Detection and Classification of Caterpillar using Image Processing with K-Nearest Neighbor Classification Technique

[1]Dr. Mukta Jagdish, [2]Andres Medina Guzman, [3]Gerber F. Incacari Sancho, [4]Aura Guerrero-Luzuriaga

[1]Lecturer, Jawaharlal Nehru Rajkeeya Mahavidyalaya, Government College, Port Blair, affiliated to Pondicherry Central University, South Andaman, Andaman and Nicobar Islands, India
[2]Universidad de la Costa, Barranquilla, Colombia
[3]Universidad Nacional del Callao, Lima, Perú
[4]Universidad Católica de Cuenca, Ecuador
[1]mukta.jagdish13@gmail.com, [2]amedina10@cuc.edu.co, [3]gfincacaris@unac.edu.pe, [4]ing.auragl@gmail.com

Abstract: Caterpillars are the developmental stage of the flying insect called butterfly. The moths are the beautiful creature of earth which comes under the class of insects. They are recognized by their beautiful and colorful forewings body and legs. Caterpillars are the larval stage of the moth which are found in the leaf and stem of the plants when the moth laid eggs on the leaves after their mating. Caterpillar after fully developed from its eggs draw out a flimsy, soft cocoon made up of dark coarse silk on leaves and stem for their shelter. Caterpillar is also a beautiful creature that is found with different colors and strips with spines and urticating hair in their body for releasing venom for self-defense from external predators. The present study works on the detection and classification of the caterpillar using image processing with a k-NN classifier.This research help in characterizing the type of caterpillar image classification for particular three classes such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and the accuracy of Io moth Caterpillar. The following stages have been considered are preprocessing, segmentation, feature extraction, and classification methods using K- Nearest Neighbor classifier. The present investigation results that SYMLET5 analysis works well in the classification of the caterpillar with an accuracy of 96% using K- Nearest Neighbor classifier compare with other measures during investigation and analysis.

Keywords: K-NN classifier, caterpillar, SYMLET analysis, machine learning, image processing.

1.   Introduction

Caterpillars are the larval stage of members of the order of butterflies. Caterpillars are herbivorous, but some are insectivorous about 1%. Caterpillars are having soft bodies that can grow rapidly between molts [1]. Their size varies between species from 1 mm up to 14 cm. Caterpillars can defend against physical conditions such as cold, hot, or dry environmental conditions. Appearances of caterpillars can often threaten a predator. Its markings on certain body parts can make it look poisonous and bigger thus helps in threatening foreign particles [2]-[3]. Some caterpillars are poisonous or tasteless because of their bright coloring & aposematic, while others may mimic dangerous caterpillars. Many caterpillars are cryptically colored to protect themselves from predators and resemble the plants on which they feed [4]. Some caterpillars use camouflage for defense in the species. If the caterpillars hatch in the spring they appear green and if they hatch in the summer they appear dark-colored. The difference in color is due to the tannin content in the diet. Some caterpillars have spines or growths that resemble plant parts such as thorns and few look like objects in the environment such as bird droppings [5]. Some species cover themselves in plant parts, while some construct and live in a bag covered in sand, pebbles, or plant material. Some caterpillars have spiny bristles or long fine hairs that will irritate by lodging in the skin or mucous membranes whereas some caterpillars acquire toxins from their host plants [6]. Most of the caterpillar defenses from urticating hairs are associated with venom glands. The venom is powerful enough to cause a human from hemorrhage to death. Caterpillars suffer predation from European pied flycatcher& Paper wasps [7]. Many caterpillars feed in protected environments, such as enclosed in silk galleries, rolled leaves, or by mining between the surfaces of leaves. Caterpillars are also called eating machines as they eat leaves voraciously [8]. They grow very quickly and their skin sheds 4 to 5 times as they grow and their weight also increases very fast. Here we specifically discuss some of the specified caterpillars like the Buck Moth Caterpillar, Saddleback Caterpillar, and Io moth caterpillar [9]-[10].

1.1 Buck Moth Caterpillar: These insects are the genus of Hemileuca from the family group of Saturniidae. The buck moth caterpillars are mostly found in oak forests of the United States from peninsular Florida to New England and west of Texas and Kansas. It is also called Dru Drury. These caterpillar larvae generate in spring which is covered with a hollow spine that is attached to a poisonous sac. These poisonous sacs of larvae cause itching, burning sensation to nausea. These larvae are feed from live oak, white oak, blackjack oak, scrub oak, and many more. Caterpillar laid their eggs in a spiral cluster on oak twigs. Mature larvae went to the litter of the

leaves to pass through their pupal stage at the end of July. The mating month of these insects is February. These moths are infamous for stinging people.

1.2 Saddleback Caterpillar: These are the larvae from the species of moth from the genus of Acharia and the family group of Limacodidae. These are native to Eastern North America such as Mexico, Yucatan. These caterpillars are leaf green color with brown color & fleshy horn at both ends and a prominent white-ringed brown dot in the center looks like a saddle for riding. The other parts of the body are covered with urticating hairs that secrete irritating venom if touched which protects it from external predators. This venom causes a swollen and painful rash on the skin and even nausea. In some cases, it leads to migraine, asthma, gastrointestinal symptoms, and sometimes leads to hemorrhaging. The larvae depend on a variety of plants like ornamental palm like Manila palm for food and nutrient for growth. The mating month of their moth is January to February. The moth laid eggs after 3 days of mating on the underside of the leaf in a series of 30-50 eggs. The size of the eggs is 1- 2 mm. Throughout the life cycle, the larvae of this caterpillar go through different stages. They are the first instars, the middle instars, the late instars. The adult comes out in June and July.

1.3 Io moth caterpillar: Io moth caterpillars are the genus of Automeris from the family group of Saturniidae. These insects are mostly found in US-North Dakota, South Dakota, New Mexico, Montana, and Southern Florida. Some species are also found in Southern Manitoba, Ontario, Quebec of Canada. The moths of these mature caterpillars are famous for their beauty. The adult moths have a 4 wings span of 2.5 – 3.5 inches. Males have yellow forewings, body, and legs whereas females have reddish forewings, body, and legs. They have a feathery antenna in front of their head and have a bluish eyespot with some white mark on its center on each wing. This provides a defense mechanism to the moth from external predators. These moths laid their eggs in a cluster of 20 on the leaves. As the eggs develop it turns black and when the orange larvae emerge from the egg after hatching it eaten up its eggshell. In the advanced stage, this larvae turns from orange to green with urticating hairs & spine all over the body for its protection. The larvae develop in 5 instars. The spine on the body of the caterpillar release very painful venom which rashes and swollen your skin. It also leads to nausea, migraine, and gastrointestinal problems. When the caterpillars are ready, they draw out a flimsy, soft cocoon made up of dark coarse silk in leaves. They depend on plant leaves for their food and nutrients. They have eaten up plant leaf voraciously during their growth period. The adult moths of this species are strictly nocturnal.

Figure 1- Buck Moth Caterpillar, Saddleback Caterpillar, and Io moth Caterpillar

2. Methodology

These researchesperformclassification methods using image processing with machine learning techniques using data of 100caterpillar images for analysis. The research method is categorized into different stages such as

preprocessing, segmentation, feature extraction, and K-Nearest Neighbor classifier. Data collected for research from multiple sources during the investigation.In this research, we discuss the classification of different types of caterpillars. Now a day's various Image Processing Techniques are used for a detailed study of different types of species and creatures. In this research, it detects and classifies a specific category of the caterpillar using Image Processing with the K-Nearest Neighbor Classification Technique.
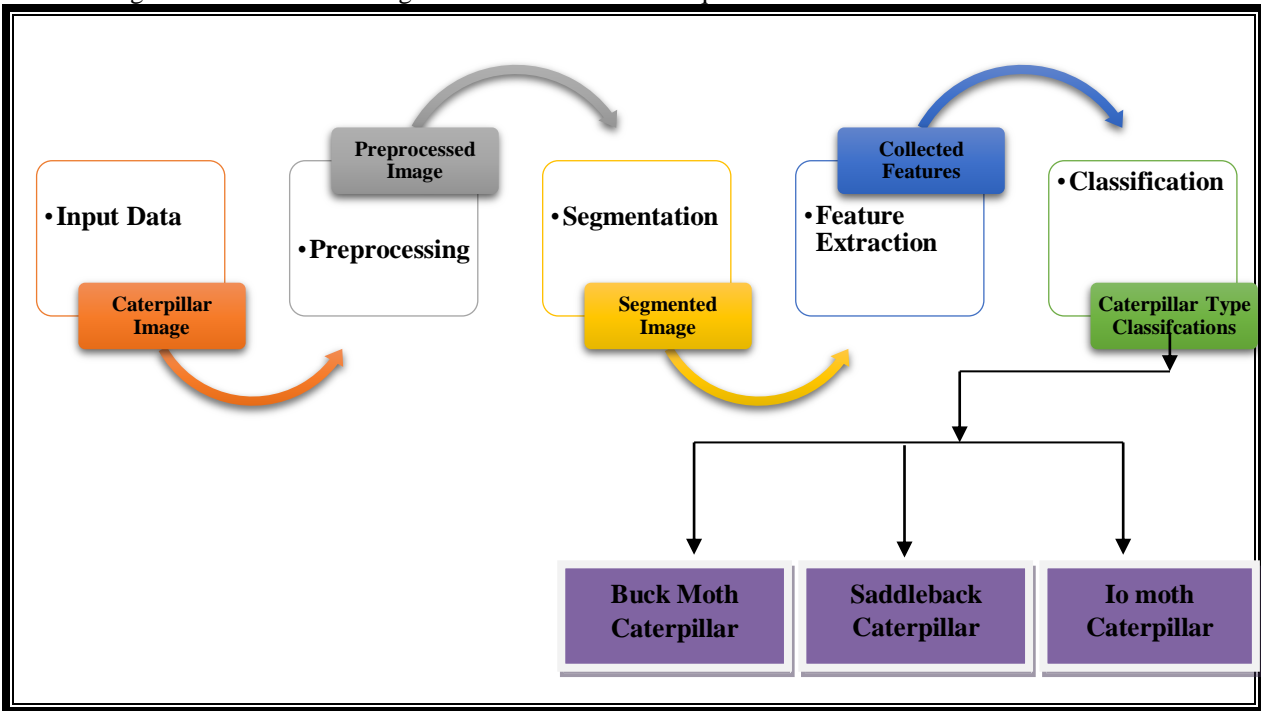

Figure 2- Methodology

3. Preprocessing

The acquisition of a caterpillar image is a very challenging task for classifying the current situations. Based on different ways of taking pictures, the image may get affected and become more noises, so it is very important to remove noise from background images during the research process and to extract detailed features from images. For this research image preprocessing play an important role in the removal of unwanted noise from the background of images to achieves better results. Preprocessing makes a noise-free image using much digital processing. In this research work, the median filter is used during preprocessing techniques.

4. Segmentation

In this research segmentation methods are used to extract region-based image segmentation for caterpillar classification. Image segmentation process the image into multiple segments or a group of pixels with relevant criteria based on some conditions or protocols. Segmentation is used to make an image with more meaningful representation. During the complete process, it will divide the image into multiple segments part for better understanding and analysis. The result of the segmentation process will collectively cover the entire caterpillar image.In the segmentation, method thresholding is used to extract caterpillar images that are completely based on threshold values.
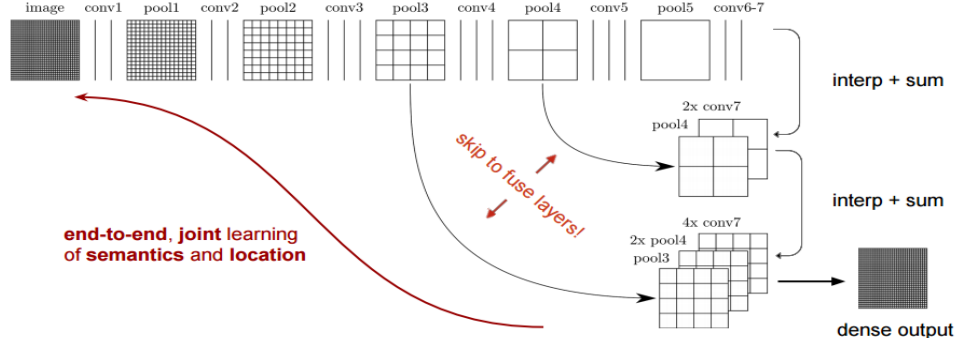

Figure 3- Segmentation

5. Feature Extraction

Features extracted for caterpillar classification are color-based analysis (RGB Colors), mean, standard deviation, ellipticity, entropy, skewness, intensity, and correlation co-efficient with wavelet such as symplet1, symlet2, symlet3, symlet4, symlet5 analysis.For SYMLET analysis it uses high computational overhead.
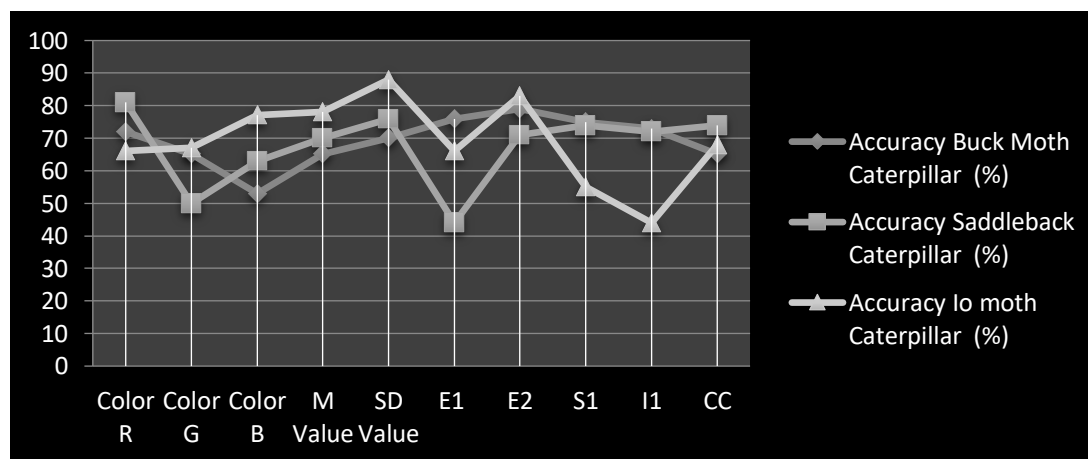
6. Classification

For this research k- Nearest Neighbor classifier method is investigated with three classes such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and the accuracy of Io moth Caterpillar. It is a classification method which makes prediction target variable class according to nearest neighbors.It calculates the distance which needs to classify with instance training data based on the majority of the vote given by the nearest neighbor pixel. The three-parameter is used they are k value, train set, and test instance during analysis. The dataset used is 100 caterpillar images with three classes such asBuck Moth Caterpillar, Saddleback Caterpillar, and Io moth Caterpillarfor classifications. This investigation contains 70 % for training data and 30% for testing data with k=4.

7. Result and Discussion

In this investigation, an advance study of caterpillar classification using image processing and machine learning methods have been considered for analysis. A 100 caterpillar image dataset has been taken for analysis and investigation. The method is broadly categories into the following stages such as preprocessing phase, segmentation phase, feature extraction such as color-based analysis (RGB Colors), mean, standard deviation, ellipticity, entropy, skewness, intensity, and correlation co-efficient with wavelets such as symlet1, symlet2, symlet3, symlet4, symlet5 analysis, and K-Nearest Neighbor classifier with 70% used for training and 30% used for testing.

Table 1:KNN classifier with symlet1 analysis for Caterpillar Image Classifications

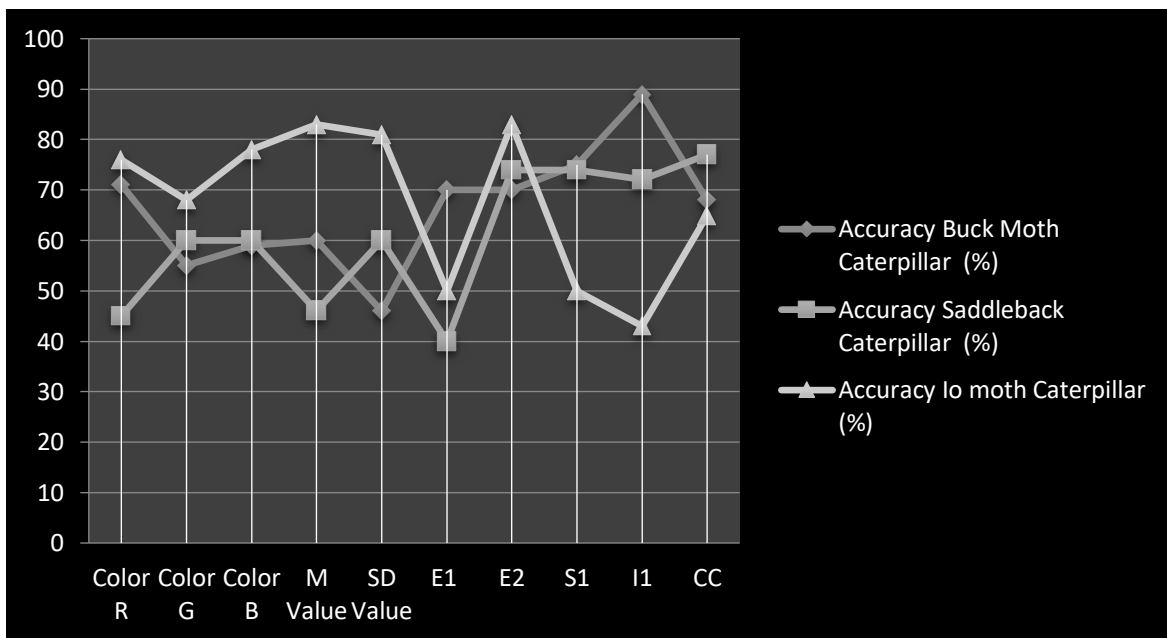| Caterpillar Image classification Using KNN Classifier									
symlet1Analysis									
Features	Color R$_1$	Color G$_1$	Color B$_1$	M Value	SD Value	E$_1$	E$_2$	S$_1$	I$_1$	CC
Accuracy Buck Moth Caterpillar (%)	72	65	53	65	70	76	79	75	73	65
Accuracy Saddleback Caterpillar (%)	81	50	63	70	76	44	71	74	72	74
Accuracy Io moth Caterpillar (%)	66	67	77	78	88	66	83	55	44	68

Figure 4- KNN classifier with Symlet1 analysis

Table 1 with figure 4 denotes the detailed result of a classification using K-NN classifier and Symlet1 analysis for a Caterpillar image classification such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and accuracy of Io moth Caterpillar.

Table 2:KNN classifier with symlet2 analysis for Caterpillar Image Classifications

Caterpillar Image classification Using KNN Classifier									
symlet2Analysis									
Features	Color R_1	Color G_1	Color B_1	M Value	SD Value	E_1	E_2	S_1	I_1	CC
Accuracy Buck Moth Caterpillar (%)	71	55	59	60	46	70	70	75	89	68
Accuracy Saddleback Caterpillar (%)	45	60	60	46	60	40	74	74	72	77
Accuracy Io moth Caterpillar (%)	76	68	78	83	81	50	83	50	43	65

Table 2 represents caterpillar classification using a k-NN classifier with symlet2 analysis with three classes such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and accuracy of Io moth Caterpillar. Features extracted are color-based analysis (RGB Colors), mean (M_Value), standard deviation (SD_Value), ellipticity (E1), entropy (E2), skewness (S1), intensity (I1), and correlation co-efficient (CC).
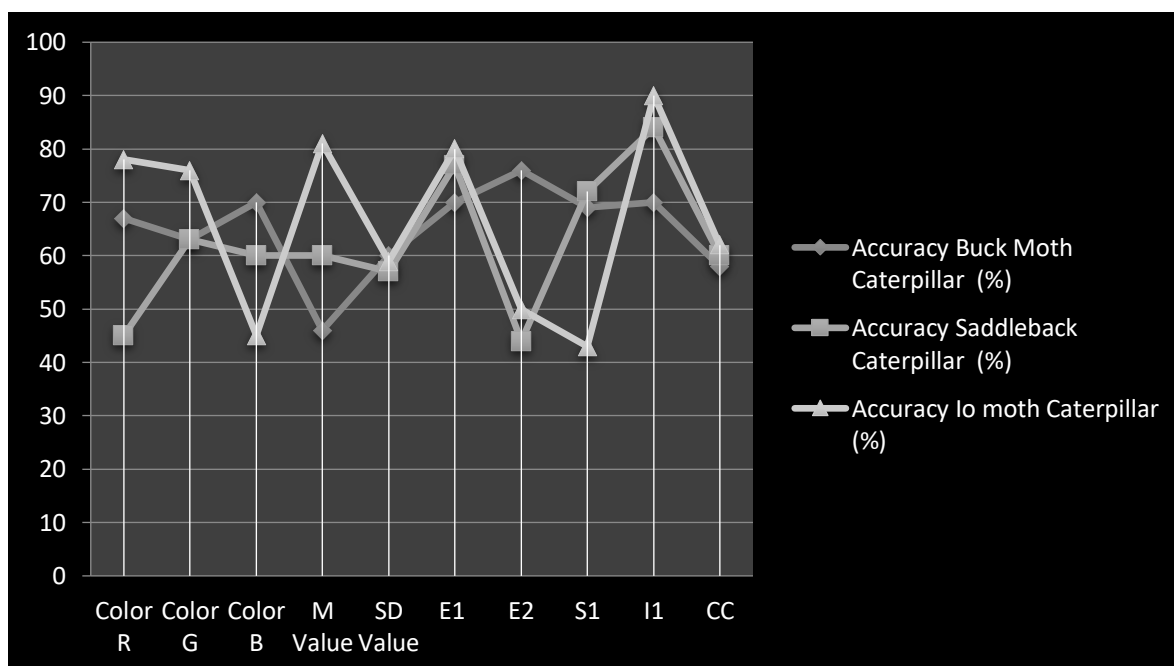

Figure 5- KNN classifier with Symlet2analysis

Figure 5 denotes the graphical result of a classification using K-NN classifier and Symlet2 analysis for a Caterpillar image classification such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and accuracy of Io moth Caterpillar.

Table 3:KNN classifier with symlet3 analysis for Caterpillar Image Classifications

Caterpillar Image classification Using KNN Classifier										
symlet3Analysis										
Features	Color R_1	Color G_1	Color B_1	M Value	SD Value	E_1	E_2	S_1	I_1	CC
Accuracy Buck Moth Caterpillar (%)	67	63	70	46	60	70	76	69	70	58
Accuracy Saddleback Caterpillar (%)	45	63	60	60	57	77	44	72	84	60
Accuracy Io moth Caterpillar (%)	78	76	45	81	59	80	50	43	90	62

Table 3 represents caterpillar classification using a k-NN classifier with symlet3 analysis with three classes such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and accuracy of Io moth Caterpillar. Features extracted are color-based analysis (RGB Colors), mean (M_Value), standard deviation (SD_Value), ellipticity (E1), entropy (E2), skewness (S1), intensity (I1), and correlation co-coefficient (CC).
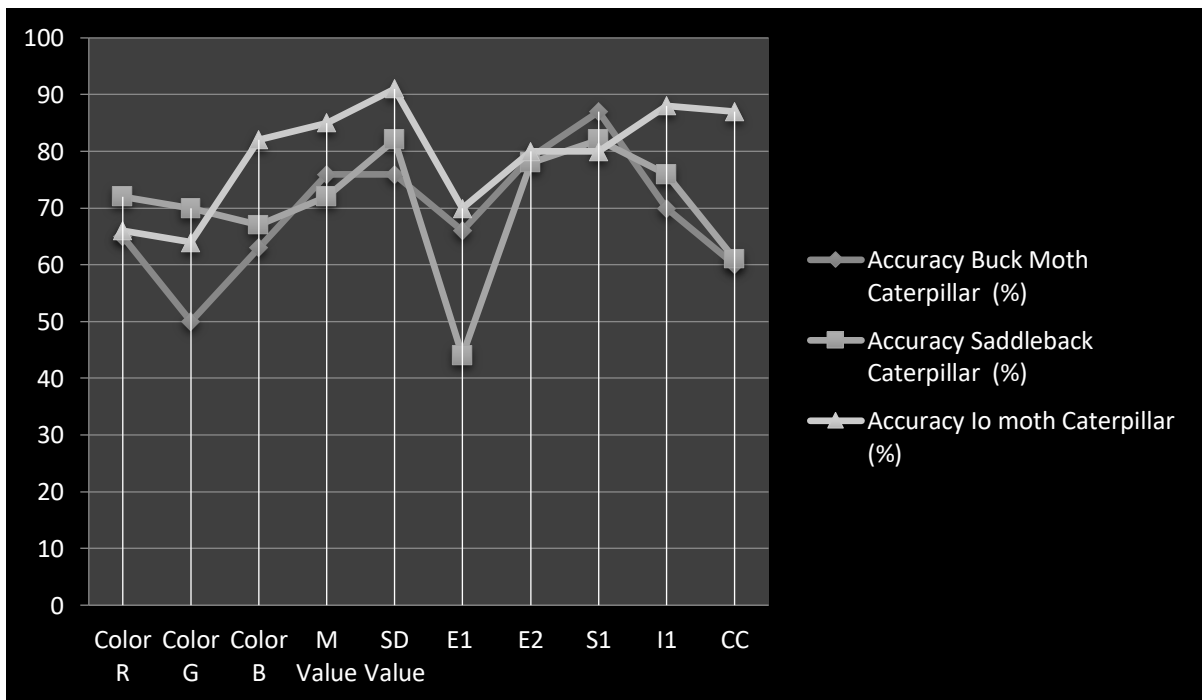

Figure 6- KNN classifier with Symlet3analysis

Figure 6 denotes the graphical result of a classification using K-NN classifier and Symlet3 analysis for a Caterpillar image classification such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and accuracy of Io moth Caterpillar.

Table 4:KNN classifier with symlet4 analysis for Caterpillar Image Classifications

Caterpillar Image classification Using KNN Classifier									
symlet4Analysis									
Features	Color R_1	Color G_1	Color B_1	M Value	SD Value	E_1	E_2	S_1	I_1	CC
Accuracy Buck Moth Caterpillar (%)	65	50	63	76	76	66	79	87	70	60
Accuracy Saddleback Caterpillar (%)	72	70	67	72	82	44	78	82	76	61
Accuracy Io moth Caterpillar (%)	66	64	82	85	91	70	80	80	88	87

Table 4 represents caterpillar classification using a k-NN classifier with symlet4 analysis with three classes such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and accuracy of Io moth Caterpillar. Features extracted are color-based analysis (RGB Colors), mean (M_Value), standard deviation (SD_Value), ellipticity (E1), entropy (E2), skewness (S1), intensity (I1), and correlation co-efficient (CC).
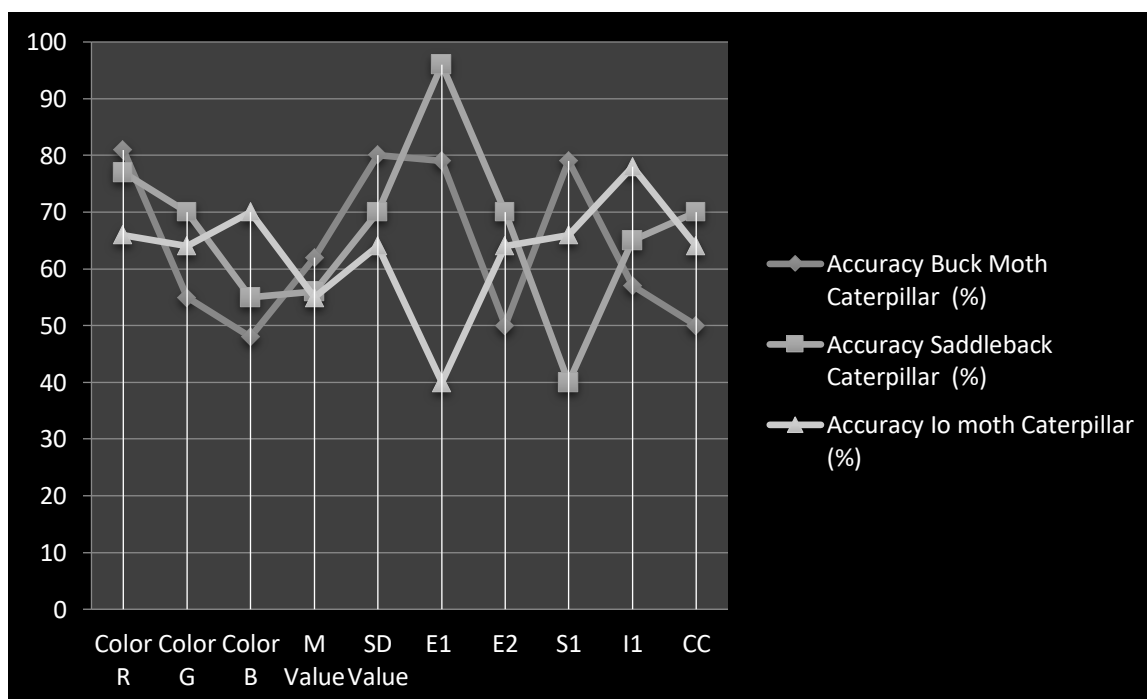

Figure 7- KNN classifier with Symlet4 analysis

Figure 7 denotes the graphical result of a classification using K-NN classifier and Symlet4 analysis for a Caterpillar image classification such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and accuracy of Io moth Caterpillar.

Table 5: KNN classifier with symlet5 analysis for Caterpillar Image Classifications

Caterpillar Image classification Using KNN Classifier									
symlet5Analysis									
Features	Color R₁	Color G₁	Color B₁	M Value	SD Value	E₁	E₂	S₁	I₁	CC
Accuracy Buck Moth Caterpillar (%)	81	55	48	62	80	79	50	79	57	50
Accuracy Saddleback Caterpillar (%)	77	70	55	56	70	96	70	40	65	70
Accuracy Io moth Caterpillar (%)	66	64	70	55	64	40	64	66	78	64

Table 5 represents caterpillar classification using a k-NN classifier with symlet5 analysis with three classes such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and accuracy of Io moth Caterpillar. Features extracted are color-based analysis (RGB Colors), mean (M_Value), standard deviation (SD_Value), ellipticity (E1), entropy (E2), skewness (S1), intensity (I1), and correlation co-efficient (CC).
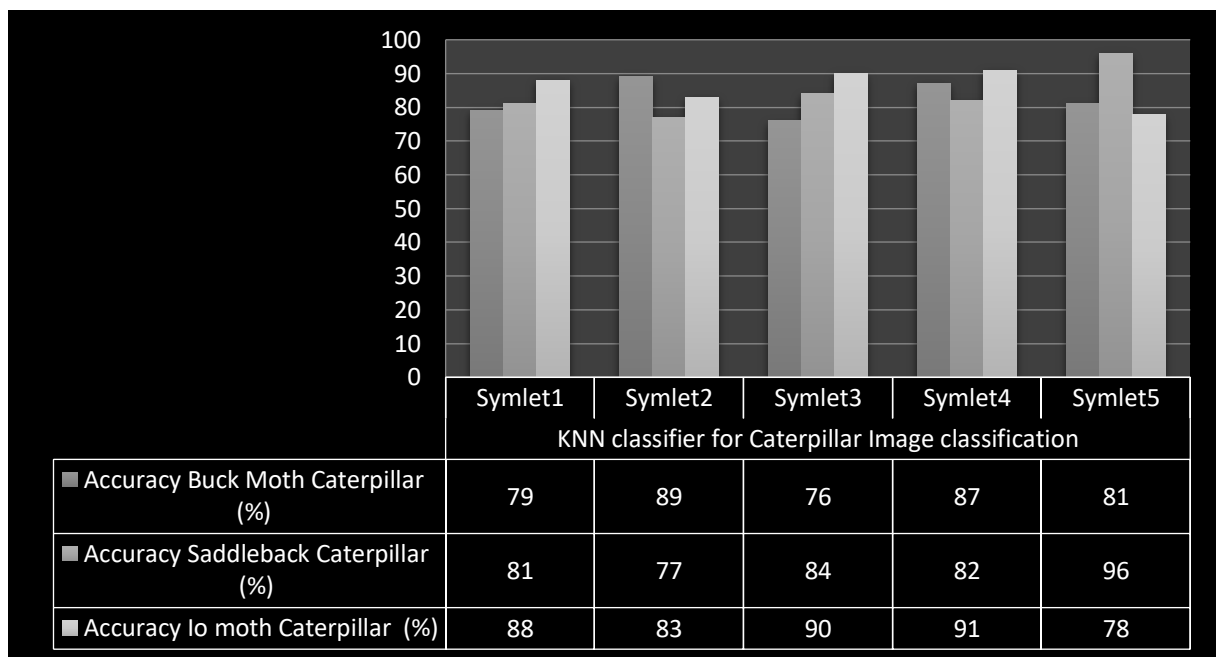

Figure 8- KNN classifier with Symlet5analysis

Figure 8 denotes the detailed result of a classification using K-NN classifier and Symlet5 analysis for a Caterpillar image classification such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and accuracy of Io moth Caterpillar.

Table 6:KNN classifier with Combine SYMLET analysis for Caterpillar Classifications

	Symlet1	Symlet2	Symlet3	Symlet4	Symlet5
Accuracy Buck Moth Caterpillar (%)	79	89	76	87	81
Accuracy Saddleback Caterpillar (%)	81	77	84	82	96
Accuracy Io moth Caterpillar (%)	88	83	90	91	78

Figure 9- Combine SYMLET analysis for Caterpillar Classifications

Table 6 represents caterpillar classification using a k-NN classifier with SYMLET analysis with three classes such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and accuracy of Io moth Caterpillar.Features extracted are color-based analysis (RGB Colors), mean (M_Value), standard deviation (SD_Value), ellipticity (E1), entropy (E2), skewness (S1), intensity (I1), and correlation co-efficient (CC). This investigation contains 70% for training data and 30% for testing data with k=4. The present investigation results that symlet5 analysis works well in the classification of caterpillar categories classification with an accuracy of 96% using the K-Nearest Neighbor classifier.

8. Conclusion

In this investigation, a caterpillar classification method and detection system using image processing and machine learning method are performed. This research help in characterizing the type of caterpillar image classification for particular three classes such as accuracy of Buck Moth Caterpillar, the accuracy of Saddleback Caterpillar, and the accuracy of Io moth Caterpillar. The following stages have been considered are preprocessing, segmentation, feature extraction, and classification methods using K- Nearest Neighbor classifier. The present investigation results that SYMLET5 analysis works well in the classification of the caterpillar with an accuracy of 96% using K- Nearest Neighbor classifier compare with other measures during investigation and analysis.

References

1. J. H. Belanger, &B. A. Trimmer, (2000). Combined-kinematic and electromyographic analyses-of-proleg function-during-crawling by the caterpillar-Manduca-sexta. Journa. of Comparative-Physiology. A Sensor., Neura., and Behavioral*Physiology186, 1031(1039).

2. J. H. Belanger, K. J. Bender, &B. A.Trimmer, (2000). Context-dependency of a limb withdrawal-reflex in the caterpillar-Manduca-sexta. Journa. of Comparative-Physiology. A Sensor., Neura., and Behavioral*Physiology 186, 1041(1048).

3. R. A. Bell, & F. A. Joachim, (1978). Techniques-for-rearing laboratory-colonies of the tobacco-hornworm, Manduc. sexta and pink-ballworms. Annals of the Entomological-Society of *America 69, 365(373).

4. J. A. Bender, A. J. Pollack, & R. E. Ritzmann, (2010). Neural-activity in the central-complex of the insect-brain is linked-to locomotor-changes. Current*Biology 20, 921(926).

5. E. A. Bernays, & H. A. Woods, (2000). Foraging-in-nature by larvae of Manduca-sexta-influenced by an endogenous-oscillation. Journa. of Insec. Physiolog.46, 825(836).

6. J. Brackenbury, (1996). Novel-locomotory-mechanisms in caterpillars*lifelineclimbing in pinotia-abbreviana (Tortricidae) and Yponomeuta-padella (Yponomeutidae).Physiological*Entomology 21, 7(14).

7. J. Brackenbury, (1997). Caterpillar-kinematics. Nature 390, 453-Brackenbury, Jo. (1999). Fast-locomotion in caterpillars. Journa. of Insect*Physiology 45, 525(533).

8. M. Calisti, M. Giorelli, G. Levy, B. Mazzolai, B. Hochner, C. Laschi,& P. Dario, (2011). An octopus*bioinspired solution-to-movement &manipulation for soft-robots. Bioinspiration*Biomimetics 6, (036002).

9. D. Carter, &M. Locke, (1993). Why-caterpillars do not-grow-short andfat. Internationa. Journa. of Insect*Morphology-Embryology 22, 81(102).

10. C. Consoulas, R. M. Johnston,H. J. Pfl¨uger,&R. B. Levine, (1999). Peripheral-distribution of presynaptic-sites of abdominal-motor-modulatory neurons inManduca sexta-larvae. Journa. of Comparative*Neurology 410, 4(19).