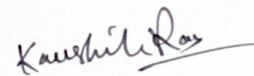


Dedicated to
My beloved Mother and Father

Declaration

I certify that

- The work contained in the dissertation is original and has been done by myself under the general supervision of my supervisors.
- The work has not been submitted to any other Institute for any degree or diploma.
- I have followed the guidelines provided by Tezpur University in writing the thesis.
- I have conformed to the norms and guidelines given in the Ethical Code of Conduct of the university.
- Whenever I have used materials (data, theoretical analysis, and text) from other sources, I have given due credit to them by citing them in the text of the dissertation and giving their details in the references.



Kaushik Ray



Department of Computer Science & Engineering
Tezpur University

Napaam, Tezpur- 784028, Assam, India.

Dr. Sarat Saharia
Professor

Phone: 03712-275105

E-Mail : sarat@tezu.ernet.in

Certificate

This is to certify that the thesis entitled "**Automatic Detection and Identification of Parasite Eggs in Noisy Microscopic Images of Fecal Samples**" submitted to Tezpur University in the Department of Computer Science and Engineering under the School of Engineering in partial fulfillment of the award of the degree of Doctor of Philosophy in Computer Science and Engineering is a record of research work carried out by **Kaushik Ray** under my supervision and guidance.

All helps received by him from various sources have been duly acknowledged. No part of this thesis has been submitted else where for award of any other degree.

A handwritten signature in black ink, appearing to read 'Sarat Saharia'.

Signature of Supervisor
(Dr. Sarat Saharia)

Professor

Department of Computer Science and Engineering
Tezpur University
Assam, India-784028



Department of Computer Science & Engineering
Tezpur University

Napaam, Tezpur- 784028, Assam, India.

Dr. Nityananda Sarma
Professor

Phone: 03712-275356
E-Mail : nitya@tezu.ernet.in

Certificate

This is to certify that the thesis entitled "Automatic Detection and Identification of Parasite Eggs in Noisy Microscopic Images of Fecal Samples" submitted to Tezpur University in the Department of Computer Science and Engineering under the School of Engineering in partial fulfillment of the award of the degree of Doctor of Philosophy in Computer Science and Engineering is a record of research work carried out by **Kaushik Ray** under my co-supervision.

All the helps received by him from various sources have been duly acknowledged. No part of this thesis has been submitted else where for award of any other degree.

A handwritten signature in blue ink, followed by the date 20/12/2024.

Signature of Co-Supervisor
(Dr. Nityananda Sarma)
Professor

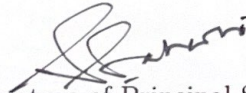
Department of Computer Science and Engineering
Tezpur University
Assam, India-784028



Certificate

This is to certify that the thesis entitled "**Automatic Detection and Identification of Parasite Eggs in Noisy Microscopic Images of Fecal Samples**" submitted by **Mr. Kaushik Ray** to Tezpur University in the Department of Computer Science and Engineering under the School of Engineering in partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy in Computer Science and Engineering has been examined by us on **20-12-2024**..... and found to be satisfactory.

The Committee recommends for award of the degree of Doctor of Philosophy.


Signature of Principal Supervisor

Signature of External Examiner

Acknowledgment

It's a great pleasure to express my gratitude to all of them who have provided me with guidance and support in bringing out the successful completion of my doctoral program at Tezpur University. It would not have been possible for me to succeed with my effort alone. There were a lot of people who extended their supportive hands toward me in the way of making my work a success. I would like to thank everyone who supported and assisted me while carrying out this work at Tezpur University.

First of all, I am grateful to my supervisor Prof. Sarat Saharia for his constant support, trust, valuable feedback, encouragement, and innumerable advice. He gave freedom to pursue my ideas and work at my own pace and was always available to discuss various problems on the way. His encouragement and guidance have provided a sound basis for completing my research work. I enjoyed spending these years with him, both at work and otherwise. Thank you, sir!

I would like to express my sincere gratitude to all the members of my doctoral research committee, particularly Prof. Nityananda Sarma, my co-supervisor, Prof. Bhogeswar Borah, Dr. Swarnajyoti Patra for their valuable help, comments, and suggestions. I genuinely thank Prof. Nityananda Sarma sir for his consistent encouragement and help during my research work, for which I will always be indebted to him.

I am very thankful to Dr. Nagappa S. Karabasanavar from Karnataka Veterinary, Animal & Fisheries Sciences University and Mr. Raafat Salih Hadi from Universiti Malaysia Pahang for providing the essential data for our research. The data was crucial, and I am at a loss for words to express how much it helped me complete my PhD. I will always be grateful to them.

I would also like to acknowledge the support and assistance received from the faculty members of the Department of Computer Science and Engineering from time to time, with special appreciation for the following individuals: Prof.

Utpal Sharma, Prof. Dhruba Kr. Bhattacharya, Dr. Sanjib Kumar Deka, Dr. Rosy Sarmah, Dr. Arindom Karmakar, Dr. Debojit Boro, and Dr. Nabajyoti Medhi.

It is a pleasure to mention the non-teaching staff namely, Mr. Gautam Kumar. Das, Mr. Ajay Kumar. Sharma, Mr. Arun Chetri, and Mr. Arup Kumar Sarkar for their help and support during my PhD work. I thank them for the way they have treated me and guided me as their brother. I am lucky to have them in my life.

I extend my sincere thanks to all other faculty members and the non-teaching staff of the department for their generous help and encouragement in various ways towards completing the work. I sincerely thank the members of my thesis review committee and the anonymous reviewers for their precious comments and feedback.

I am deeply grateful to all my friends, seniors, and juniors from the department and the university, especially Dr. M. Kathing, Meenakshi Ba, Barnali Ba, Prathana, Risang, and other colleagues in the department. Their constant encouragement and support have been invaluable throughout my Ph.D. journey. I would also like to express my heartfelt thanks to my friend Dhriti Khaund for her support and encouragement during this journey.

A very special mention goes to my mother, father, sister, and brother, who have been a constant source of motivation and support in every step of my life. I am truly grateful to have such a wonderful family..

I would like to thank the Almighty for giving me the opportunity and blessings to fulfill my dream.

Finally, I would like to thank those who have directly or indirectly helped me complete my research work in different ways.

Kaushik Ray

List of Figures

1-1	Various steps involved in a manual identification process of parasite eggs.	4
1-2	Integration of automated parasite egg identification process.	5
1-3	Various stages involved in an automatic identification process of parasite eggs.	6
1-4	Some common species of parasite eggs found in fecal or stool samples [1].	7
1-5	A few commonly found parasite eggs in blood samples of animals and humans.	8
1-6	A few commonly found parasite eggs in urine samples of animals and humans [1].	8
1-7	Microscopic images of fecal samples containing different species of parasite eggs. Image credit - "E-Varaha Information System for Safe Pork Production in North-East India" - ITRA, Digital India Corporation-sponsored project and R.S Hadi [2].	9
2-1	Methods used for Parasite Egg Segmentation	18
3-1	A few microscopic images of fecal samples containing parasite eggs of Asacaris (first row) and Neactor (second row), provided by Dr. Nagappa	51
3-2	A few microscopic images of fecal samples containing parasite eggs of Asacaris (First row) and Trichuris (Second row), provided by Mr. R. S. Hadi [2]	52

List of Figures

3-3	Example of blurry images that are discarded from the dataset . . .	52
3-4	A few segmented images of Ascaris (Roundworm) eggs	53
3-5	A few segmented images of Necator (Hookworm) eggs	53
3-6	A few segmented images of Trichuris (Whipworm)	53
3-7	A few segmented images of non-egg objects or debris	53
3-8	Original Image (Row 1) and Corresponding Segmentation Mask (Row 2) indicating Ascaris or Roundworm Eggs in Green and Non- Egg objects in Red	55
3-9	Original Image (Row 1) and Corresponding Segmentation Mask (Row 2) indicating Necator or Hookworm Eggs in Green and Non- Egg objects in Red	56
3-10	Original Image (Row 1) and Corresponding Segmentation Mask (Row 2) indicating Trichuris or Whipworm Eggs in Green and Non- Egg objects in Red	56
3-11	Original Image (Row 1) and Corresponding Binary Segmentation Mask (Row 2) indicating parasite eggs and other non-egg objects in white	56
3-12	Example of annotating different classes of objects using bounding boxes in LabelImg	57
4-1	Examples of input and thresholded output images: (a) Image con- taining parasite eggs with a little load of debris, (b) with moderate load of debris, and (c) with high load of debris	66
4-2	Segmentation result of various edge detection methods	69
4-3	Output Images from the Distance Transform-based Watershed Seg- mentation Method	70
4-4	A few examples of output images using K-means segmentation method	70
4-5	Proposed Segmentation Approach	71

List of Figures

4-6	Input and output images of different stages of the proposed segmentation approach	74
4-7	Comparison of original ground truth masks and edge-detected images.	75
5-1	Process of Extracting Pixel Intensity-based Features. Here, f1, f2, f3, f4, and f5 represent mean, maximum, minimum pixel intensity, variance, and standard deviation, respectively.	92
6-1	Change of classification accuracy with K-value for KNN with Hu's moments	95
6-2	Test results of ANN using Hu's seven moments	96
6-3	Test results of XGBoost using Hu's seven moments	96
6-4	Change of classification accuracy with moment orders for different classifiers	100
6-5	Test result of SVM using 10 th order Chebyshev moment	101
6-6	Test result of XGBoost using 15 th order Chebyshev moments	101
6-7	Change of classification accuracy with the number of PCA components for SVM with 50 th order Chebyshev Moments	102
6-8	Performance analysis of different classifiers using three types of image moments with PCA and LDA	105
6-9	Test result of SVM using Texture and Shape-based features	106
6-10	Test result of XGBoost using Texture and Shape-based features . . .	106
6-11	Change of Classification Accuracy with Size of the Blocks or Dimension of Pixel-Intensity-based Feature set	107
6-12	Test result of ANN using pixel intensity-based features	108
6-13	Test result of SVM using pixel intensity-based features	109
6-14	Analysis of test accuracy of different classifiers with individual vs combined feature sets	111

List of Figures

6-15	Analysis of Silhouette score for finding optimum number of clusters using KMeans	112
6-16	Analysis of Silhouette score for finding optimum number of clusters using Gaussian Mixture Model	112
6-17	Non-Egg objects in cluster-1 with 3 optimal clusters	113
6-18	Non-Egg objects in cluster-2 with 3 optimal clusters	113
6-19	Non-Egg objects in cluster-3 with 3 optimal clusters	113
6-20	Analysis of classification accuracy of different classifiers with different features while using two classes of non-eggs and three parasite eggs	115
6-21	Confusion matrix obtained from testing SVM and XGBoost models using five classes of objects	115
6-22	Classification report obtained from testing SVM and XGBoost models using five classes of objects	115
6-23	Confusion matrix obtained from testing Random Forest and SVM using six classes of objects	117
6-24	Classification report obtained from testing Random Forest and SVM using six classes of objects	117
6-25	Analysis of classification accuracy of different classifiers using single and multiple non-egg classes of objects	117
6-26	Analysis of classification accuracy of different classifiers using a single non-egg object and three parasite eggs	118
6-27	A comparison of the classification accuracy of different parasite egg classes using single and multiple non-egg classes	119
7-1	Training and validation loss of UNet using transfer learning with ResNet50 for binary segmentation.	126
7-2	Validation result of Unet using transfer learning: input image (left), ground truth (middle), and output segmented image (right)	127

7-3	Validation results of UNet training from scratch using random weights: image (left), ground truth (middle), and output segmented image (right)	128
7-4	UNet multi-class segmentation output: input image (left), ground truth (middle), and output segmented image (right)	129
7-5	Modification of the VGG16 Model Used for Training Parasite Eggs Using the Transfer Learning Approach	130
7-6	Change of Training and Validation Accuracy of Transfer Learning Approach - four classes	131
7-7	Confusion Matrix from validation of CNN classification model using transfer learning—four classes	132
7-8	A summarised view of the custom CNN model designed for the classification of parasite eggs	132
7-9	Change of Training and Validation Accuracy of Custom CNN Model - Four classes	133
7-10	Confusion Matrix from validation of own CNN classification model - Four classes	133
7-11	Change of Training and Validation Accuracy of Custom CNN model - Six classes	134
7-12	Confusion Matrix of the Validation Process of Custom CNN Model - Six classes	135
7-13	Comparison of classification accuracy of parasite eggs using CNN models with single and multiple Non-egg classes	135
7-14	Comparison of classification accuracy of ML-based and CNN-based models - four classes	136
7-15	Comparison of classification accuracy of ML-based and CNN-based models - six classes	136
7-16	Working principle of R-CNN [3]	138

List of Figures

7-17 Visualization of different Losses from the training of Faster-RCNN model	139
7-18 Examples of detecting various objects in images using the object detection model.	140

List of Tables

2.1	Parasite egg segmentation works that use Thresholding based method	20
2.2	Parasite egg segmentation works that used Edge detection-based methods	25
2.3	Parasite egg segmentation works that used ACM-based methods . .	29
2.4	Parasite egg detection works that used Watershed-based methods .	31
2.5	Works that used Ellipse matching-based methods	34
2.6	Parasite egg segmentation/detection works that applied Convolutional Neural Network-based methods	37
2.7	Various types of Features and Classification algorithms used in Automatic Identification of Parasite Eggs	42
2.8	Parasite egg Classification and Detection works that applied Convolutional Neural Network-based Approaches	46
3.1	Dataset containing microscopic images of different types of parasite egg	52
3.2	Number of Segmented Objects	54
3.3	Image dataset containing different types of parasite eggs for CNN-based segmentation	55
4.1	Advantages and disadvantages of some well-known image segmentation methods	61

6.1	Different classification algorithms and hyper-parameter settings for classifying parasite egg images	94
6.2	Classification Results Obtained by Various Classifiers Using Hu's Seven Moments	96
6.3	Classification Results Obtained by ANN using Various types of Image Moments	97
6.4	Classification Results Obtained by SVM Using Various Types of Image Moments	97
6.5	Classification Results Obtained by kNN using Various types of Image Moments	98
6.6	Classification Results Obtained by Decision Tree Using Various Types of Image Moments	98
6.7	Classification Results Obtained by Random Forest Using Various Types of Image Moments	98
6.8	Classification Results Obtained by XGBoost Using Various Types of Image Moments	99
6.9	Highest classification accuracy obtained by different classifiers using different moment orders	99
6.10	Highest training accuracy yielded by different classifiers using PCA on different types of image moments	103
6.11	Highest test accuracy of different classifiers using moment-based features and PCA analysis	103
6.12	Highest training accuracy yielded by different classifiers using LDA on different types of image moments	104
6.13	Highest test accuracy of different classifiers using moment-based features and LDA analysis	104
6.14	Overall classification accuracy of various classifiers using texture and shape-based features	105

List of Tables

6.15	Classification accuracy (test) for each class using texture and shape-based features	105
6.16	Training accuracy of different classifiers using Pixel intensity-based features for Four Classes	107
6.17	Test accuracy of different classifiers for each class using pixel intensity-based features	108
6.18	Classification accuracy for different classifiers using PCA analysis on pixel intensity-based features	109
6.19	Classification accuracy (Test) for each class using pixel intensity-based features using PCA	109
6.20	Training classification accuracy using combinations of different feature sets	110
6.21	Highest classification results of the classifiers with combination of different features	110
6.22	Number of segmented images in each class for classification task using multiple non-egg classes	113
6.23	Training results of various classifiers using Two classes of Non-eggs and Three parasite eggs	114
6.24	Test accuracy for each class using texture and shape-based features while trained for five classes of objects	114
6.25	Training results of various classifiers while using three classes of Non-eggs and three parasite eggs	116
6.26	Test accuracy for each class using texture and shape-based features while trained for six classes of objects	116
6.27	Classification accuracy of different classes of parasite eggs with single and multiple non-egg classes	118
6.28	Comparison of results with similar works in the past	119
7.1	Number of images used in UNet-based segmentation process	125

List of Tables

7.2	Hyper-parameters used during the training of UNet using the transfer learning technique	126
7.3	UNet segmentation results as binary classification with different pre-trained weights	127
7.4	UNet segmentation results as binary classification with random weights	127
7.5	Hyper-parameters used during the training of UNet using transfer learning technique with multi-class objects	129
7.6	UNet segmentation results as multi-classification with random weights	129
7.7	Values of Different Parameters Used in Training Using Transfer Learning Approach	131
7.8	Validation result of CNN model using transfer learning - four classes	132
7.9	Validation result using a custom CNN model - Four classes	133
7.10	Validation result using a custom CNN model - Six classes	134
7.11	Comparison of results with the works that used CNN for classification of parasite eggs	137
7.12	Validation results of Faster-RCNN objects detection model	139
7.13	Comparison of our result with the previous works that used CNN-based object detection models for the detection of parasite eggs. . .	140