Chapter 6

Identification of Parasite Eggs using Machine Learning Algorithms

After segmenting and extracting features, the detected objects are classified into multiple classes using machine learning-based classifiers. These classes include three types of parasite eggs: Ascaris, Necator, and Trichuris, and one class of non-egg objects. The experimentation involves evaluating several classifiers, such as Artificial Neural Network (ANN), Support Vector Machine (SVM), k-nearest Neighbours (kNN), Decision Tree (DT), Random Forest (RF), and eXtreme Gradient Boosting (XGBoost). These classifiers are tested with various hyper-parameter configurations, employing an empirical approach to determine optimal settings for different feature sets. Table 6.1 shows some important hyper-parameter settings used in various classifiers for our classification tasks.

Certain parameters, such as the value of K in k-nearest neighbours (kNN), maximum depth, minimum sample split, etc., in decision trees and random forest classifiers, are determined using the cross-validation method. This approach involves performing classification with k-fold cross-validation for different values of each parameter and selecting the value that yields optimal performance. Figure 6-1 illustrates an example of determining the value of K (number of neighbours) in the kNN classification algorithm using Hu's seven invariant moments. The figure shows that the highest classification accuracy is achieved when K=7. Similar procedures are applied to optimize parameters for other feature sets.

In the following sections, the findings of our classification experiments

Table 6.1: Different classification algorithms and hyper-parameter settings for classifying parasite egg images

Classifier	**Parameters**
ANN	Number of Hidden Layers: 2 (Hu moments, Texture and pixel intensity feature), 3 (for other moments) Learning rate: Adaptive learning rate based on Adam (Adaptive Moment Estimation) algorithm Activation Function: ReLU (Rectified Linear Unit) Optimizer: SGD (Stochastic Gradient Descent) and Adam (Adaptive Moment Estimation) Batch Size: 64.
SVM	Kernel type: Radial Basis Function (RBF) Degree: 2 Gamma Parameter: Class Weight: total samples / (number of classes $\times$ number of samples of each class) Decision Function: One-Vs-Many.
kNN	No. of Neighbours: Determined using cross-validation method Distance Metric: Euclidean distance (for Hu moments), and Manhattan distance (for others).
Decision Tree	criterion: Entropy Max depth, Min Samples Split, Min Samples Leaf: Determined using cross-validation Max Features: square root of total features.
Random Forest	No. of estimators, max depth, min samples split, min samples leaf, max features: Determined using cross-validation.
XGBoost	learning rate: 0.1 No. of estimators, max depth: determined using cross-validation gamma: Tuned using cross-validation within the range of [0,1] alpha (L1 regularization), (L2 regularization): Empirically chosen between [0, 10].

using different sets of features are provided. At first, all the objects that are not parasite eggs are grouped into a single class, and the results are examined. Later on, the effectiveness is evaluated by dividing non-egg objects into multiple classes and comparing the outcomes with our initial approach. This step-by-step analysis allows us to gain insights into the effectiveness of the classification methods in different scenarios.

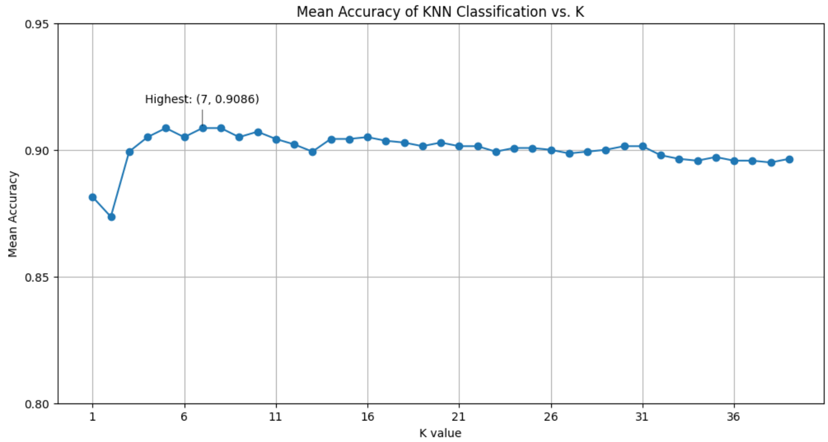


Figure 6-1: Change of classification accuracy with K-value for KNN with Hu's moments

6.1 Classification Using Four Classes of Objects

For the classification task, using a single class of non-egg object and three parasite eggs, the training dataset is prepared using various data augmentation techniques, as discussed in Chapter 3. All classifiers are trained using a randomly selected 80% of the total data, and validation is done using 10-fold cross-validation method. The remaining 20% of samples are used for testing and evaluating the models.

6.1.1 Classification Results Obtained Using Image Moments

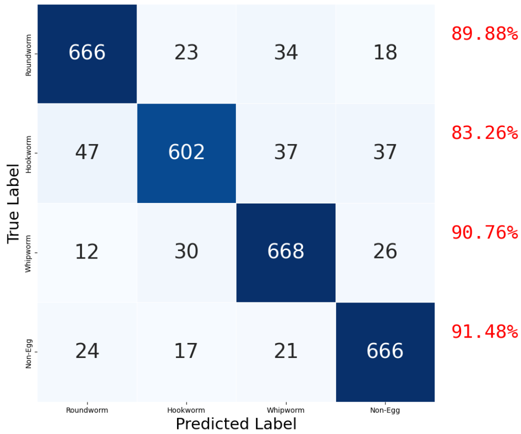
6.1.1.1 Classification Results using Hu's Moments

To obtain an effective ML-based trained model, several experiments are performed by fine-tuning several hyper-parameters, as mentioned in 6.1. The highest overall classification accuracies obtained by various classifiers are recorded in Table 6.2. The confusion matrices and classification reports obtained from the testing of ANN and SVM models are shown in Figures 6-2 and 6-3.

The results show that Hu's moments fail to produce satisfactory outcomes across all classification methods. Particularly, it exhibits challenges in low identification accuracy for Hookworm eggs within our dataset. This limitation may be due to the similarity between the internal textures of Hookworm eggs and the background, resulting in resemblances with other non-egg objects present in the images.

Table 6.2: Classification Results Obtained by Various Classifiers Using Hu's Seven Moments

Classifier	**Overall Accuracy (Training)**	**Overall Accuracy (Test)**
ANN	89.72%	88.87%
SVM	88.92%	87.85%
kNN	84.45%	82.77%
Decision Tree	85.34%	84.21%
Random Forest	89.25%	87.84%
XGBoost	90.86%	89.52%

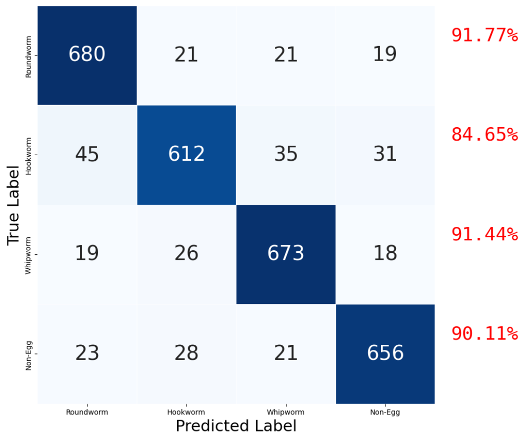


(a) (a) Confusion Matrix (Testing)

	precision	recall	f1-score	support
Roundworm	0.89	0.90	0.89	741
Hookworm	0.90	0.83	0.86	723
Whipworm	0.88	0.91	0.89	736
Non-Egg	0.89	0.91	0.90	728
accuracy			0.89	2928
macro avg	0.89	0.89	0.89	2928
weighted avg	0.89	0.89	0.89	2928

(c) (b) Classification Report (Testing)

Figure 6-2: Test results of ANN using Hu's seven moments



(a) (a) Confusion Matrix (Testing)

	precision	recall	f1-score	support
Roundworm	0.89	0.92	0.90	741
Hookworm	0.89	0.85	0.87	723
Whipworm	0.90	0.91	0.91	736
Non-Egg	0.91	0.90	0.90	728
accuracy			0.90	2928
macro avg	0.90	0.89	0.89	2928
weighted avg	0.90	0.90	0.89	2928

(c) (b) Classification Report (Testing)

Figure 6-3: Test results of XGBoost using Hu's seven moments

6.1.1.2 Classification Results Obtained Using Chebyshev, Legendre, and Krawtchouk Moments

The feature dimension grows with the order of moments as $\frac{(n+1)(n+2)}{2}$, where n is the moment order. Various experiments are conducted, exploring different moment orders, to observe how accuracy changes with moment orders for all three types of moments. The overall classification accuracy obtained by different classifiers using different orders of Chebyshev, Legendre, and Krawtchouk Moments presented in Table 6.3, 6.4, 6.5, 6.6, 6.7, and 6.8.

Table 6.3: Classification Results Obtained by ANN using Various types of Image Moments

Moment Order	**No. of Moments**	**Legendre Moment**	**Chebyshev Moments**	**Krawtchouk Moments**
5	21	89.9	87.5	78.7
10	66	92.7	89.6	77.4
15	136	92.1	93.2	84.5
20	231	91.3	91.7	88.1
25	351	90.0	91.2	88.6
30	496	88.7	90.1	89.5
35	666	88.1	89.6	90.8
40	861	86.3	88.3	92.1
45	1081	85.5	86.8	91.9
50	1326	83.7	85.2	92.0

Table 6.4: Classification Results Obtained by SVM Using Various Types of Image Moments

Moment Order	**No. of Moments**	**Legendre Moment**	**Chebyshev Moments**	**Krawtchouk Moments**
5	21	90.7	86.8	76.0
10	66	93.1	93.5	77.9
15	136	93.0	92.5	83.0
20	231	91.6	91.4	87.5
25	351	90.5	90.5	85.7
30	496	89.3	89.4	89.6
35	666	88.6	88.1	90.8
40	861	85.9	87.6	91.3
45	1081	84.4	86.3	91.6
50	1326	83.2	84.6	91.5

Based on the results presented in the above tables, a graph is drawn as shown in Figure 6-4, highlighting the impact of moment order on the performance

Table 6.5: Classification Results Obtained by kNN using Various types of Image Moments

Moment Order	**No. of Moments**	**Legendre Moment**	**Chebyshev Moments**	**Krawtchouk Moments**
5	21	87.7	89.3	79.1
10	66	90.4	90.3	80.3
15	136	89.1	89.2	82.1
20	231	88.2	87.4	84.8
25	351	87.3	87.5	85.7
30	496	86.8	86.9	87.0
35	666	85.2	84.7	88.9
40	861	83.6	83.2	89.5
45	1081	81.5	82.6	89.8
50	1326	81.3	80.7	89.9

Table 6.6: Classification Results Obtained by Decision Tree Using Various Types of Image Moments

Moment Order	**No. of Moments**	**Legendre Moment**	**Chebyshev Moments**	**Krawtchouk Moments**
5	21	87.5	89.6	80.3
10	66	89.4	90.4	81.4
15	136	90.2	89.7	82.8
20	231	88.5	88.8	84.5
25	351	87.7	87.3	85.7
30	496	87.2	86.7	86.9
35	666	85.6	85.3	88.5
40	861	84.6	84.6	89.6
45	1081	82.5	82.7	89.8
50	3126	81.4	81.1	89.8

Table 6.7: Classification Results Obtained by Random Forest Using Various Types of Image Moments

Moment Order	**No. of Moments**	**Legendre Moment**	**Chebyshev Moments**	**Krawtchouk Moments**
5	21	89.5	88.3	79.2
10	66	92.1	91.6	79.9
15	136	92.7	92.8	82.6
20	231	91.5	92.2	84.3
25	351	90.7	90.5	86.9
30	496	90.3	88.6	88.6
35	666	87.9	87.3	90.3
40	861	86.2	86.6	90.7
45	1081	83.5	85.1	91.0
50	3126	82.3	83.7	90.7

Table 6.8: Classification Results Obtained by XGBoost Using Various Types of Image Moments

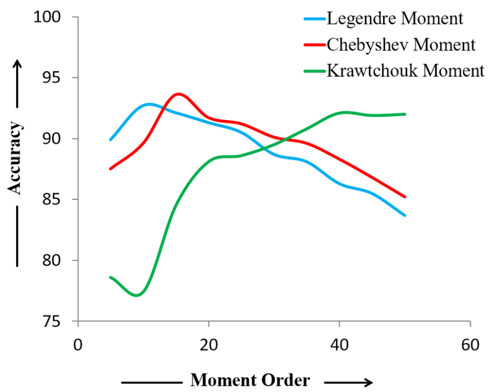
Moment Order	No. of Moments	Legendre Moment	Chebyshev Moments	Krawtchouk Moments
5	21	91.2	88.8	79.0
10	66	92.9	93.1	81.7
15	136	92.7	93.4	84.2
20	231	92.4	92.3	87.5
25	351	90.7	91.2	88.9
30	496	89.3	89.8	89.5
35	666	88.6	88.3	91.5
40	861	86.2	87.5	92.0
45	1081	83.7	86.3	91.6
50	1326	82.4	83.6	91.2

of various classifiers. It is observed that the classification accuracy generally increases with the moment order for Krawtchouk moments up to a certain level. However, for Legendre and Chebyshev moments, the classification accuracy tends to decrease as the moment order increases. As Legendre and Chebyshev moments are prone to sensitivity to noise, the introduction of additional information through higher moment orders may lead to ambiguity, resulting in a decrease in classification accuracy as the moment order increases. Krawtchouk moments exhibit less sensitivity to noise compared to the other two types of moments. This robustness enables Krawtchouk moments to effectively capture the underlying structure of the data even in the presence of noise. However, the classification accuracy increases only up to a certain level due to the effect of increasing the feature dimension. The highest overall classification results obtained by different classifiers are provided in Table 6.9.

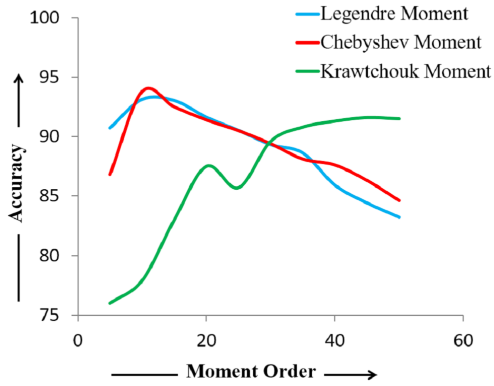
Table 6.9: Highest classification accuracy obtained by different classifiers using different moment orders

Classifier	Training Accuracy	Test Accuracy
ANN	93.2 (Chebyshev, order 15)	91.6%
SVM	93.5 (Chebyshev, order 10)	92.2%
kNN	90.4 (Legendre, order 10)	89.2%
Decision Tree	90.4 (Chebyshev, order 10)	89.5%
Random Forest	92.7 (Chebyshev, order 15)	91.2%
XGBoost	93.4 (Chebyshev, order 15)	91.9%

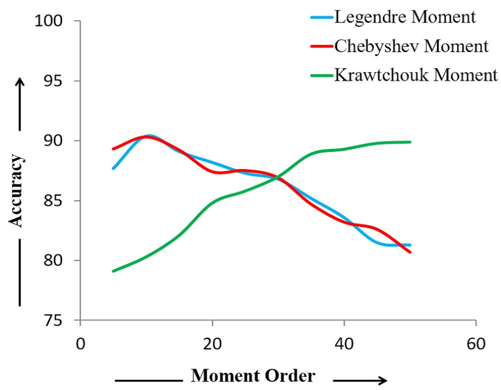
The test results for SVM and XGBoost in the form of confusion matrices and classification reports are shown in Figures 6-5 and 6-6.



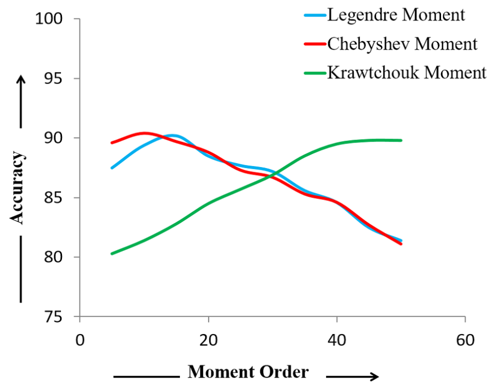
(a) Accuracy Vs Moment Order for ANN



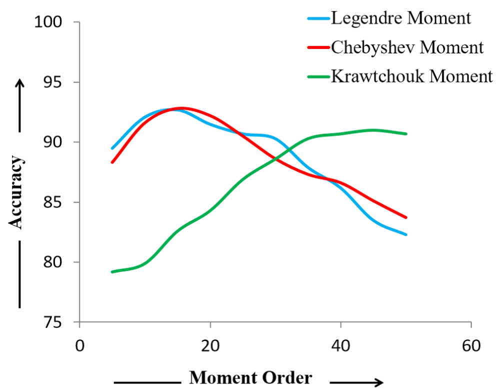
(c) Accuracy Vs Moment Order for SVM



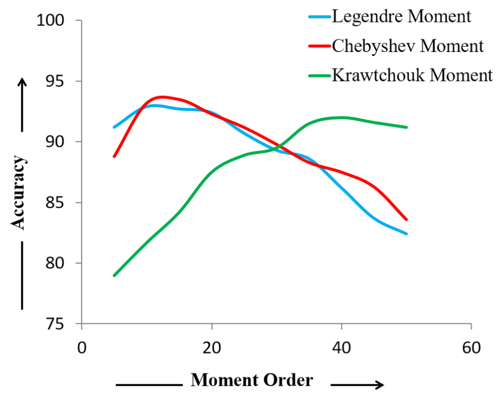
(e) Accuracy Vs Moment Order for kNN



(g) Accuracy Vs Moment Order for Decision Tree

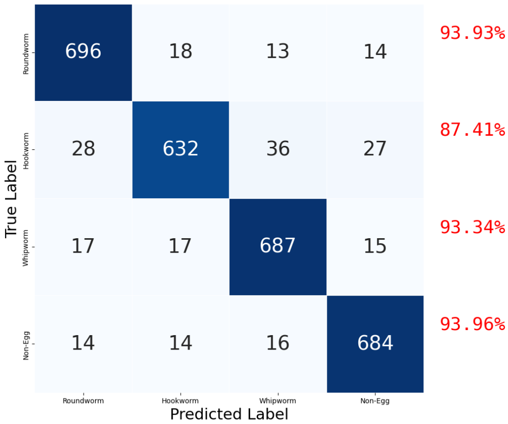


(i) Accuracy Vs Moment Order for Random Forest



(k) Accuracy Vs Moment Order for XGBoost

Figure 6-4: Change of classification accuracy with moment orders for different classifiers

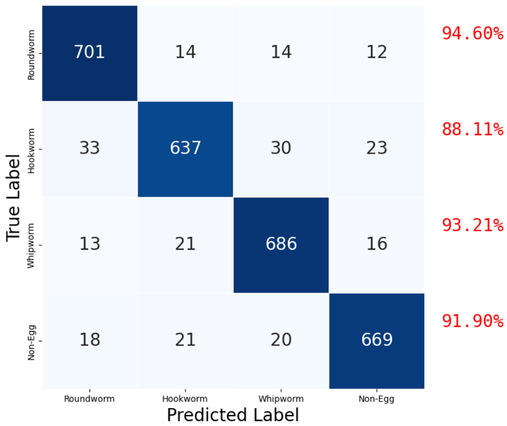


(a) (a) Confusion matrix

	precision	recall	f1-score	support
Roundworm	0.92	0.94	0.93	741
Hookworm	0.93	0.87	0.90	723
Whipworm	0.91	0.93	0.92	736
Non-Egg	0.92	0.94	0.93	728
accuracy			0.92	2928
macro avg	0.92	0.92	0.92	2928
weighted avg	0.92	0.92	0.92	2928

(c) (b) Classification report

Figure 6-5: Test result of SVM using 10^{th} order Chebyshev moment



(a) (a) Confusion matrix

	precision	recall	f1-score	support
Roundworm	0.92	0.95	0.93	741
Hookworm	0.92	0.88	0.90	723
Whipworm	0.91	0.93	0.92	736
Non-Egg	0.93	0.92	0.92	728
accuracy			0.92	2928
macro avg	0.92	0.92	0.92	2928
weighted avg	0.92	0.92	0.92	2928

(c) (b) Classification report

Figure 6-6: Test result of XGBoost using 15^{th} order Chebyshev moments

6.1.1.3 Use of Dimensionality Reduction Techniques for Higher-Order Moments

To optimize classification accuracy across different types of image moments, two widely used feature dimension reduction techniques, Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are applied. For Legendre and Chebyshev moments, both dimensionality reduction methods are applied to feature sets generated by the $10^t h$ and $15^t h$ orders, respectively, as these orders achieve the highest classification accuracy. Likewise, for Krawtchouk moments, the $45^t h$ order is used.

The optimal number of PCA components for a classifier is determined by evaluating its performance across different numbers of components within the range of 1 to 300. An example illustrating how the performance of a Random Forest classifier changes with the number of PCA components when using Chebyshev moments is shown in Figure 6-7. The highest training classification results achieved by various classifiers are presented in Table 6.10, along with the improvement rates with previous results. The improvement rate in classification accuracy for a classifier is calculated by comparing the highest accuracy achieved by a classifier using a specific moment without PCA to the accuracy obtained with PCA. The formula for computing this rate is represented by Equation 6.1.

$$\text{Percentage} = \frac{result_{PCA} - result_0}{result_{PCA}} \times 100 \tag{6.1}$$

Where $result_{PCA}$ represents the highest classification result obtained by the specific classifier using the specific image moment with PCA components, and $result_0$ represents the highest classification accuracy obtained by the same classifier using the same image moment (which may differ in moment order) without using PCA analysis.

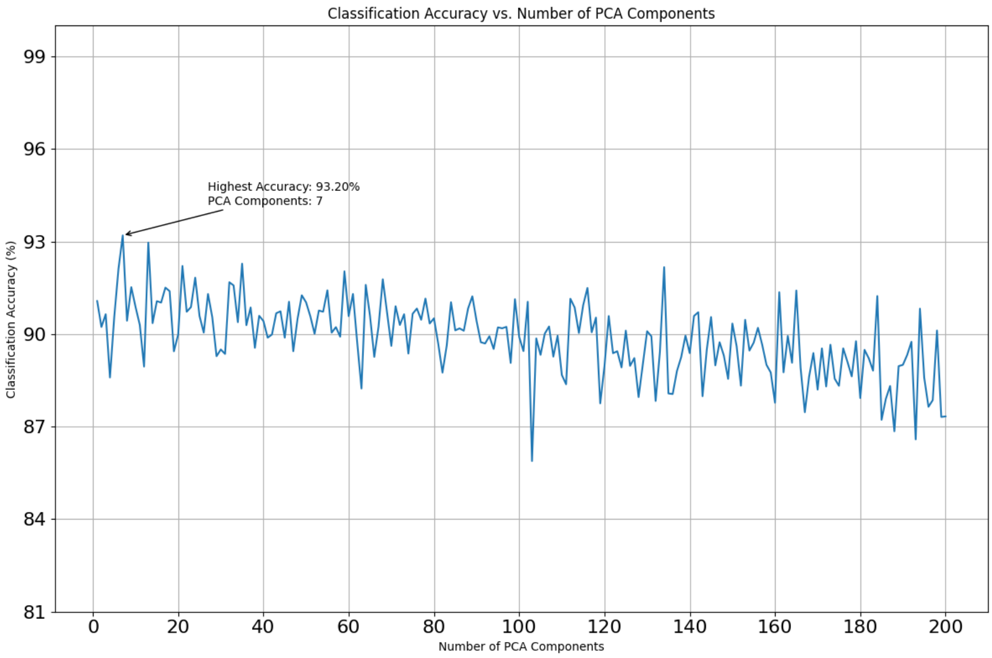


Figure 6-7: Change of classification accuracy with the number of PCA components for SVM with 50^{th} order Chebyshev Moments

Test results of the classification models with the highest training accuracy are presented in Table 6.11.

Similar to PCA, LDA is also employed to reduce the feature dimension. The reduced feature set is subsequently used to evaluate the performance of various classification algorithms. Table 6.12 shows the highest training accuracies achieved by various classifiers using different types of feature sets obtained after applying LDA, while Table 6.13 displays the test results of the best-fitted models.

Table 6.10: Highest training accuracy yielded by different classifiers using PCA on different types of image moments

Moment	**Classifier**	**No. of PCA components**	**Accuracy**	**Improvement in accuracy**
Legendre Moment	ANN	10	93.4%	0.75%
	SVM	8	94.5%	1.48%
	kNN	23	90.9%	0.55%
	Decision Tree	19	91.5%	1.42%
	Random Forest	20	93.2%	0.54%
	XGBoost	22	94.2%	1.38%
Chebyshev Moment	ANN	10	94.7%	1.58%
	SVM	9	95.2%	1.79%
	kNN	22	91.0%	0.77%
	Decision Tree	21	91.8%	1.53%
	Random Forest	19	93.6%	0.85%
	XGBoost	23	94.5%	1.16%
Krawtchouk Moment	ANN	15	93.3%	1.29%
	SVM	12	93.7%	2.24%
	kNN	28	91.3%	1.53%
	Decision Tree	22	91.7%	2.07%
	Random Forest	23	91.9%	0.98%
	XGBoost	27	93.2%	1.29%

Table 6.11: Highest test accuracy of different classifiers using moment-based features and PCA analysis

Classifier	**Train Accuracy**	**Test Accuracy**
ANN	94.7 (Legendre Moment)	93.6
SVM	95.2 (Chebyshev Moment)	94.4
kNN	91.3 (Krawtchouk Moment)	89.8
Decision Tree	91.8 (Chebyshev Moment)	90.3
Random Forest	93.6 (Chebyshev Moment)	92.4
XGBoost	94.5 (Chebyshev Moment)	93.5

The performances of all the classifiers using both dimensionality reduction methods are evaluated and compared for their effectiveness. Both the methods perform nearly identically, although the analysis as shown in Figure 6-8 suggests that PCA may be more beneficial than LDA when applied to the specified three types of image moments.

Table 6.12: Highest training accuracy yielded by different classifiers using LDA on different types of image moments

Moment	**Classifier**	**Accuracy**	**Improvement in accuracy**
Legendre Moment	ANN	93.8%	1.17%
	SVM	94.1%	1.06%
	kNN	90.9%	0.55%
	Decision Tree	91.1%	0.99%
	Random Forest	93.1%	0.43%
	XGBoost	93.7%	0.85%
Chebyshev Moment	ANN	94.1%	0.74%
	SVM	94.3%	1.27%
	kNN	91.5%	1.31%
	Decision Tree	91.5%	1.20%
	Random Forest	93.4%	0.64%
	XGBoost	94.9%	1.58%
Krawtchouk Moment	ANN	93.1%	1.07%
	SVM	92.9%	1.40%
	kNN	90.5%	0.66%
	Decision Tree	91.3%	1.64%
	Random Forest	91.9%	0.98%
	XGBoost	92.8%	1.19%

Table 6.13: Highest test accuracy of different classifiers using moment-based features and LDA analysis

Classifier	**Train Accuracy**	**Test Accuracy**
ANN	94.1 (Chebyshev Moment)	93.0
SVM	94.3 (Chebyshev Moment)	93.1
kNN	91.5 (Chebyshev Moment)	90.2
Decision Tree	91.5 (Chebyshev Moment)	89.6
Random Forest	93.4 (Chebyshev Moment)	91.7
XGBoost	94.9 (Chebyshev Moment)	93.2

6.1.2 Classification using Texture and Shape-based Features

Texture and shape-based features are extracted from the segmented greyscale images, as mentioned in Chapter 5. In Table 6.14, a summary is presented, showing the overall training and testing accuracy achieved through the training of various classifiers using the texture and shape-based feature set. For a more granular insight into the classification performance, a detailed report of test accuracy for each class is provided in Table 6.15. The experiments reveal that all classifiers produce satisfactory results, surpassing the experiments of earlier sections. In

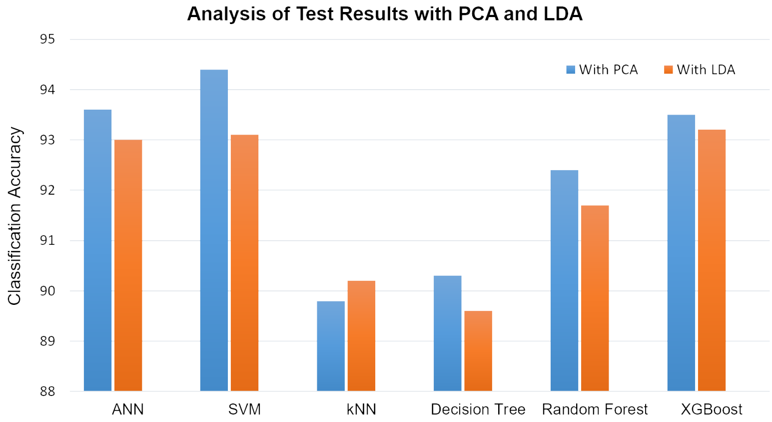


Figure 6-8: Performance analysis of different classifiers using three types of image moments with PCA and LDA

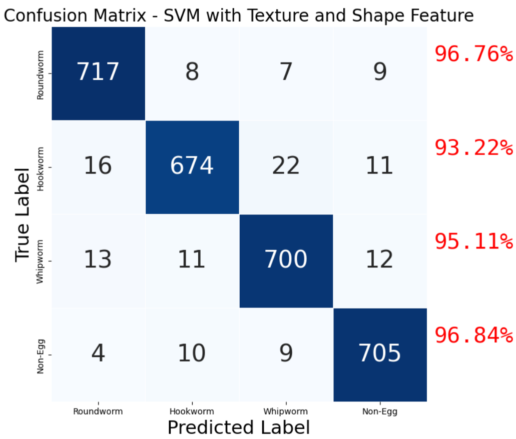
general, SVM and XGBoost exhibit a slight edge over the other classifiers in overall performance. The confusion matrix and classification report generated by the testing process of both the models are shown in Figure 6-9.

Table 6.14: Overall classification accuracy of various classifiers using texture and shape-based features

	ANN	**SVM**	**kNN**	**Decision Tree**	**Random Forest**	**XG-Boost**
Training	95.8%	96.8%	94.5%	94.8%	95.3%	96.2%
Testing	94.7%	95.5%	93.2%	93.2%	94.3%	94.7%

Table 6.15: Classification accuracy (test) for each class using texture and shape-based features

	ANN	**SVM**	**kNN**	**Decision Tree**	**Random Forest**	**XG-Boost**
Round-worm (Ascaris)	96.22	96.76	94.20	94.20	95.95	95.01
Hook-worm (Neca-tor)	90.87	93.22	91.01	90.59	91.15	92.39
Whip-worm (Trichuris)	95.24	95.11	94.16	96.20	94.70	96.20
Non-Egg	96.57	96.84	93.68	94.78	95.19	95.05

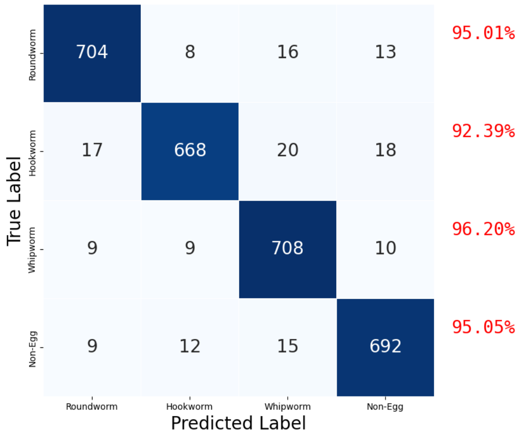


(a) (a) Confusion Matrix

	precision	recall	f1-score	support
Roundworm	0.96	0.97	0.96	741
Hookworm	0.96	0.93	0.95	723
Whipworm	0.95	0.95	0.95	736
Non-Egg	0.96	0.97	0.96	728
accuracy			0.95	2928
macro avg	0.95	0.95	0.95	2928
weighted avg	0.95	0.95	0.95	2928

(c) (b) Classification report

Figure 6-9: Test result of SVM using Texture and Shape-based features



(a) (a) Confusion Matrix

	precision	recall	f1-score	support
Roundworm	0.95	0.95	0.95	741
Hookworm	0.96	0.92	0.94	723
Whipworm	0.93	0.96	0.95	736
Non-Egg	0.94	0.95	0.95	728
accuracy			0.95	2928
macro avg	0.95	0.95	0.95	2928
weighted avg	0.95	0.95	0.95	2928

(c) (b) Classification report

Figure 6-10: Test result of XGBoost using Texture and Shape-based features

6.1.3 Classification using Pixel Intensity-based Features

Pixel intensity-based features, such as standard deviation, variance, maximum, minimum, and mean pixel values, are computed by dividing segmented images into multiple blocks of tiles and extracting features from each block, as discussed in Chapter 5. In this study, the segmented images are resized to 120×120 pixels, and experiments are performed using various block sizes for partitioning. The overall classification accuracies of different classifiers are presented in Table 6.16. An analysis is conducted to assess the impact of the block size or feature dimension on classification accuracy, which can be visualised through Figure 6-11. It is observed that the classification accuracy tends to decrease as the block size increases. It suggests that smaller blocks provide more detailed information about the intensity distribution. Extracting information from multiple smaller regions contributes to the creation of a more robust feature vector. Conversely, larger block sizes cover a larger area, which leads to extracting less granular information. The test results of the classifiers using 10 block-sized feature sets are presented in Table 6.17. Among all the classifiers, ANN and SVM achieve the highest accuracy during the testing process. The results of both classifiers can be visualised in the form of a confusion matrix and classification report as shown in Figures 6-12 and 6-13.

Table 6.16: Training accuracy of different classifiers using Pixel intensity-based features for Four Classes

Block Size	**Feature Dimension**	**ANN**	**SVM**	**kNN**	**Decision Tree**	**Random Forest**	**XG-Boost**
10×10	720	95.3%	95.4%	92.5%	94.1%	94.7%	95.3%
20×20	180	94.7%	94.6%	92.6%	92.9%	93.9%	94.9%
30×30	80	92.6%	93.1%	91.3%	91.3%	92.3%	93.5%
40×40	45	91.3%	91.8%	90.4%	90.7%	91.3%	91.6%
60×60	20	88.8%	89.7%	87.8%	88.5%	89.0%	90.0%

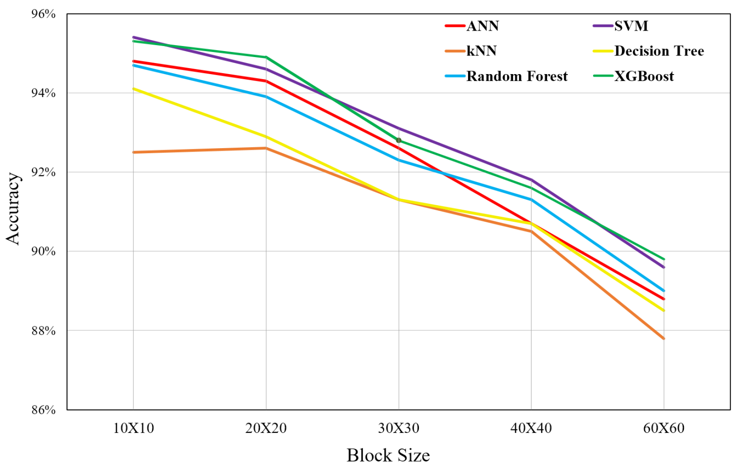
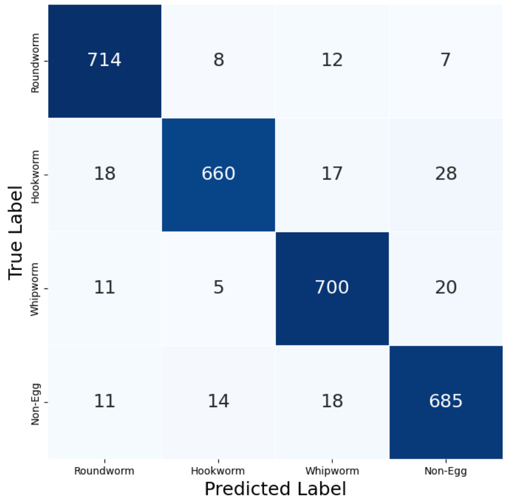


Figure 6-11: Change of Classification Accuracy with Size of the Blocks or Dimension of Pixel-Intensity-based Feature set

The feature vector generated by 10×10 block sizes has a high dimension of 720, which may affect the performance of a classification algorithm. Therefore, PCA analysis is applied to evaluate whether the accuracy improves or not. The number of PCA components for each classification algorithm is determined empirically as discussed in Section 6.1.1.3. This improves the results a little; however, no significant improvement in the classification accuracy is found. The results of PCA analysis are shown in Table 6.18. Test accuracy of each class of objects is also recorded and provided in Table 6.19.

Table 6.17: Test accuracy of different classifiers for each class using pixel intensity-based features

	ANN	SVM	kNN	Decision Tree	Random Forest	XG-Boost
Round-worm	96.36%	94.47%	92.12%	93.26%	95.11%	94.94%
Hook-worm	91.29%	92.81%	89.25%	90.15%	91.52%	90.93%
Whip-worm	95.11%	95.79%	90.84%	92.44%	93.36%	% 95.36
Non-Egg	94.09%	93.41%	91.15%	92.78%	92.85%	% 94.72
Overall Accuracy	94.23%	94.12%	90.84%	92.16%	93.21%	93.98%



(a) (a) ANN

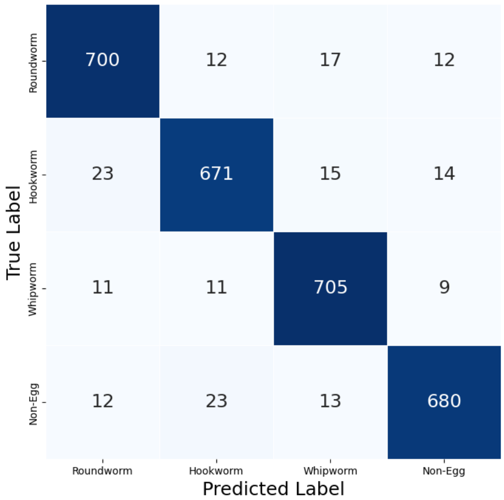
	precision	recall	f1-score	support
Roundworm	0.95	0.96	0.96	741
Hookworm	0.96	0.91	0.94	723
Whipworm	0.94	0.95	0.94	736
Non-Egg	0.93	0.94	0.93	728
accuracy			0.94	2928
macro avg	0.94	0.94	0.94	2928
weighted avg	0.94	0.94	0.94	2928

(c) (b) SVM

Figure 6-12: Test result of ANN using pixel intensity-based features

6.1.4 Classification Using Combinations of Different Feature Sets

The classification results presented in the previous sections show that the texture and shape-based and pixel intensity-based features are better choices for any classifier compared to image moment-based features. Although satisfactory test results are achieved using these feature sets, further experimentation is conducted to improve the performance of the classifiers by combining different types of feature sets. In this work, two or more distinct feature vectors are combined to create a single feature vector. As the Chebyshev, Legendre, and Krawtchouk moments produced varying results with different moment orders, the feature sets that yielded the highest accuracy through PCA analysis are utilized. Similarly,



(a) (a) ANN

```
              precision    recall  f1-score   support

   Roundworm       0.94      0.94      0.94       741
    Hookworm       0.94      0.93      0.93       723
    Whipworm       0.94      0.96      0.95       736
     Non-Egg       0.95      0.93      0.94       728

    accuracy                           0.94      2928
   macro avg       0.94      0.94      0.94      2928
weighted avg       0.94      0.94      0.94      2928
```

(c) (b) SVM

Figure 6-13: Test result of SVM using pixel intensity-based features

Table 6.18: Classification accuracy for different classifiers using PCA analysis on pixel intensity-based features

	ANN	**SVM**	**kNN**	**Decision Tree**	**Random Forest**	**XG-Boost**
PCA Component	16	13	24	23	24	18
Train Accuracy	96.2	96.0	93.4	94.9	95.8	96.4
Test Accuracy	94.2	94.9	92.3	92.5	93.9	94.4
Improvement rate (Train)	0.94%	0.62%	0.96%	0.84%	1.15%	1.14%
Improvement rate (Test)	0.74%	0.96%	0.66%	0.84%	0.85%	0.1%

Table 6.19: Classification accuracy (Test) for each class using pixel intensity-based features using PCA

	ANN	**SVM**	**kNN**	**Decision Tree**	**Random Forest**	**XG-Boost**
Round-worm (Ascaris)	94.47	96.22	93.12	93.39	95.28	95.28
Hook-worm (Necator)	92.67	92.95	90.46	90.18	91.84	92.95
Whip-worm (Trichuris)	95.24	95.38	93.75	94.97	94.29	95.24
Non-Egg	94.37	95.05	92.03	91.35	93.96	94.23

for the pixel intensity-based feature set, the feature vector obtained using PCA analysis is also used. Table 6.20 shows overall training accuracy obtained by various classifiers using different combinations of features. Based on the results, the highest classification results achieved by the different classifiers are summarized in Table 6.21. It is observed that a combination of Hu moments, pixel intensity, and texture and shape-based features achieves the highest results with the majority of classifiers. An analysis is conducted to visualise the improvement in accuracy of different classifiers compared to the highest results from previous experiments, as illustrated in Figure 6-14. The analysis shows a little improvement in overall classification results using a combined feature set.

Table 6.20: Training classification accuracy using combinations of different feature sets

Feature Combination	**ANN**	**SVM**	**kNN**	**Decision Tree**	**Random Forest**	**XGBoost**
Hu + Pixel intensity	94.8	95.9	93.0	94.6	93.6	95.4
Hu+Texture	95.7	96.1	93.1	94.1	95.3	96.3
Legendre + Pixel intensity	95.4	95.6	92.6	93.8	93.8	94.8
Legendre + Texture	94.7	95.8	92.6	93.7	93.8	95.6
Chebyshev + Pixel intensity	95.2	95.7	92.4	91.1	94.7	95.9
Chebyshev + Texture	93.6	94.9	92.9	93.3	95.1	95.2
Krawtchouk + Pixel intensity	93.3	93.8	93.2	91.8	92.8	92.9
Krawtchouk + Texture	94.2	93.7	92.3	89.4	91.7	94.3
Pixel intensity + Texture	95.9	96.7	94.5	95.6	95.8	96.3
Hu + Pixel intensity + Texture	96.1	96.8	94.4	95.7	95.8	96.8

Table 6.21: Highest classification results of the classifiers with combination of different features

Classifier	Feature	Training Accuracy	Test Accuracy
ANN	Hu + Pixel intensity+Texture	96.1	94.7
SVM	Hu + Pixel intensity+Texture	96.8	95.6
kNN	Pixel intensity + Texture	94.5	93.1
Decision Tree	Hu + Pixel intensity + Texture	95.7	94.2
Random Forest	Pixel intensity + Texture	95.8	94.5
XGBoost	Hu + Pixel intensity + Texture	96.8	95.3

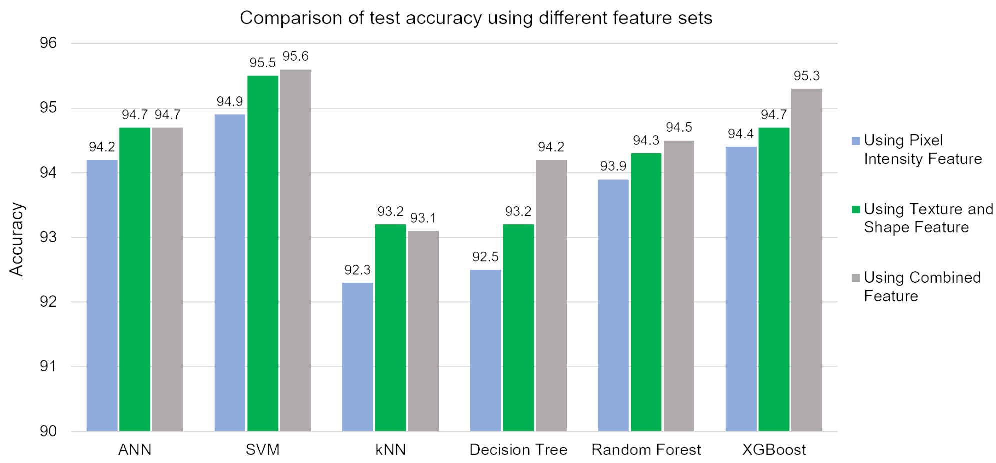
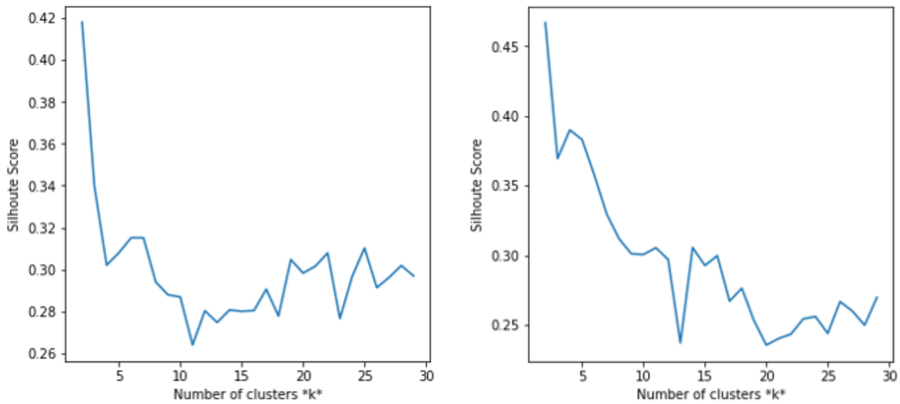


Figure 6-14: Analysis of test accuracy of different classifiers with individual vs combined feature sets

6.2 Classification using Multiple Classes of Non-Egg Objects

During the research, it is observed that many previous works do not consider the classification accuracy of non-egg objects or exclude them from their classification tasks. However, it is crucial to include non-egg objects to effectively distinguish them from various types of parasite eggs. So far, this study uses only a single class of non-egg objects obtained from the segmentation process. Since non-egg objects lack the regular shape or texture of parasite eggs, assigning all non-egg objects to a single class may increase the misclassification rate, especially with the introduction of new species of parasite eggs in the future. Moreover, it is observed that the identification rate of Hookworm eggs is relatively lower than the other two types of parasite eggs. To address these issues, the non-egg objects are divided into multiple classes. In order to do so, clustering methods, including K-means and the Gaussian Mixture Model (GMM) are used using pixel intensity and texture-shape-based features. These features are chosen due to their robustness, as observed in the previous sections. The results are then analyzed to determine the optimal number of clusters for the non-egg objects based on the silhouette score. Figures 6-15 and 6-16 show the change in silhouette scores with respect to the number of clusters for K-Means and Gaussian Mixture Model, respectively.

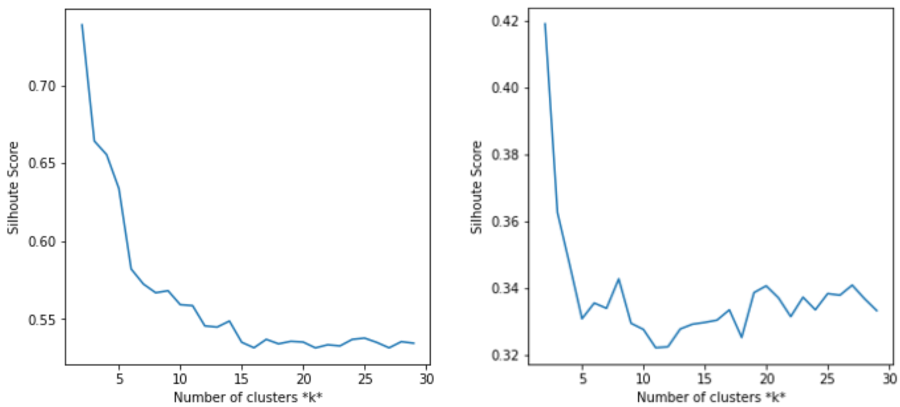
The analysis of silhouette scores indicates that the optimal number of clusters falls between two and three, as suggested by both methods. In this study, all segmented images of Non-Egg objects from the preceding sections are used. In the case of two clusters, Cluster-1 comprises 2218 images, while Cluster-2 en-



(a) (a) Silhouette score Vs number of clusters using KMeans with Texture and Shape-based features

(c) (b) Silhouette score Vs number of clusters using KMeans with Pixel Intensity-based feature sets

Figure 6-15: Analysis of Silhouette score for finding optimum number of clusters using KMeans



(a) (a) Silhouette score Vs number of clusters using GMM with Texture and Shape-based features

(c) (b) Silhouette score Vs number of clusters using GMM with Pixel Intensity-based features

Figure 6-16: Analysis of Silhouette score for finding optimum number of clusters using Gaussian Mixture Model

compasses 1505 images. For three clusters, the distribution is: 1008 images for Cluster-1, 1256 images for Cluster-2, and 1449 images for Cluster-3. A few images of different clusters are shown in Figures 6-17, 6-18 and 6-19. To increase the number of samples in each class, a few data augmentation techniques are applied. These include horizontal and vertical flipping, rotation, and blurring. Following the data augmentation, some augmented images that looked similar to the original images, such as those containing air bubbles and round-shaped objects, are carefully discarded. Finally, a total of 18,117 segmented objects, including parasite eggs and non-egg objects are selected while considering two clusters of non-egg objects. In the three-cluster approach, the total number of objects used in classification is 21,717. Table 6.22 presents the distribution of objects across each class for classification tasks.

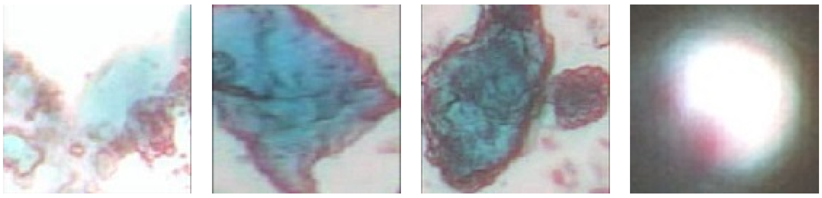


Figure 6-17: Non-Egg objects in cluster-1 with 3 optimal clusters

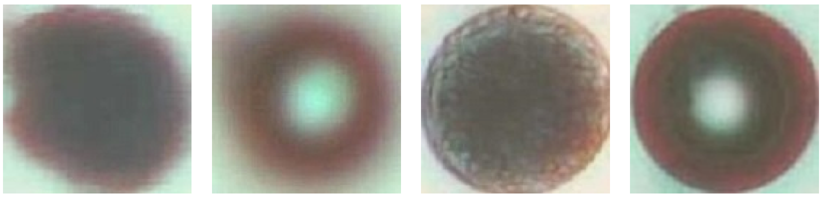


Figure 6-18: Non-Egg objects in cluster-2 with 3 optimal clusters

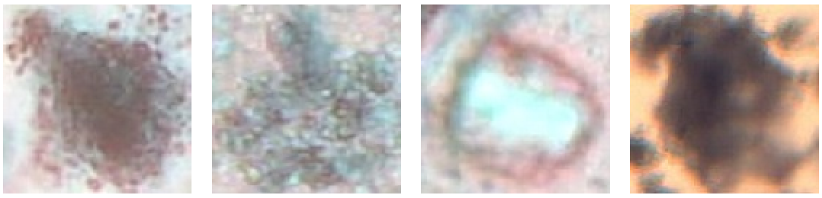


Figure 6-19: Non-Egg objects in cluster-3 with 3 optimal clusters

Table 6.22: Number of segmented images in each class for classification task using multiple non-egg classes

Class →	**Round-worm**	**Hook-worm**	**Whip-worm**	**NonEgg-1**	**NonEgg-2**	**NonEgg-3**
Two clustering approach	3655	3618	3644	3600	3600	x
Three clustering approach	3655	3618	3644	3600	3600	3600

Using two classes of Non-Egg objects: The highest training classification results from our experimentation with five classes of objects are shown in Table 6.23. Based on the results, an analysis of the classification results of different classifiers with various feature sets is performed as shown in Figure 6-20. From the analysis, it is observed that the texture and shape-based feature set yields better classification results than the others. The test results of various classifiers with the texture-based feature set are presented in Table 6.24. The confusion matrix

Table 6.23: Training results of various classifiers using Two classes of Non-eggs and Three parasite eggs

Classifier / Feature	ANN	SVM	kNN	Decision Tree	Random Forest	XGBoost
Legendre Moments	86.2	85.5	78.8	76.5	82.3	85.7
Chebyshev Moments	85.1	86.2	77.5	75.8	81.6	86.5
Krawtchouk Moments	82.1	83.3	80.2	76.0	81.3	81.8
Hu Moments	85.7	82.8	84.3	77.1	81.2	83.6
Pixel Intensity	89.1	85.5	85.1	78.6	88.2	90.2
Texture and Shape	89.3	91.3	84.6	82.5	89.6	90.8

Table 6.24: Test accuracy for each class using texture and shape-based features while trained for five classes of objects

	ANN	SVM	kNN	Decision Tree	Random Forest	XG-Boost
Round-worm	92.62	95.49	87.16	85.79	93.44	94.13
Hook-worm	90.06	93.23	85.22	84.53	91.44	93.09
Whip-worm	92.46	95.06	86.56	84.36	93.55	94.79
Non-Egg 1	82.92	83.75	78.89	75.14	83.19	84.17
Non-Egg 2	80.64	83.98	76.46	76.04	81.89	82.31
Overall Accuracy	87.77	90.34	82.86	81.17	88.74	89.73

and classification report, indicating precision, recall, and F1 score for each class, obtained from SVM and XGBoost, are shown in Figures 6-21 and 6-22.

Using three classes of Non-Egg objects: Training results of various classifiers using six classes of objects: three parasite eggs and three non-eggs, are presented in Table 6.25. Similar to the previous experimentation, analysis of classification accuracies shows that the texture and shape-based feature set also yields better performance when using six classes of objects. The test results of all the classifiers are tabulated in Table 6.26. It is observed that all the classifiers performed equally, with ANN, SVM, and Random Forest showing slight advantages. Examples of confusion matrices and classification reports for two models are shown in Figures 6-23 and 6-24, respectively.

A comparison of the overall classification results of the classifiers when

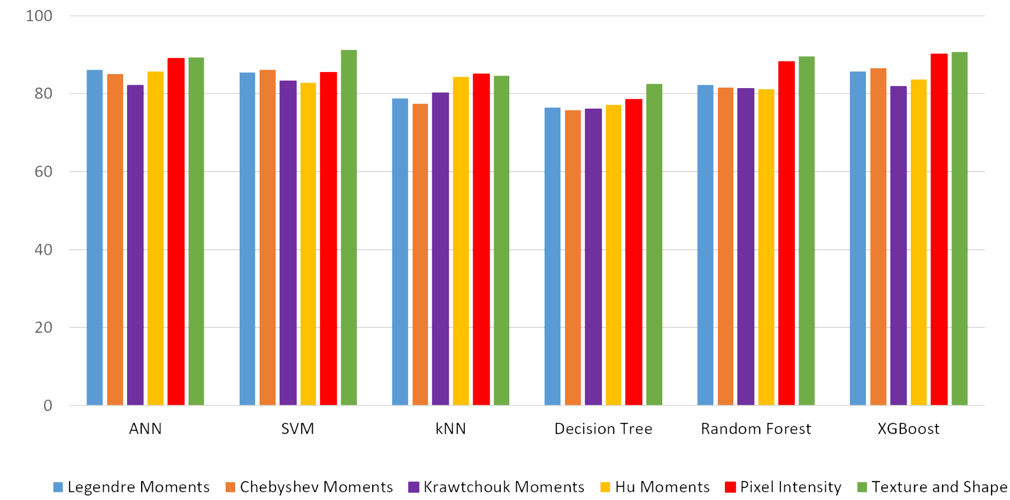
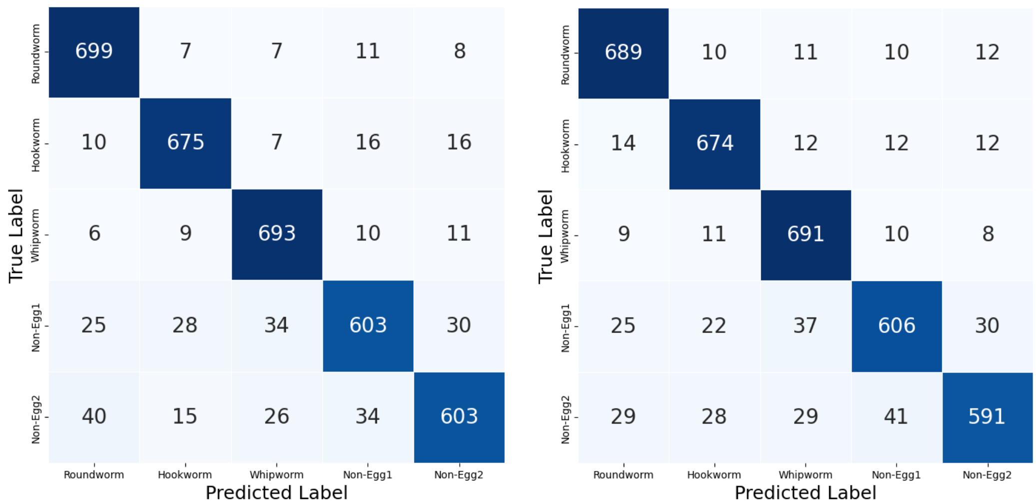


Figure 6-20: Analysis of classification accuracy of different classifiers with different features while using two classes of non-eggs and three parasite eggs



(a) SVM (c) XGBoost

Figure 6-21: Confusion matrix obtained from testing SVM and XGBoost models using five classes of objects

	precision	recall	f1-score	support
Roundworm	0.90	0.95	0.92	732
Hookworm	0.92	0.93	0.93	724
Whipworm	0.90	0.95	0.93	729
Non-Egg1	0.89	0.84	0.87	720
Non-Egg2	0.90	0.84	0.87	718
accuracy			0.90	3623
macro avg	0.90	0.90	0.90	3623
weighted avg	0.90	0.90	0.90	3623

(a) SVM

	precision	recall	f1-score	support
Roundworm	0.90	0.94	0.92	732
Hookworm	0.90	0.93	0.92	724
Whipworm	0.89	0.95	0.92	729
Non-Egg1	0.89	0.84	0.87	720
Non-Egg2	0.91	0.82	0.86	718
accuracy			0.90	3623
macro avg	0.90	0.90	0.90	3623
weighted avg	0.90	0.90	0.90	3623

(c) XGBoost

Figure 6-22: Classification report obtained from testing SVM and XGBoost models using five classes of objects

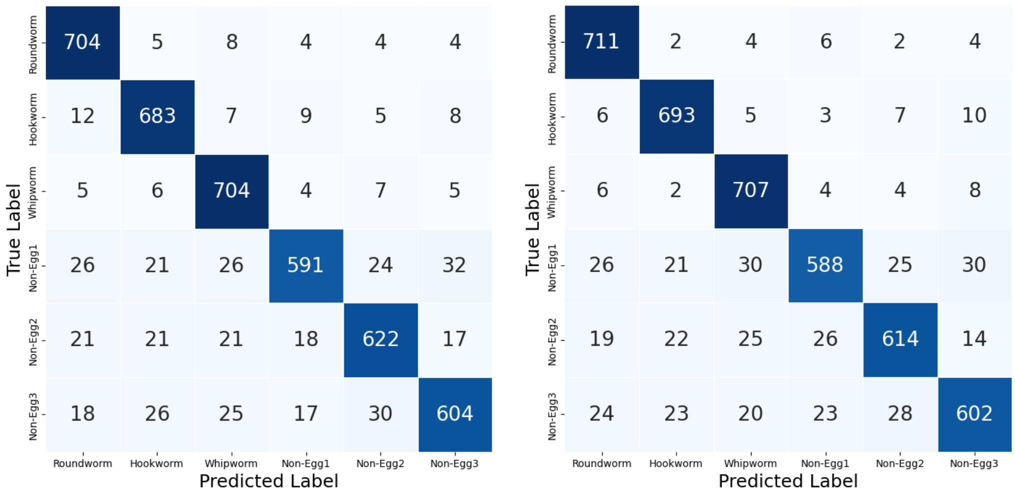
Table 6.25: Training results of various classifiers while using three classes of Non-eggs and three parasite eggs

Classifier / Feature	ANN	SVM	kNN	Decision Tree	Random Forest	XGBoost
Legendre Moments	83.2	82.3	80.8	76.5	83.3	84.2
Chebyshev Moments	84.1	84.2	81.9	81.8	82.8	85.4
Krawtchouk Moments	80.7	81.5	79.1	82.1	80.1	84.8
Hu Moments	83.3	82.7	79.4	80.6	82.2	79.5
Pixel Intensity	89.4	90.2	88.3	87.5	88.3	88.6
Texture and Shape	91.4	91.3	89.5	90.3	91.3	90.9

Table 6.26: Test accuracy for each class using texture and shape-based features while trained for six classes of objects

	ANN	SVM	kNN	Decision Tree	Random Forest	XG-Boost
Round-worm	96.57	97.53	94.65	93.28	96.57	96.42
Hook-worm	94.06	95.72	92.04	92.44	94.34	94.78
Whip-worm	95.90	96.72	93.30	94.39	96.31	96.71
Non-Egg 1	83.61	81.67	83.01	82.19	82.08	81.33
Non-Egg 2	86.25	85.28	84.03	85.97	86.39	86.27
Non-Egg 3	82.78	83.61	81.53	81.53	83.89	83.11
Overall Accuracy	89.89	90.12	88.12	88.33	89.95	89.80

utilizing single, two, and three non-egg classes of objects is made and shown in Figure 6-25. Experimenting with multiple non-egg classes reveals that the overall accuracy of the classifiers decreases compared to the previous experiments using a single non-egg class. However, a comparative analysis of the identification rates of various classes of parasite eggs indicates an improvement in the identification rate of hookworm eggs. The difference in the identification rate of each class of parasite egg while using single and multiple non-egg classes is illustrated in Table 6.27.



(a) Random Forest **(c)** SVM

Figure 6-23: Confusion matrix obtained from testing Random Forest and SVM using six classes of objects

	precision	recall	f1-score	support
Roundworm	0.91	0.97	0.94	729
Hookworm	0.90	0.95	0.93	724
Whipworm	0.88	0.96	0.92	731
Non-Egg1	0.90	0.84	0.87	720
Non-Egg2	0.92	0.86	0.89	720
Non-Egg3	0.90	0.82	0.86	720
accuracy			0.90	4344
macro avg	0.90	0.90	0.90	4344
weighted avg	0.90	0.90	0.90	4344

(a) Random Forest

	precision	recall	f1-score	support
Roundworm	0.90	0.98	0.93	729
Hookworm	0.91	0.96	0.93	724
Whipworm	0.89	0.97	0.93	731
Non-Egg1	0.90	0.82	0.86	720
Non-Egg2	0.90	0.85	0.88	720
Non-Egg3	0.90	0.84	0.87	720
accuracy			0.90	4344
macro avg	0.90	0.90	0.90	4344
weighted avg	0.90	0.90	0.90	4344

(c) SVM

Figure 6-24: Classification report obtained from testing Random Forest and SVM using six classes of objects

Figure 6-25: Analysis of classification accuracy of different classifiers using single and multiple non-egg classes of objects

Table 6.27: Classification accuracy of different classes of parasite eggs with single and multiple non-egg classes

Parasite Egg	Highest accuracy using a single Non-Egg class	Highest accuracy using multiple Non-Egg class
Roundworm	96.76	97.53
Hookworm	93.22	95.72
Whipworm	96.20	96.72

6.3 Result Analysis and Discussion

The performances of the classifiers with different feature sets are analyzed and illustrated in Figure 6-26. The analysis shows that, among all the classifiers, SVM and XGBoost demonstrated marginally superior performance. It is observed that texture and shape-based features yield the highest test accuracy with SVM. Moreover, combining Hu moments, texture and shape features, and pixel intensity-based features results in a slightly higher accuracy with SVM and XGBoost.

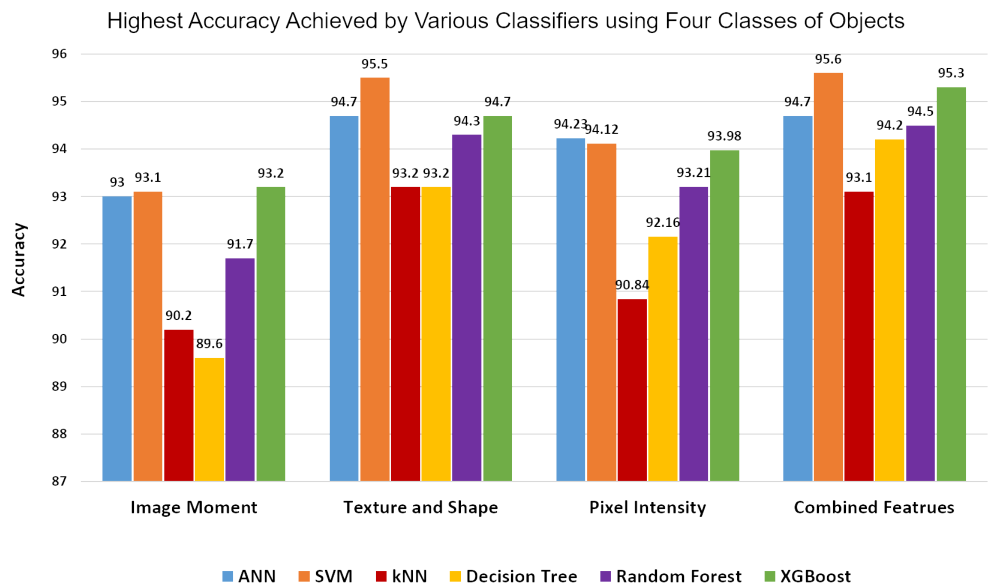
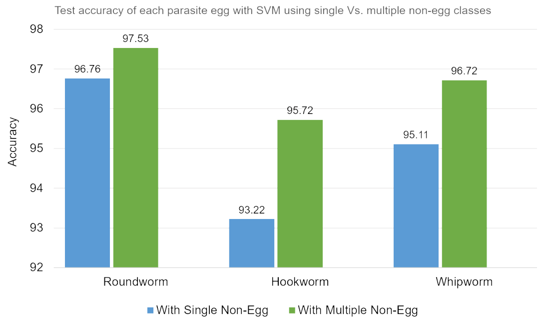


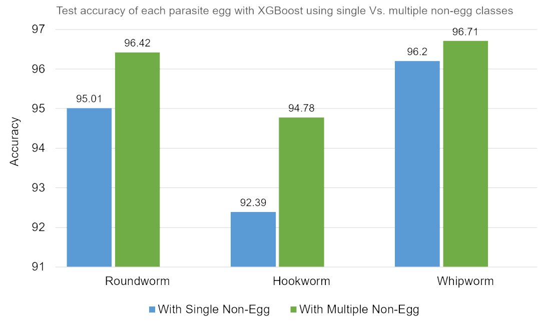
Figure 6-26: Analysis of classification accuracy of different classifiers using a single non-egg object and three parasite eggs

The incorporation of multiple non-egg classes in classification tasks reveals a decrease in overall classification accuracy. However, it notably enhances the classifier's ability to identify parasite eggs, particularly in the case of hookworms. A comparison of the classification accuracy of each parasite egg class when using multiple non-egg classes versus a single non-egg class is made as shown in Figure 6-27. This comparative analysis focuses on the results obtained by SVM and

XGBoost, as these classifiers demonstrate superior performance using texture and shape-based features.



(a) The classification accuracy of various parasite egg classes achieved by SVM using single and multiple non-egg classes



(b) The classification accuracy of various parasite egg classes achieved by XGBoost using single and multiple non-egg classes

Figure 6-27: A comparison of the classification accuracy of different parasite egg classes using single and multiple non-egg classes

This decline in classifier accuracy using multiple non-egg classes is likely due to the irregular properties of non-egg objects, which vary widely in size, shape, colour, and texture. However, the findings suggest that instead of grouping all non-egg objects into a single class, dividing them into two or more classes may improve the classification accuracy of parasite eggs.

6.4 Comparison with Previous Works

The classification accuracy obtained from this study is compared with several well-known works in this field. Table 6.28 shows the results of the various previous works and the findings from our study.

Table 6.28: Comparison of results with similar works in the past

Article	**Number of Classes**	**Classifier**	**Result**	**Remarks**
Yang et al. (2001) [9]	7	ANN	90%	Only a small number of images were used in the experiment

Dogantekin et al. (2008) [33]	16	Adaptive network-based fuzzy interface system	93.5%	It is not clear how many unique images are used by the authors to train the classifier. Datasets for training and testing are created by rotating the images from 0^0 to 165^0 in the steps of 15^0
Avci and Varol (2009) [34]	16	SVM	97%	It is not clear how many unique images are used by the authors to train the classifier. Datasets for training and testing are created by rotating the images from 0^0 to 165^0 in the steps of 15^0, which produced 120 images for each type of parasite egg.
Bruun et al. (2012) [76]	2	Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA)	92.7%	282 segmented images of a single type of parasite egg are used, where one class contains 249 and the other one contains 33 images.
Hadi et al. (2012) [2]	2	Threshold with Logical Classification Method	93% and 94%	100 images for each type of parasite egg are used.
Zhang et al. (2014) [63]	3	SVM	83.3%	
Li et al. (2015) [77]	6		95%	It is not clear how many images are used for training, but the article mentioned that images from 20 cases for each of the two types of parasite eggs and impurities are used.

Abdalla and Sekar (2017) [39]	11	ANN	96.6%	Two datasets are used, where one contains 4402 images of 7 species of Eimeria and the other contains 2902 images of 11 species of Eimeria. However, it is observed that both datasets are not well balanced for each class.
N. A. Khairudin et al. (2020) [74]	2	Logical if-else method based on area and size	84% and 76%	50 images of each type of parasite egg are used in classification.
Sandra Valeria Inacio et al. (2020) [93]	4	SVM	Kappa index = 0.7636	The study used a sufficient number of images, which is 10.699. However, the collection has a significantly low amount of parasite egg images. Almost 70% of the samples are fecal impurities, and the rest contain images of three types of parasite eggs.
Our study (with a single non-egg class)	4 (3 parasite eggs and 1 non-egg)	SVM	Overall accuracy 96.8%	The dataset contains a satisfactory quantity of images that are balanced for each class by using various data augmentation techniques.

Our study (with three non-egg classes)	6 (3 parasite eggs and 3 non-egg)	SVM	97.5%, 95.7% and 96.7% respectively for the three different types of parasite eggs	The dataset has a satisfactory quantity of images, which is balanced for each class by using various data augmentation techniques.

6.5 Conclusion

In this chapter, the feasibility and effectiveness of several machine-learning algorithms for classifying parasite eggs in noisy microscopic images of fecal samples from pigs are discussed. Various classifiers are examined, and their effectiveness is analyzed using different feature sets. The key contributions and observations of the chapter are outlined below:

- Several classifiers, including ANN, SVM, kNN, Decision Tree, Random Forest, and XGBoost are examined using diverse feature sets and dimensionality reduction methods such as PCA and LDA.
- Various comparative analysis are conducted, assessing both overall accuracy and individual class accuracy. The analysis shows the importance of feature selection in improving the classification accuracy of different types of parasite eggs.
- The experimentation shows that texture- and shape-based feature sets outperform other types of feature sets. Pixel intensity-based features also produce promising results. Specifically, SVM achieves nearly 96.5% accuracy with texture- and shape-based features and 95% accuracy with the pixel intensity-based features.
- An experiment is conducted by merging multiple different types of feature

sets together. SVM achieves 96.8% overall accuracy with combined texture-shape and pixel intensity-based features.

- The non-egg objects are divided into multiple classes using clustering methods. This approach is proposed to improve the identification rate of hookworm eggs and reduce the overall misclassification rate of parasite eggs.

In summary, this chapter provides valuable insights into using machine learning classifiers for parasite egg classification and highlights the effectiveness of various features. The methodologies and findings presented can serve as a foundation for further exploration and enhancement of machine learning techniques in this field.