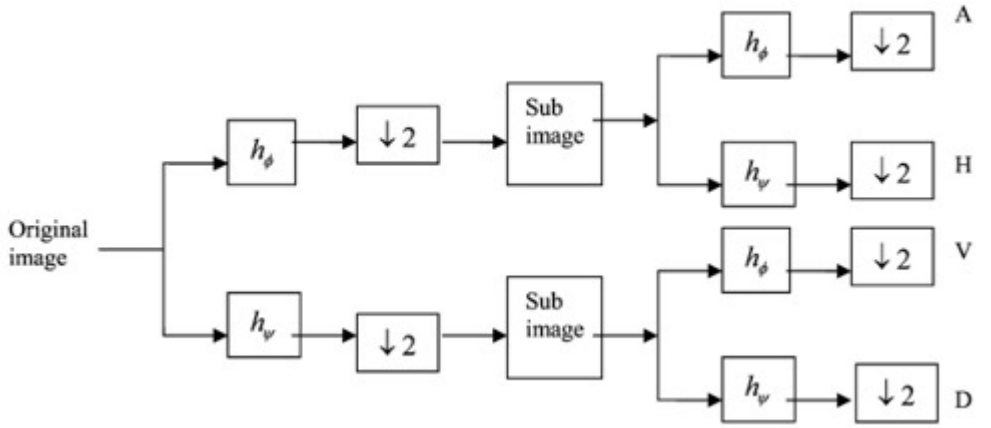
M is selected such that $M=M^{j}$ where *J* is the number of levels for the transformation. The basis functions $\left\{\phi_{j,k}\left[n\right]\right\}$ which is the scaling function and $\left\{\psi_{j,k}\left[n\right]\right\}$ thus the wavelet function is defined as:

$$\phi_{j,k}\left[n\right]=2^{j/2}\phi\left[2^{j}n-k\right] \tag{9}$$

$$\psi_{j,k}\left[n\right]=2^{j/2}\psi\left[2^{j}n-k\right] \tag{10}$$

For an efficient implementation of a two-dimensional wavelet transform, the filter bank structure is often used. Figure 2 illustrates how the discrete wavelet transformed images are obtained. The first step is to convolve the image with the scaling and wavelet function. The resultant output is

Figure 2. A 2D discrete wavelet transform (Pradham, Younan, King, & Stathaki, 2008)

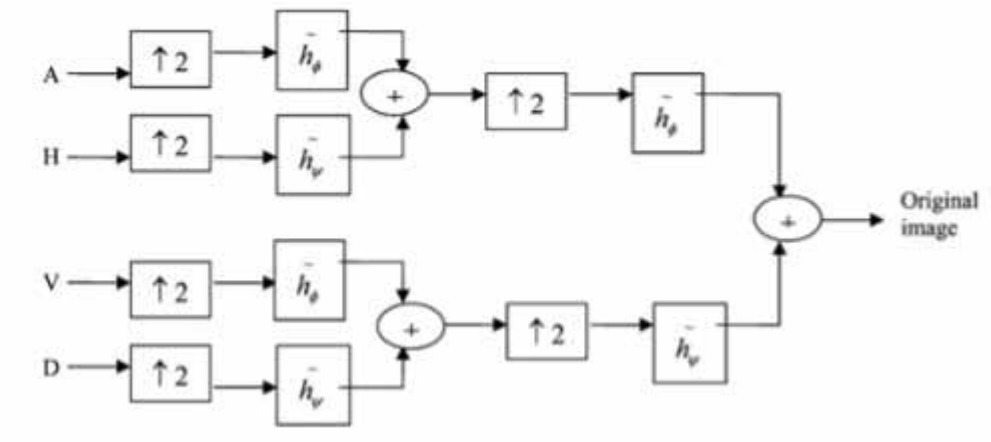


downsampled by a factor two to obtain new image maps. Both maps are further convolved with the same scaling and wavelet function and downsampled again. This results in four maps which are explained as follows (Castleman, 1996):

1. Approximation coefficients (A) – this is the grey level variation of the image at each level.
2. Horizontal coefficients (H) – this is the column variations of the image at a given level.
3. Vertical coefficients (V) – this is the row variations of the image at a given level.
4. Diagonal coefficients (D) – this gives the diagonal variations of the image at a given level.

On the other hand, the reconstruction of the original image is by the inverse discrete wavelet transform given the four sub-images obtained from the discrete wavelet transform. Figure 3 gives a pictorial representation of reconstructing the original image from its discrete wavelet transform components. The four sub-images are upsampled where the vertical sub-image is convolved with the scaling function and the diagonal sub-image is convolved with the wavelet function. The approximation

Figure 3. A 2D inverse discrete wavelet transform (Pradham, Younan, King, & Stathaki, 2008)



and horizontal subimages are also convolved with the wavelet and scaling function respectively. Each pair of the subimage is further upsampled and convolved with the same wavelet and scaling function respectively. The resultant images are finally summed to estimate the original image.

3.3.2 The Gabor Filters

The 2D Gabor filter is a frequency and orientation representation of an edge detector. The filters are functions of Gaussian kernels modulated by a sinusoidal wave oriented at an angle. By practice and theory, Gabor filters has demonstrated significant performance in texture features extraction over other domain-specific filters (Xie, Jiang, & Tsui, 2005). In (Bankman, Spisz, & Pavlopoulos, 2009), the general Gabor filter family is implemented as Eqn. 11:

$$G\left(x,y\right)=exp\left(-\frac{x'^2}{2\sigma_x^2}\right)exp\left(-\frac{y'^2}{2\sigma_y^2}\right)cos\left(\frac{2\pi x'^2}{\lambda}+\varphi\right) \tag{11}$$

with:

$$x'=\left(x-m_x\right)\cos\gamma-\left(y-m_y\right)\sin\gamma$$
$$y'=\left(x-m_x\right)\sin\gamma-\left(y-m_y\right)\cos\gamma$$

where m_x and m_y is the center of the Gabor receptive field, σ_x and σ_y is its size, γ is orientation, φ is a phase offset, and 1¤λ is the preferred spatial frequency of the Gabor filter. In (Xie, Jiang, & Tsui, 2005; Kruizinga & Petkov, 1999), a set of eight orientations ($\gamma = \{22.5°, 45°, \ldots, 180°\}$) and three spatial frequencies ($\gamma = \{5.47,8.20,10.93\}$) is usually used. This results in 24 features for each position of the filter. In like manner, the parameters σ_x and σ_y are set to $\sigma_y = 2\,\sigma_x = 1.12\lambda$, as suggested in (Devalois, Albrecht, & Thorell, 1982). Traditionally, they sum of squares metric is used to quantify the texture of a region in a filtered image.

3.4 Image Classification Model

Good features are technically difficult to obtain and are not sufficient to appropriately classify images but instead required to work with a good classification model that suitably characterize the images. Most classification methods perform model construction based on feature vectors (Mohanty, Lee-St.John, Manmatha, & Rath, 2013). For this task, several classification methods have been developed in the literature. However, in this study, the support vector machine is considered as its theory provides the most principled approach to the design of neural networks, eliminating the need for domain knowledge [33, 34] and its popularity in machine learning methods. Given two classes labelled by 1 or -1, the binary classifier f: $X \rightarrow Y = \{1, -1\}$ is a function mapping the input space X t*o* Y w*h*ich groups X i*n*to two classes. A small error is expected if the classifier performs well. This misclassification error R(f*) is* measured in Eqn. 12 as the probability of predicting an event wrongly $\{f(x) \neq y\}$ (Wu, et al., 2006):

$$R\left(f\right)=Prob_{z\in Z}\left\{f\left(x\right)\neq y\right\}=\int P\left(y\neq f\left(x\right)|x\right)d\rho X \tag{12}$$

Clearly, the best classifier is estimated by minimizing Eqn. 12 with the Bayes rule defined by $f_c\left(x\right)=sgn\left(f_\rho\right)$ where:

$$sgn\left(f\right)\left(x\right)=\begin{cases}1, & \text{if } f\left(x\right)\geq 0,\\ -1, & \text{if } f\left(x\right)<0\end{cases} \tag{13}$$

The task of estimating this good approximation f_z of the Bayes rule is to draw a random sample $z=\left\{\left(x_i,y_i\right)\right\}_{i=1}^{m}$ from a probability distribution ρ. Here, the regularization schemes associated with reproducing a kernel Hilbert space is considered where:

$$f_{z,\lambda}=arg\min_{f\in\mathcal{H}_K}\left\{\frac{1}{m}\sum_{i=1}^{m}\phi\left(y_i f\left(x_i\right)\right)+\lambda f_K^2\right\} \tag{14}$$

where $\varnothing: R\rightarrow R+$ is the loss function for classification.

Given the general classification scheme in Eqn. 14, suppose $X\subset\mathbb{R}^n$ where the sample set, z has two classes that is; $I=\{i: yi=1\}$ and $II=\{i: yi=-1\}$, then these classes are separable by a hyperplane $wx=b$ with $w\in\mathbb{R}^n$, $|w|=1$, and $b\in\mathbb{R}$ if:

$$\begin{cases}w\cdot x_i>b, if\ \ i\in I\\ w\cdot x_j<b, if\ \ j\in II\end{cases} \tag{15}$$

From Eqn. 15, we define $b_1\left(w\right)=\min_{i\in I}\left\{w\cdot x_i-b\right\}>0$ and $b_2\left(w\right)=\max_{j\in II}\left\{w\cdot x_j-b\right\}<0$. To maintain a balance, the hyperplane is shifted to $wx=c(w)$, where $c\left(w\right)=\frac{1}{2}\left\{b_1\left(w\right)+b_2\left(w\right)\right\}$. For $\Delta\left(w\right)=\frac{1}{2}\left\{b_1\left(w\right)+b_2\left(w\right)\right\}>0$, there holds:

$$\begin{cases}w\cdot x_i-c\left(w\right)\geq\min_{\ell\in I}w\cdot x_\ell-c\left(w\right)=\Delta\left(w\right),\ \ if\ \ i\in I\\ w\cdot x_j-c\left(w\right)\geq\max_{\ell\in II}w\cdot x_\ell-c\left(w\right)=-\Delta\left(w\right),\ \ if\ \ j\in II\end{cases} \tag{16}$$

where the two separating hyperplanes are expected to be $2\Delta(w)$ distance apart where $\Delta(w)$ is the margin from the hyperplane $wx=c(w)$. The primary task of the support vector machine is to find the best hyperplane where $\Delta(w)$ is the maximum by solving the optimization problem in Eqn. 17:

$$\Delta\left(w^*\right)=\max_{|w|=1}\frac{1}{2}\left\{\min_{y_i=1}w\cdot x_i-\max_{y_j=-1}w\cdot x_j\right\} \tag{17}$$

If $\Delta(w*)>0$ the[n] z is separable and $w*$ is a solu[t]ion to the optimization problem. This gives a classifier defined as $f(x)=sgn(w*x-c(w*))$. Unfo[r]tunately, [m]ost data are not separable by a hyperplane leading to the introduction of slack variable $\xi=\left(\xi_i\right)_{i=1}^{m}$ into support vector machine termed as soft margin classifier defined as Eqn. 18:

$$\left(\tilde{w},\tilde{b}\right)=\arg\min_{w\in\mathbb{R}^n,b\in\mathbb{R}}\min_{\xi\in\mathbb{R}^n}\left\{\left|w\right|^2+\frac{2C}{m}\sum_{i=1}^{m}\xi_i\right\} \tag{18}$$

such that $y_i\left(w\cdot x_i-b\right)\geq 1-\xi_i,\geq 0, i=1-,\ \ldots,\ m.$ where $C>0$ is the trade-off. This defines $sgn\left(\tilde{w}\cdot x-\tilde{b}\right)$ as the soft margin classifier.

3.5 Performance Metrics

The classification metric is easily illustrated with elements of the confusion matrix. The confusion matrix of a binary classifier where Y and N represent the two classes can be viewed as one of the four possible ways as enumerated below (Kotu & Deshpande, 2015):

1. **True Positive (TP):** Where the predicted label (Y) is the same as the actual label (Y).
2. **False Positive (FP):** Where the predicted label (Y) is different from the actual label (N).
3. **False Negative (FN):** Where the predicted label (N) is different from the actual label (Y).
4. **True Negative (TN):** Where the predicted label (N) is the same as the actual label (N).

These four cases are used to explain the classification performance. The general expectation is that a perfect classifier will have the number of FP = number of FN = 0.

1. Sensitivity/Recall is expressed as a ratio, usually presented as a percentage in Eqn. 19. The metric selects all positive instances tagged with "positive" labels:

$$\frac{TP}{TP+FN} \tag{19}$$

2. Specificity of a model rejects all negative instances that are tagged as "negative" and it is expressed as Eqn. 20:

$$\frac{TN}{TN+FP} \tag{20}$$

3. The precision metric explains the portion of positive instances that are relevant to the study and this is computed as:

$$\frac{TP}{TP+FP} \tag{21}$$

4. The accuracy is defined by Eqn. 22. An accuracy of 100% indicates that *FP = FN = 0*. This metric selects all positive instances while rejecting all negative instances:

$$\frac{TP+TN}{TP+FP+TN+FN} \tag{22}$$

Finally, the error of the classification model is computed as the the complement of accuracy that is; (1 – accuracy).

3.6 Proposed Framework

Figure 4 is an overview of the proposed framework used to carry out this study. Images of the dataset are first loaded with unwanted borders cropped to maintain a uniform image size. The image is further greyed, and the Laplacian pyramid is applied to it. The resultant image is passed to a three levelled two-dimensional discrete wavelet transform. Gabor filters are then applied to both the approximation and detail image maps generated from each level of the wavelet transform. The Gabor features are put together to form the feature vector. Due to the largeness of the feature space, a principal component analysis is applied to reduce the feature space. A refined soft margin support vector machine model is build using the result feature space. The model is then validated using metric developed from the standard confusion matrix for classification.

4. RESULTS AND DISCUSSION

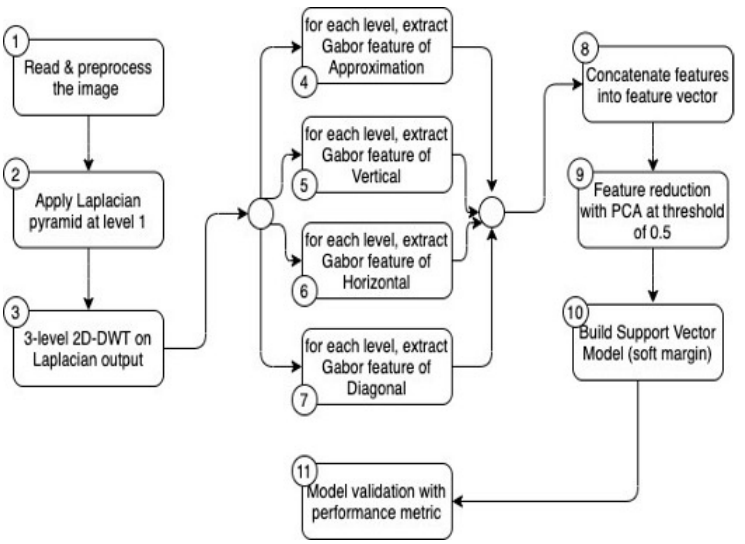
4.1 Environmental Setup

This study was carried out using MATLAB^{R} v2016 on a CPU machine with 16GB memory, Intel^{R} Core^{TM} i7-7800U processor at 2.80GHz x 4 running 64-bit Ubuntu 18.04 LTS.

4.2 Feature Composition

Using a Gabor filter of five scales, eight orientations, and size of 39 x 39 on the wavelet image map produces 40 features. Given that we have three image maps per wavelet level, excluding the approximation coefficient, a total of 40 x 3 x 3 = 360 features are extracted from each image in the dataset. The structure and the distribution of the images are such that the features extracted are highly correlated. Considering the high dimension of these feature vector, a dimensionality reduction technique known as principal component analysis (PCA) was employed. In effect, only 5 out of 360 features were found to be better explain the dataset, thereby reducing the overhead cost in the training phase.

Figure 4. The proposed framework



4.3 Analysis of Observation

This section of the study discusses the various experimental scenarios observed in analyzing the detection of the intestinal worm of the given dataset.

4.3.1 Effect of Knowledge Drift on Performance

The original dataset used for this study has unbalanced characteristics where only 147 observations out of a total of 1217 observations had intestinal worm present. This poses a challenge of knowledge drift in the data distribution.

At a glance of Table 1, as has been the practice of most study, one may conclude that the linear SVM performs well compared to the other classification models based on the face value of accuracy metric. Several works of literature only report the performance of their model based on the overall accuracy of their model without taking into account other factors that are of practical interest to the research and industrial community. However, careful observation of the table in relation to other metrics more especially, sensitivity/recall and precision indicate that there is some hidden information that needs to be revealed as shown in Table 1. In Table 1, linear SVM refers to the linear support vector machine, LDA is linear discriminant analysis, kNN is k-nearest neighbourhood, tree is decision tree, and Bayes is the naïve Bayes.

From Table 2, it is evident that reporting the performance of a classification algorithm using the accuracy metric, as shown in Table 1 is misleading to the research community. In Table 2, A-NW represent actual image without the intestinal worm, A-W is an actual image with the intestinal worm, P-NW is images predicted not to have intestinal worm by the classifier, and lastly, P-W is the images predicted to have an intestinal worm. The table is an augmentation of five confusion matrix. Observing Table 2 carefully, one will notice that the accepted classify thus Linear SVM based on accuracy as reported in Table 1 is biased towards non-intestinal worm images. It fails to discriminate the wormed images from the nonwormed images. This is usually attributed to the knowledge drift problem introduced into the dataset. As the classifier makes an attempt to discriminate appropriately, the performance in the context of accuracy metric turns to decrease making kNN reporting low performance even though its performance is more appreciative compared to the rest using the confusion matrix. This problem could be addressed using the concept of knowledge balance, as discussed in Section 4.3.2.

Table 1. Performance of classification algorithm

Metric	Linear SVM %	LDA %	kNN %	Tree %	Bayes %
Accuracy	87.92	87.76	78.96	80.36	87.76
Sensitivity/Recall	NaN	25.00	10.79	16.67	37.50
Specificity	87.92	87.96	87.76	88.51	88.09
Precision	0	0.68	10.20	15.65	2.04

Table 2. Confusion matrix of the classification algorithm

	Linear SVM		LDA		kNN		Tree		Bayes	
	P-NW	P-W	P-NW	P-W	P-NW	P-W	P-NW	P-W	P-NW	P-W
A-NW	1070	0	1067	3	952	118	939	131	1064	6
A-W	147	0	147	0	135	12	127	20	144	3

4.3.2 Effect of Balanced Knowledge on Performance

It is usually essential that the distribution of classes in a given dataset is reasonably balanced. Enforcing this requirement is expected to resolve the issue of knowledge drift, as discussed in Section 4.3.1. For the sake of computational cost and the demonstration of a proof of concept, the number of non-intestinal wormed images was reduced to match up with the intestinal wormed images instead of employing the concept of data augmentation. In effect, the number of intestinal wormed images remain at 147 while the non-intestinal wormed images become 152. This gives a marginal variation of five images and a total of 299 images.

From Table 3, it is clear that the accuracy value reported in Table 1 was overly estimated and did not give a true reflection of the classifier's performance. Though the overall performance of the classifiers in Table 3 is relatively low using the accuracy metric, it gives a true reflection of the classifier. This is best for industrial practice as it engages the organizational research unit to take further investigation to propose new classifiers or feature set that best explains the dataset. Again, from Table 3, one will notice that the Bayes classifier recorded 93.20% with its precision metric though it had the lowest performance value with accuracy metric. This though gives mixed emotion on the acceptance or rejection of the Bayes model. However, from Eqn. 20 and Eqn. 21 it is stated that precision is more sensitive to "truly positive" while specificity is more sensitive to "truly negative" which is evident in Table 4.

In Table 4, the Bayes model finds it challenging to discriminate between wormed and non-wormed images by treating almost all images as wormed images. This results in poor performance, in general, using the accuracy metric. However, LinearSVM, LDA, and kNN make a fair trade-off though their performance is relatively low due to the underlining challenges established in the dataset. Even though LDA recorded the highest performance, SVM is more parametric and can be fine-tuned. Besides, it recorded the second position in performance in its most basic form.

4.3.3 Parameter Tuning of SVM

From Eqn. 18, not all dataset is separable with the linearity assumption in SVM, hence the need for a data transformation with kernel functions. A number of kernel functions exist; however, Table 5 shows the dynamics of the polynomial kernel as used in SVM. The accuracy metric was considered.

From Table 5, the cubic polynomial kernel function with 9-folds proves a considerable improvement in the use of SVM as a classifier for this study. Again, from the table, an increase in

Table 3. Performance of classification algorithm

Metric	Linear SVM %	LDA %	kNN %	Tree %	Bayes %
Accuracy	58.19	59.53	50.84	51.17	49.83
Sensitivity/Recall	57.97	60.00	50.00	50.35	49.46
Specificity	58.39	59.17	51.57	51.90	54.55
Precision	54.42	53.06	47.62	48.30	93.20

Table 4. Confusion matrix of the classification algorithm

	Linear SVM		LDA		kNN		Tree		Bayes	
	P-NW	P-W	P-NW	P-W	P-NW	P-W	P-NW	P-W	P-NW	P-W
A-NW	94	58	100	52	82	70	82	70	12	140
A-W	67	80	69	78	77	70	76	71	10	137

Table 5. Dynamics of the polynomial kernel on SVM

Kfold	2nd Order	3rd Order	4th Order	5th Order
3	59.20	59.87	47.49	52.51
5	**63.55**	60.20	54.18	48.83
7	61.20	63.88	58.19	51.17
9	62.21	**65.22**	54.18	54.85

the order of the polynomial kernel leads to a decrease in performance. This observation perhaps could be attributed to the fact that the classifier turns to overfit the implicit train data resulting in low performance in the implicit test data. Furthermore, an increase in the order of the polynomial kernel also increases the computational time required. In this instance, one may by choice compromise on performance and select the quadratic polynomial kernel function with 5-folds which gives 63.55%.

4.3.4 Deep Learning Techniques

In this study, concepts of deep learning were not explored due to the limitation of the physical devices. This is therefore treated as a threat to validity of the study, nonetheless it is expected to be adopted when building the hybrid models in future studies.

5. CONCLUSION

In this study, we have been able to analysis intestinal worm dataset curated by Makerere AI Research unit using a mobile phone at its preliminary stage. This was done with the fusion of Laplacian pyramid, Gabor filters and reverse bio-orthogonal wavelet features. Despite the fundamental challenges posed by the dataset, our proposed framework attained an overall accuracy of 65.22% using the third-order polynomial kernel of support vector machine at 9-folds. It was also concluded that research in this field should not only rely on the accuracy value reported by their framework, but to also check the behaviour of their classifiers subject to other parameters and how they contribute to the overall performance of their model. This study based on the characteristic of the given dataset serves as a benchmark for further study to improve upon performance. In future studies, we seek to create hybrid models that leverages on the advantages of the proposed framework with other existing frameworks to increase performance of the benchmark set in this study.

FUNDING AGENCY

The publisher has waived the Open Access Processing fee for this article.

REFERENCES

Aggarwal, D., Mittal, S., & Bali, V. (2021). Significance of Non-Academic Parameters for Predicting Student Performance Using Ensemble Learning Techniques. *International Journal of System Dynamics Applications*, *10*(3), 38–49.

Alva, A., Cangalaya, C., Quiliano, M., Krebs, C., Gilman, R. H., & Sheen, P. (2017). Mathematical algorithm for the automatic recognition of intestinal parasites. *PLoS One*, *12*(4), e0175646.

Avci, D., & Varol, A. (2009). An expert diagnosis system for classification of human parasite eggs based on multi- class SVM. *Expert Systems with Applications*, *36*(1), 43–48.

Bankman, I. N., Spisz, T. S., & Pavlopoulos, S. (2009). Two-Dimensional Shape and Texture Quantification. Handbook of Medical Image Processing and Analysis, 261-277.

Burt, P. J., & Adelson, E. H. (1993). The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, *31*(4), 532–540.

Castañon, C., Fraga, J., Fernandez, S., Gruber, A., & Costa, L. (2007). Biological shape characterization for automatic image recognition and diagnosis of protozoan parasites of the genus Eimeria. *Pattern Recognition*, *40*(7), 1899–1910.

Castleman, K. R. (1996). *Digital Image Processing*. Prentice Hall.

Crowley, R. S., Naus, G. J., Stewart, J., & Friedman, C. P. (2003). Development of Visual Diagnostic Expertise in Pathology - An Information-processing Study. *Journal of the American Medical Informatics Association*, *10*(1), 39–51.

Daugschies, A., Imarom, S., & Bollwahn, W. (1999). Differentiation of porcine Eimeria spp. by morphologic algorithms. *Veterinary Parasitology*, *81*(3), 201–210.

Devalois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, *22*, 545–559.

Dogantekin, E., Yilmaza, M., Dogantekin, A., Avcic, E., & Sengurc, A. (2008). A robust technique based on invariant moments–ANFIS for recognition of human parasite eggs in microscopic images. *Expert Systems with Applications*, *35*(3), 728–738.

Garcia, L. S., Arrowood, M., Kokoskin, E., Paltridge, G. P., Pillai, D. R., Procop, G. W., & Visvesvara, G. et al. (2017). Laboratory Diagnosis of Parasites from the Gastrointestinal Tract. *Clinical Microbiology Reviews*, *31*(1).

Ghazali Kamarul, H., Hadi Raafat, S., & Mohamed, Z. (2013). Automated system for diagnosis intestinal parasites by computerized image analysis. *Modern Applied Science*, *7*(5), 98–114.

Goundar, S., & Bhardwaj, A. (2021). Property Valuation Using Linear Regression and Random Forest Algorithm. *International Journal of System Dynamics Applications*, *10*(4), 1–16.

Goundar, S., Prakash, S., Sadal, P., & Bhardwaj, A. (2020). Health Insurance Claim Prediction Using Artificial Neural Networks. *International Journal of System Dynamics Applications*, *9*(3), 40–57.

Gupta, A., Bharadwaj, N., & Rastogi, V. (2021). Computational Framework of Various Semi-Active Control Strategies for Road Vehicles Thorough Bondgraphs. *International Journal of System Dynamics Applications*, *10*(4), 1–29.

Gupta, A. K., & Shanker, U. (2021). Prediction and Anticipation Features-Based Intellectual Assistant in Location-Based Services. *International Journal of System Dynamics Applications*, *10*(4), 1–25.

Haque, R. (2007). Human Intestinal Parasites. *Journal of Health, Population and Nutrition*, *25*(4), 387–391. PMID:18402180

Harhay, M. O., Horton, J., & Olliaro, P. L. (2010). Epidemiology and control of human gastrointestinal parasites in children. *Expert Review of Anti-Infective Therapy*, *8*(2), 219–234.

Haykin, S., Sinha, N. K., & Gupta, M. M. (2000). Neural Networks: A Guided Tour. *Soft Computing and Intelligent Systems*, 71-80.

Jahne, B. (2005). *Digital Image Processing*. Springer.

Kotu, V., & Deshpande, B. (2015). Model Evaluation. *Predictive Analytics and Data Mining, Concepts and Practice with Rapidminer*, 257-275.

Kruizinga, P., & Petkov, N. (1999). Nonlinear operator for oriented texture. *IEEE Transactions on Image Processing*, *8*, 1395–1407.

Lim, J. S. (1990). *Two-Dimensional Signal and Image Processing*. Prentice Hall.

Mohanty, N., Lee-St. John, A., Manmatha, R., & Rath, T. M. (2013). Shape-Based Image Classification and Retrieval. Handbook of Statistics, 31, 249-267.

Momcilovic, S., Cantacessi, C., Arsic-Arsenijevic, V., Otranto, D., & Tasic-Otasevic, S. (2019). Rapid diagnosis of parasitic diseases: Current scenario and future needs. *Clinical Microbiology and Infection*, *25*(3), 290–309.

Moody, A. H., & Chiodini, P. L. (2001). Methods for the detection of blood parasites. *Clinical and Laboratory Haematology*, *22*, 189–202.

Nkamgang, O. T., Tchiotsop, D., Tchinda, B. S., & Fotsin, H. B. (2018). A neuro-fuzzy system for automated detection and classification of human intestinal parasites. *XXX*, *13*, 81–91.

Panda, M. (2019). Software Defect Prediction Using Hybrid Distribution Base Balance Instance Selection and Radial Basis Function Classifier. *International Journal of System Dynamics Applications*, *8*(3), 53–75.

Pradham, P., Younan, N. H., King, R. L., & Stathaki, T. (2008). Concepts of image fusion in remote sensing applications. *Image Fusion*, 393-428.

Saha, T. B., Tchiotsop, D., Tchinda, R., Wolf, D., & Noubom, M. (2015). Automatic recognition of human parasite cysts on microscopic stools images using principal component analysis and probabilistic neural network. *Int J Adv Res Artif Intell*, *4*(9), 26–33.

Singh, R.-N. P., Sinha, N. K., & Gupta, M. M. (2000). An Intelligent Approach to Positive Target Identification. *Soft Computing and Intelligent Systems*, 549-570.

Sommer, C. (1996). Digital image analysis and identification of eggs from bovine parasitic nematodes. *Journal of Helminthology*, *70*(2), 143–151.

Suter, B. W. (1998). Wavelet Signal Processing. *Wavelet Analysis and Its Applications*, 167-190.

Suzuki, C. T., Gomes, J. F., Falcao, A. X., Papa, J. P., & Hoshino Shimizu, S. (2013). Automatic segmentation and classification of human intestinal parasites from microscopy images. *IEEE Transactions on Biomedical Engineering*, *60*(3), 803–812.

Tavares, R. G., Staggemeier, R., Borges, A. L., Rodrigues, M. T., Castelan, L. A., Vasconcelos, J., & Spalding, S. M. et al. (2011). Molecular techniques for the study and diagnosis of parasite infection. *The Journal of Venomous Animals and Toxins Including Tropical Diseases*, *17*(3), 239–248.

Vapnik, V. N. (1998). *Statistical Learning Theory*. Wiley.

Weatherall, D., Greenwood, B., Chee, H. L., & Wasi, P. (2006). Science and Technology for Disease Control: Past, Present, and Future. In *Disease Control Priorities in Developing Countries* (pp. 1–44). The International Bank for Reconstruction and Development / The World Bank.

Wu, Q., Ying, Y., Zhou, D.-X., Jetter, K., Buhmann, M. D., Haussmann, W., & Stöckler, J. et al. (2006). Learning Theory: From Regression to Classification. *Studies in Computational Mathematics*, *12*, 257–290.

Xie, J., Jiang, Y., & Tsui, H. (2005). Segmentation of kidney from ultrasound images based on texture and shape priors. *IEEE Transactions on Medical Imaging*, *24*, 45–57.

Yang, Y. S., Park, D. K., Kim, H. C., Choi, M. H., & Chai, J. Y. (2001). Automatic identification of human helminth eggs on microscopic fecal specimens using digital image processing and an artificial neural network. *IEEE Transactions on Biomedical Engineering*, *48*(6), 718–730.

Justice Kwame Appati is a lecturer in the School of Physical and Mathematical Science (SPMS) and the Department of Computer Science. He began his teaching career at Kwame Nkrumah University of Science and Technology in Kumasi as a graduate assistant and then later moved to the University of Ghana in 2017 as a lecturer. Justice earned a PhD, in Applied Mathematics from Kwame Nkrumah University Science and Technology in 2016. He also graduated in 2010 and 2013 with a BSc. Mathematics and MPhil Applied Mathematics from the same institution. His current research includes data science, mathematical intelligence, image processing and scientific computing, He has singly and jointly supervised undergraduate and postgraduate students from Kwame Nkrumah University of Science and Technology (KNUST), National Institute of Mathematical Sciences (NIMS), African Institute of Mathematical Sciences (AIMS) and the University of Ghana Currently, Justice handles course like Design and Analysis of Algorithm, Artificial Intelligence, Formal Methods and Computer Vision. He looks forward to working with everyone interested in his field of study more especially, Intelligence and Data Science.

Winfred Yaokumah is a researcher and senior faculty at the Department of Computer Science of the University of Ghana. He has published several articles in highly rated journals including Information and Computer Security, Information Resources Management Journal, IEEE Xplore, International Journal of Human Capital and Information Technology, International Journal of Human Capital and Information Technology Professionals, International Journal of Technology and Human Interaction, International Journal of e-Business Research, International Journal of Enterprise Information Systems, Journal of Information Technology Research, International Journal of Information Systems in the Service Sector, and Education and Information Technologies. His research interest includes cyber security, cyber ethics, network security, and information systems security and governance. He serves as a member of the International Review Board for the International Journal of Technology Diffusion.

Paul Nii Tackie Ammah is a software and systems engineer with vast experience in the development and delivery of enterprise software systems and solutions for private organizations, government agencies and non-profit organizations. Paul is an experienced software engineer with over 10+ years track record for developing modern systems and re-engineering legacy solutions. Paul is a recognised expert in relational databases, data migration, data transformation and data modeling. Paul holds a BSc. Computer Science and Statistics, MEng. Computer Engineering and is currently pursuing a PhD in Computer Science researching in Convolutional Neural Networks and its application in Medical Image processing.