Journal of Agricultural Sciences (Tarim Bilimleri Dergisi) 2024, 30 (4) : 685 – 697 DOI: 10.15832/ankutbd.1436890

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

Machine Learning-Based Grasshopper Species Classification using Neutrosophic Completed Local Binary Pattern

Nuh Alpaslan[a], Mustafa Ilcin[b*]

[a]*Bingol University, School of Engineering and Architecture, Department of Computer Science, Bingol, TURKEY*

[b]*Bingol University, Faculty of Agriculture, Department of Plant Protection, Bingol, TURKEY*

ARTICLE INFO
Research Article
Corresponding Author: Mustafa Ilcin, E-mail: milcin@bingol.edu.tr
Received: 14 February 2024 / Revised: 24 April 2024 / Accepted: 30 April 2024 / Online: 22 October 2024

Cite this article
Alpaslan N, Ilcin M (2024). Machine Learning-Based Grasshopper Species Classification using Neutrosophic Completed Local Binary Pattern. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 30(4):685-697. DOI: 10.15832/ankutbd.1436890

ABSTRACT
Locusts are seen as a major threat to the ecosystem because they devastate crops and contribute to thousands of tons food lost every year. Numerous well-trained agents are needed for the efficient control of these insects. However, this is a challenging process. Grasshopper detection methods are being developed using traditional forecasting methods by expert entomologists. The maximum potential of these methods has not yet been completely realized. Hence the majority of work is still done manually. In this paper, a neutrosophic CLBP (completed local binary pattern) based grasshopper species classification framework is proposed. Our proposed system comprises a novel grasshopper species database of over 7.392 images for grasshopper species classification. The grasshopper image is first converted to a neutrosophic field. These discriminatory features are merged with rotation invariant LBP. Our proposed system could achieve up to 99.7% classification accuracy even while working with challenging datasets of wide image quality and size range. The proposed methodology involved diagnosing 11 species and subspecies. It demonstrates the impracticability of conventional diagnostic techniques in the later stages. It could have a big impact on data analysis, enabling more effective handling of global pest.

Keywords: Neutrosophic set, Local binary pattern, Locust, Deep learning, Machine learning

1. Introduction

Insects are important in the ecological and economic life of humans and other living things. For example, they may cause direct product losses by feeding on agricultural and forest products (Gullan & Cranston 2014). Locust infestations have deleterious consequences on the food accessibility and socio-economic fabric of rural communities. The financial outlay for the eradication of these insects may extend into the millions and requires a certain amount of expert training. Additionally, the extensive use of chemical insecticides might have detrimental environmental repercussions (Zhang et al. 2019). Therefore, precise automated locust detection systems hold the potential to attenuate the reliance on a multitude of pesticides in the context of locust management. Grasshoppers are a pest that is wreaking havoc all over the planet. They could starve millions of people to death and destroy crops worth billions of dollars every year. Depending on factors like temperature and vegetation, they could rapidly multiply twenty times (FAO 2020). However, the amount of data to be estimated by traditional approaches is limited and unreliable. To combat locusts effectively, experts are moving away from a reliance on heavy insecticide use and instead integrating quantitative metrics with on-site observations from specialists. This data fusion approach is crucial for the development of precise, automated locust control systems that can target specific types, developmental stages, and crop vulnerabilities.

A paucity exists in the realm of non-invasive and cost-effective methodologies for the detection and classification of insect diversity, rendering the current state devoid of an efficacious approach. Conventional methodologies form the basis of comparative morphological and insect taxonomy studies. In taxonomy, molecular and molecular cytogenetic research techniques have recently become very prominent (Sreedevi et al. 2015). Classical taxonomy offers the most practical and accurate classification based on general commonalities and the most obvious characteristics of species. It is crucial for managing biological collections and identifying species (using identification keys). Initial endeavors in classification were constrained to taxonomic attributes of the organism, neglecting considerations of interspecies relatedness. The phylogenetic classification emerged later in the 1950s to handle the evolutionary background of the creature. Here are the various schools of conventional taxonomy that, although adhering to the same principle of morphological similarity between species, diverge in how they see phylogenetic classification (Sinev 2012). There are also many different sampling and monitoring methods. Direct observations, direct sampling, and DNA-based methods are some of these methods. The pitfall traps are one of the most criticized but frequently used methods (Engel et al. 2017). They are invasive because they wipe study instances out of the environment. Moreover, the explication of variations across habitats poses challenges, given that each trapping technique is imbued with distinct biases and methodological idiosyncrasies (Skvarla et al. 2014). The direct observation method is non-invasive. Nonetheless, it necessitates the discernment of organisms by proficient entomologists within the study locale during the designated sampling period. This predicament significantly diminishes the potential pool of instances under consideration. (Hansen et al. 2020).

The enormous species diversity of insects (more than 1.02 million species reported so far) makes identification difficult (Zhang 2011). Besides, the substantial variety within species is caused by factors like sex, color morph, life stage, etc. Higher taxonomic categories, such as orders, could be distinguished with practice. However, once we get down to the family level, even for specialists, the work becomes rather difficult unless we narrow the problem down to a certain life stage, location, or insect order. The challenge of identification grows harder the lower the taxonomic level. Reliable species identification may necessitate years of education and specialization in a single insect taxon. These highly qualified taxonomists generally have insufficient time, especially for less conspicuous and aesthetically pleasing groupings. They need to use their time better elsewhere rather than on routine identifications (Chudzik et al. 2020). It is critical to identify insects according to species or higher taxonomic categories quickly and accurately in many situations. A significant fraction of the biological diversity in our world is made up of insects, and advancement in our understanding of the make-up and operation of the earth's ecosystems depends partly on our capacity to locate and recognize the insects that live there successfully. Concerns about the safety of human food and health also call for the quick and precise identification of insects (Valan et al. 2019).

In many sectors, machine learning and image processing have immense potential. Self-driving automobiles, improvements in medical technologies, and transportation are just a few examples of beneficial applications for this technology. There are still a lot of applications for it, though, and it has not yet reached its full potential. The image processing-based methods provide significant advantages over conventional entomology approaches. They also sample more individuals without destroying habitats or depleting the population of the species. Moreover, they need less maintenance (Collett & Fisher 2017). These systems were employed to detect and classify insects and other arthropods infrequently but progressively (Martineau et al. 2017; Hansen et al. 2020). However, it is difficult to distinguish insects in nature due to diversity and their responses to the environment. These issues pose a significant challenge for computer vision systems. The computer vision-based grasshopper and locust detection systems are still in their infancy. It is a manual process carried out by specialist agents or farmers. Liu et al. proposed a computer vision-based pest detection and classification system (Liu et al. 2016). They used saliency map and deep CNN learning methods. They obtained a dataset from the internet. They have carried out their experiments on a high-performance computer. Liu et al. proposed a novel pest detection network in laboratory conditions (Liu et al. 2019). The proposed network gives 75% mean accuracy and precision. Xia et al. proposed a fast region-based convolutional neural network (Fast RCNN) based insect detection system (Xia et al. 2018). They have obtained 89% detection accuracy. They have also obtained datasets from the internet. Ding & Taylor proposed a moth detection approach (Ding & Taylor 2016). A sliding window technique was employed, covering the entire image in their approach. In their strategy, anchor boxes were utilized to handle the entire image in a single iteration.

Species identification is very popular for mammals and birds (Norouzzadeh et al. 2018). However, arthropod-focused computer vision experiments have concluded that it is almost impossible for humans to identify species solely on images (Hansen et al. 2020). There is a need for automatic species detection and classification systems to take full advantage of image processing-based systems (Hansen et al. 2020). The species-level arthropod classification is carried out from museum collections (Hansen et al. 2020). They used a convolutional neural network (CNN) for classifying arthropods. Kasinathan et al. proposed an insect pest detection method (Kasinathan et al. 2021). The Wang, Xie, Deng, and IP102 datasets have been used. They have obtained 91.5% classification accuracy for nine class with the CNN model. Cheng et al. proposed a pest identification algorithm (Cheng et al. 2017). They used deep residual learning. They have obtained 98.67% classification accuracy for 10 classes. Xia et al. used an improved CNN model for the classification of crop insects. They have used an improved CNN model. Xie et al. used multi-level unsupervised feature learning methods (Xie et al. 2018) for pest identification. Nanni et al. proposed different Adam optimization methods for pest identification (Nanni et al. 2022). They have obtained 95.52% and 74.11% classification accuracy on Deng and IP102 datasets respectively.

It is important to extract textural features for insect species identification. The performance of the insect species classification is directly affected by the feature extraction step. All the methods mentioned above use deep learning approaches in both feature extraction and classification stages. In recent times, deep learning approaches have become increasingly popular in many areas. However, their major drawbacks are thought to be the substantial computational expense, the demand for many-dimensional input, and the costly hardware prerequisites. Consequently, the demand for hand-crafted approaches remains. The LBP is used to extract the spatial characteristics of grasshopper images. There are many variants of the LBP method. It has a strong discriminative ability and low computational complexity. While LBP has achieved significant performance in image processing, there is still a need for further exploration and investigation into its working principles. Numerous versions of LBP have been introduced in recent years. Several queries persist regarding LBP. One such inquiry involves understanding the extent to which the seemingly straightforward LBP feature conveys highly discriminative information about local structures. Additionally, there is a need to identify the information that may be lacking in the LBP feature and determine methods to articulate imperfect data within the LBP framework. In this study, the neutrosophic theory is combined with LBP. The neutrosophic theory plays a crucial role in differentiating the boundary areas among distinct texture motifs. It involves examining details on a small and subtle scale.

The aim of this paper is to propose an innovative hybrid approach based on neutrosophic set theory for the characterization of intra and interclass variations in grasshopper species classification. Additionally, the paper aims to introduce a novel open Grasshopper Classification Dataset (GHCD11) to facilitate research and development in the field of automated grasshopper species classification. For this reason, the neutrosophic completed local binary pattern (CLBP) is used to solve this problem and to increase classification accuracy.

2. Material and Methods

2.1. Material

The proposed dataset consists of 11 different types of grasshopper species imaged in the Southeast region of Turkey. This study mainly focuses on classifying grasshopper species using unique features and machine learning techniques based solely on a photo of the grasshopper. We have collected 28 images for each grasshopper species class. Each grasshopper image has different scale, size, rotation, shapes, backgrounds, and views. It is difficult to identify because of these variances.

2.1.1. Dataset

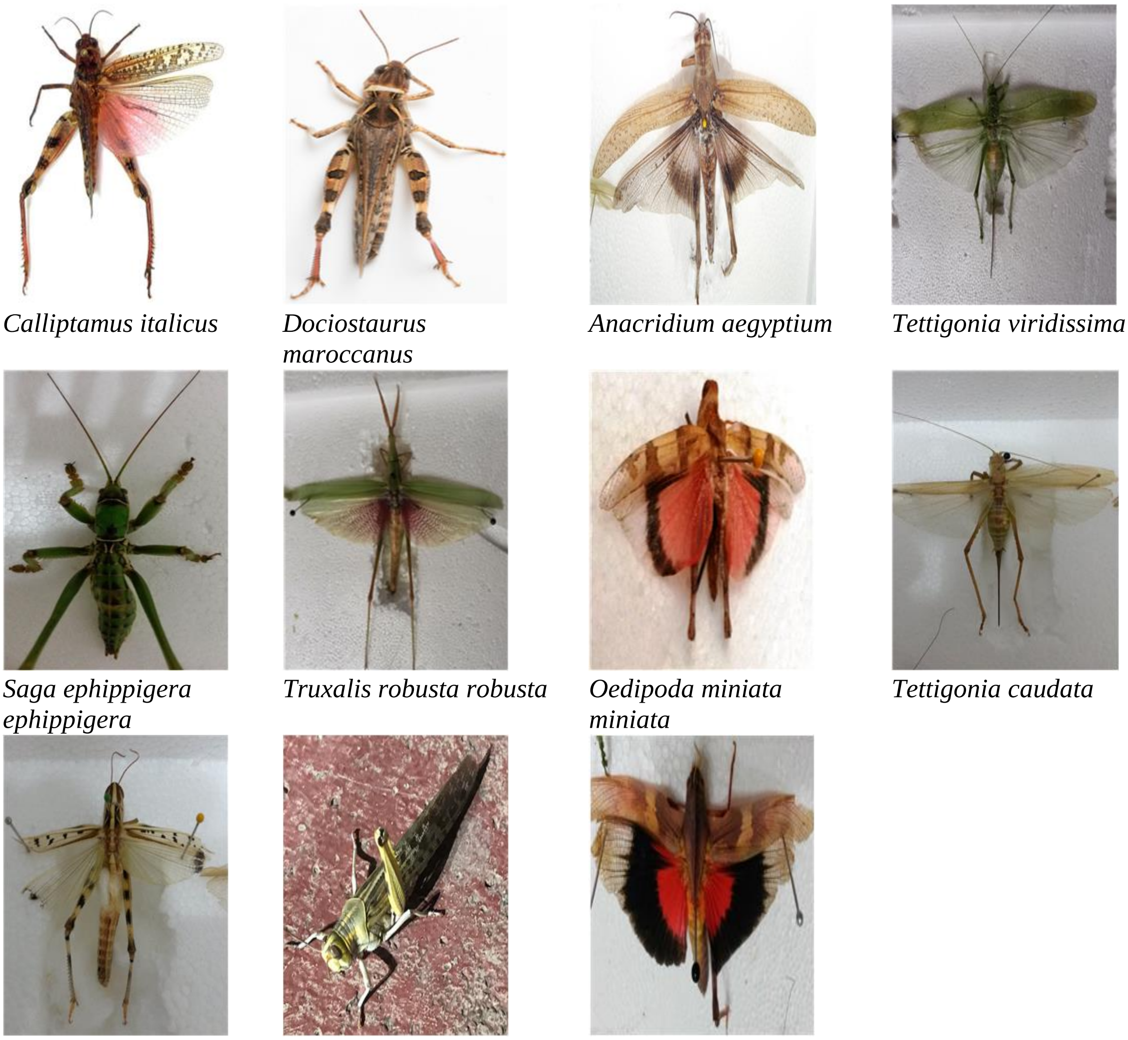
Calliptamus italicus, a species that damages plant crops in many parts of Turkey and has a swarm forming feature, caused regional damage to many agricultural lands in the Adakli district of Bingöl province in 2021 and caused economic loss (Qin et al. 2013; İlçin & Satar 2020; İlçin et al. 2021). *Dociostaurus maroccanus* species, known as the Moroccan locust, causes damage in areas such as gardens and orchards and is effective in many regions of Turkey, but it has been determined that it causes economic damage in the rural areas and cereal fields of the Bingöl Province (İlçin et al. 2021). It has been recorded that *Anacridium aegyptium* species, known as the corn locust, causes economic losses in cereals. *A. aegyptium*, which is a common species in the areas where cereals are grown in our country, is an important plant pest. *Tettigonia viridissima* species have been detected in the meadow, pasture, and cereal areas and are a common species locally. Although *Tettigonia caudata* species is omnivorous, it has the same genus as *T. viridissima* species and has typical similarities. This species causes damage, especially to cotton and alfalfa plants. *Truxalis robusta robusta* has been detected in many localities in the Eastern and Southeastern Anatolia regions, especially in wooded areas and in areas where cotton plants are dense. *Oedipoda miniata miniata* is one of the most common grasshopper species and it has been determined that it causes low damage in fields and gardens. *Notostaurus anatolica* is a species that could be recognized by the specific wing and leg bands collected in the nymph (instar) stage and its body patterns are unique. *Shistocerca gregaria* is one of the rare species of grasshoppers that have the ability to form swarms and spread over large areas, and according to the data of FAO, it caused 20.2 million people to face hunger in countries located on the African continent in 2020 (FAO 2020). It causes serious economic loss. *Pyrgodera armata* is a species with specific features in terms of the shape and color of its head and other body organs and is a grasshopper species whose damage status could not be determined. The locality was collected from the rural areas of Batman province, Southeastern Anatolia region. The species *Saga ephippigera ephippigera* is a predatory grasshopper known as the carnivorous grasshopper. It is an effective and useful grasshopper species observed in many places, especially in the Eastern and Southeastern Anatolian regions (İlçin & Satar 2018; İlçin 2019). Table 1 shows the sample images from 11 different grasshopper species.

Table 1- Details of the GHCD11 dataset

Grasshopper species	*Number of images*
Calliptamus italicus	672
Dociostaurus maroccanus	672
Anacridium aegyptium	672
Tettigonia viridissima	672
Saga ephippigera	672
Truxalis robusta	672
Oedipoda miniata miniate	672
Tettigonia caudata	672
Notostaurus anatolica	672
Schistocerca gregaria	672
Pyrgodera armata	672

2.1.2. Image resizing

Images in the dataset have different sizes. Every image has been resized to a 480 x 480 pixels size to standardize the image size and to trim down the computational expenses.



Calliptamus italicus | *Dociostaurus maroccanus* | *Anacridium aegyptium* | *Tettigonia viridissima*

Saga ephippigera ephippigera | *Truxalis robusta robusta* | *Oedipoda miniata miniata* | *Tettigonia caudata*

Notostaurus anatolica | *Schistocerca gregaria* | *Pyrgodera armata*

Figure 1- Sample images from 11 different grasshopper species

2.1.3. Data augmentation

In this paper, data augmentation is applied to artificially increase dataset size. It helps overcome the "not enough data" issue, prevents overfitting. The original dataset has 28 images for each grasshopper type. It is applied to the original dataset based on geometric transformations. Affine transformations are used in data augmentation to introduce variability to the training dataset. This helps prevent overfitting by exposing the model to a more diverse set of examples during training. Data augmentation through affine transformations, the model becomes more robust and generalizes better to variations present in real-world scenarios. The introduction of variations in the training data helps the model learn essential features and patterns without memorizing specific instances. The augmentations carried out via rotation, scaling, and reflection to enhance classification performance. The images in original dataset rotated by 45° angles in the range [0, 360] degrees and resized images by 2 random scale factors in the range [1.2, 2.1]. Thus, the dataset's size is expanded to be 24 times larger through augmentation of the number of images. Table 1 expresses the count of images for each grasshopper species in GHCD11 dataset.

2.2. Methods

This paper introduces a grasshopper species classification framework based on neutrosophic set without high-cost extra operations. This feature is widely embraced for handling diverse types of uncertainties. Additionally, it considers both spatial and boundary information, effectively addressing and managing uncertainties. A thorough texture analysis is conducted by amalgamating various features inherent in locust images. The LBP method exhibits insufficient robustness to outer variables like noise and low contrast, thereby adversely impacting the feature extraction procedure. Unwanted elements like noise and fluctuations in light within the input locust image can influence the quality of features. Consequently, there is a risk of information loss, leading to a decrease in performance during classification stage. Effectively suppressing noise components through a neutrosophic set is crucial for precise edge detection and the extraction of effective features. It is imperative to guarantee the preservation of discriminative features in diverse image regions. Otherwise, the classification performance may suffer due to the inability to extract these key features from the image. Failure to extract discriminative features from the image may result in a decline in classification performance.

2.2.1. Completed Local Binary Patterns (CLBP)

LBP encodes local texture information. To improve the LBP's performance in different textural situations, various variants are proposed (Ojala et al. 2002). The rotation, translation, and noise sensitivity make grasshopper image classification a difficult task. It's crucial to extract distinctive descriptors to effectively classify textures. The CLBP represents a modification of the LBP. It uses CLBP_M, CLBP_C, and CLBP_S operands as well as the local difference sign-magnitude transform (LDSMT). These three operands indicate the magnitude, center, and sign components (Guo et al. 2010). A general overview of the CLBP approach is shown in Figure 2. To express local patterns, the LBP only uses the sign vector. This could produce some inaccurate classification results. Compared to the magnitude components, the sign components are more accurate. Additionally, combining the sign and magnitude components yields significantly better results. The sign component conserves a lot of information about local differences. The magnitude component might offer additional discriminant data.

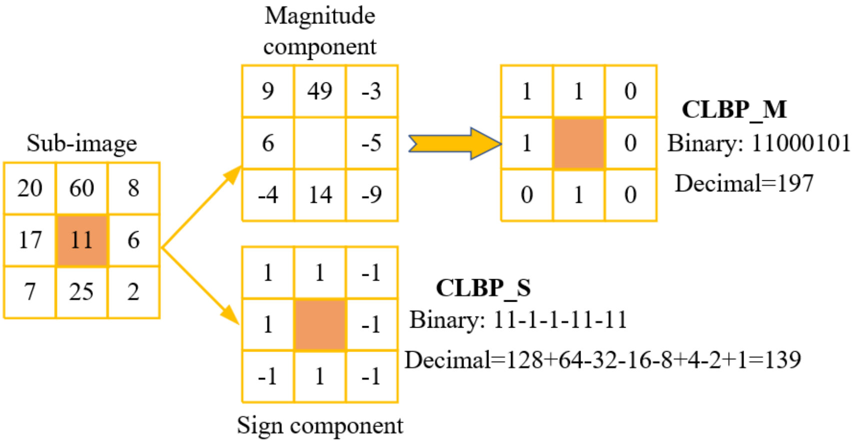


Figure 2- Illustration of completed local binary pattern

2.2.2. Neutrosophic set

The neutrosophic set (NS) theory looks into the nature and the extent of neutrality (Sengur et al. 2019). The degree of belonging is expressed using membership functions in standard fuzzy sets. In neutrosophy, every incident has an indeterminacy (I), falsity (F), and truth (T) level. The three memberships are utilized to determine an element's degree of T, I, and F. They are in [0, 1] interval. The degree of these functions could be used to quantify the uncertainty. The definitions of NS and uncertainty using NS were provided in (Alpaslan 2022).

The neutrosophic transformations of sample grasshopper images (*Calliptamus italicus*, *Dociostaurus maroccanus*) are demonstrated in Figure 3.

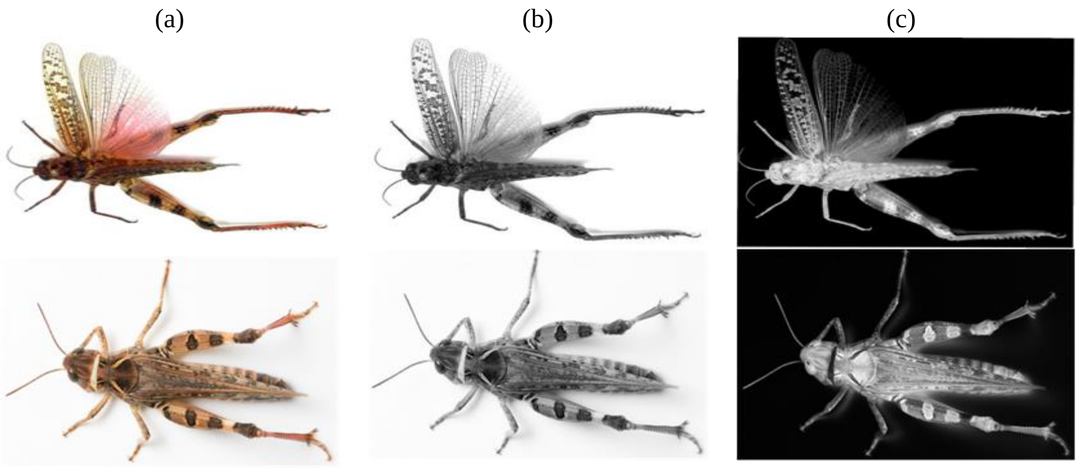


Figure 3- Subset of grasshopper images in the neutrosophic field. (a) Input image, (b) *T* (c) *F* sets

2.2.3. neutroCLBP

The local texture is encoded by most of the texture analysis methods in the literature. Unfortunately, these approaches do not successfully capture the boundaries. Additionally, the conventional methods are unable to eliminate the ambiguities in an image. Because they do not have any uncertainty management instruments. The neutroCLBP method could overcome these limitations without high-cost extra operations. The *T* and *F* sets in the neutrosophic field were used to characterize the grasshopper images. The *F* set was utilized with the *T* set in neutroCLBP. As a result, the complementary use of both false and truth components improves the LBP's distinctiveness.

There are two key stages in neutroCLBP. The input images are converted to neutrosophic field in the first stage. The second stage eliminates the deficiency of *I* set using LBP's local texture analysis capabilities. The neutroCLBP method makes advantage of the neutrosophic set and the CLBP's statistical texture analysis capabilities to suppress the image's noise. As a result, inter-class similarities and intra-class variations in grasshopper images are more defined. The neutroCLBP consists of three main stages: (1) neutrosophic components calculation, (2) extraction of neutroCLBP-M and neutroCLBP-S features, (3) cross-scale joint coding and histogram conjunction. The $neutroCLBP-M^{h}_{(P,R)}$ gives better results. For this reason, the $neutroCLBP-M^{h}_{(P,R)}$ was used in this study. The frame diagram of neutroCLBP is shown in Figure 4 (Alpaslan 2022).

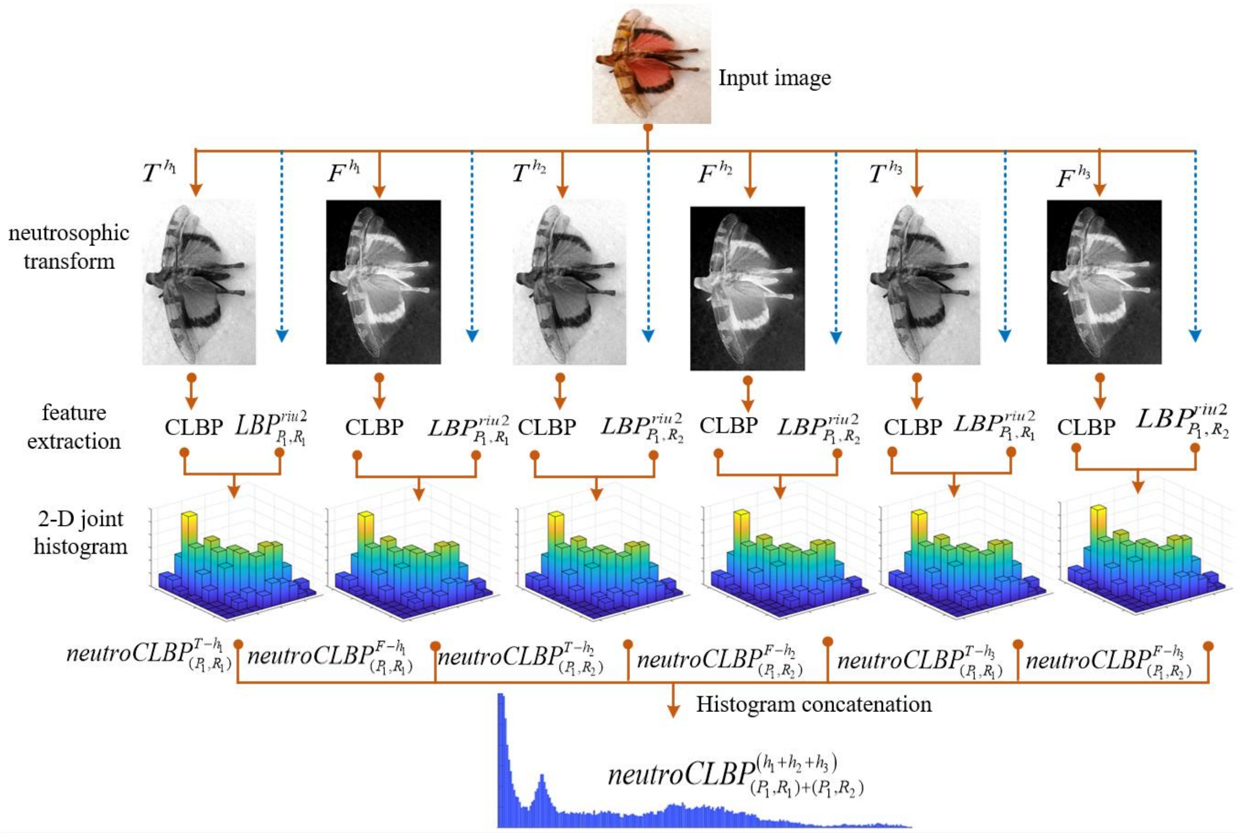


Figure 4- Calculation of neutroCLBP features

In order to encode neutrosophic membership components, the $LBP^{riu2}_{P,R}$ and the CLBP are combined (El Merabet et al. 2019). As a result, the CLBP_C features of *T* and *F* components were calculated.

The neutrosophic components of the grasshopper image are computed using various parameters. Various (P,R) and h values are used in multi-resolution texture analysis. The diameter of filters is determined with three different h. The radius R is also used for reduction of feature dimension of the neutroCLBP. The cross-scale joint coding is carried out as follows:

1) $LBP^{riu2}_{P,R}$ and $CLBP(x_c, y_c)^{riu2\_h}_{P,R}$ of T^{h_1} with (P_1,R_1) and h_1 are merged. The result is showed as $neutroCLBP(x_c, y_c)^{T-h_1}_{P_1,R_1}$.
2) $LBP^{riu2}_{P,R}$ and $CLBP(x_c, y_c)^{riu2\_h}_{P,R}$ of F^{h_1} with (P_1,R_1) and h_1 are merged. The result is showed as $neutroCLBP(x_c, y_c)^{F-h_1}_{P_1,R_1}$.
3) In $CLBP^{h_1}_{P_1,R_1}$ of T^{h_1}, h_1 is changed with h_2, and novel statement is denoted as $neutroCLBP^{T-h_2}_{P_1,R_1}$.

4) In $CLBP^{h_1}_{P_1,R_1}$ of F^{h_1}, h_1 is changed with h_2, and novel statement is denoted as $neutroCLBP^{F-h_2}_{P_1,R_1}$

5) $neutroCLBP(x_c, y_c)^{F-h_1}_{P_1,R_1}$, $neutroCLBP^{F-h_2}_{P_1,R_1}$, $neutroCLBP^{T-h_2}_{P_1,R_2}$ and $neutroCLBP^{T-h_3}_{P_1,R_2}$ are concatenated to construct the multi-resolution structure. This structure is denoted as $neutroCLBP^{(h_1+h_2+h_3)}_{(P_1,R_1)+(P_1,R_2)}$.

3. Experimental Results

Several experiments have been carried out to validate the effectiveness of the neutroCLBP on grasshopper species classification. The nearest neighbor (1-NN) and support vector machines (SVM) classifier were used to demonstrate the discriminative ability of the method (Gul et al. 2021). The one-versus-one approach and cubic kernel function were employed in our experimental studies. The SVM model performs noticeably better when nonlinear kernels are used in complicated problems. The 30% and 70% of the dataset were used for testing and training, respectively. Additionally, 5-fold cross-validation was applied.

The GHCD11 dataset was utilized in the first experiment to assess the effects of various parameters. The accuracy is used in experiments. The calculation for accuracy is shown in Eq (1).

$$accuracy = \frac{1}{p}\sum_{i=1}^{p}\frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \tag{1}$$

Where; *TP*, *TN*, *FP*, *FN* and *p* denotes the number of true positives, true negatives, false positives, false negatives, and classes respectively.

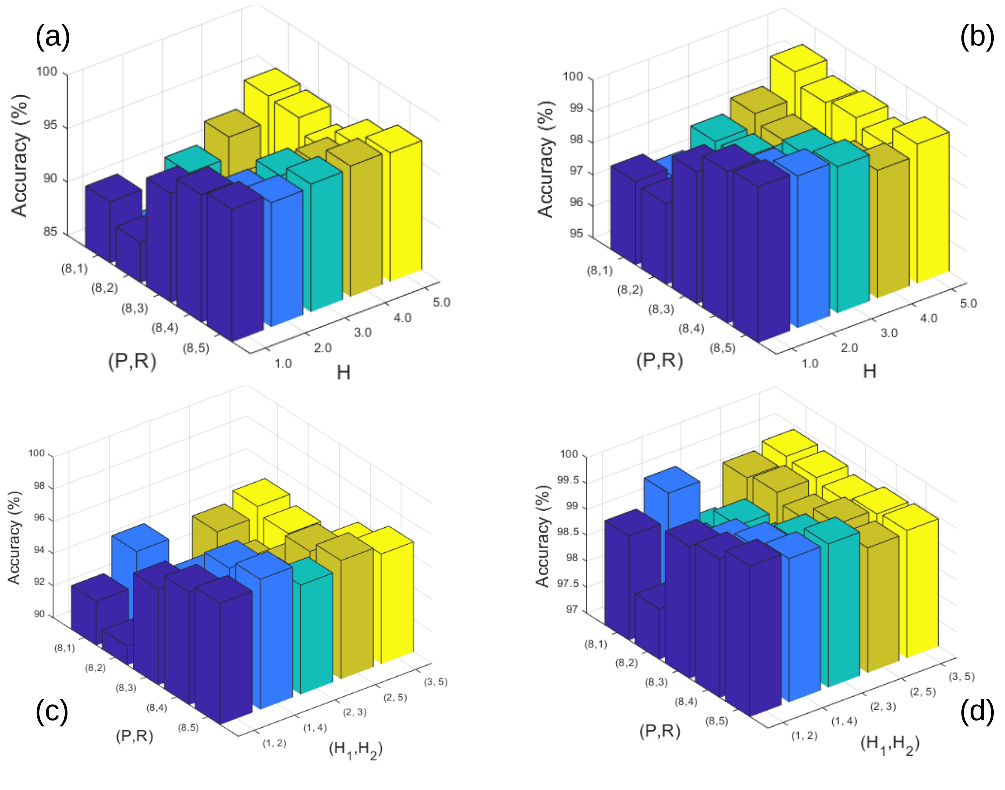


Figure 5- Experimental results (%) of $neutroCLBP^{h}_{(P,R)}$ for (a) kNN and (b) SVM classifier and $neutroCLBP^{h_1+h_2}_{(P_1,R_1)}$ for (c) kNN and (d) SVM classifiers

The experimental studies were carried out with Intel Xeon(R) E3-1241 v3 CPU, and 8 GB of RAM.

3.1. Analysis of neutroCLBP parameters

There are three different parameters in the $neutroCLBP^{h}_{(P,R)}$. These parameters are *h*, *R* and *P*. The *h* refers to neutrosophic filter; *R* refers to sampling radius; and *P* refers to sampling neighborhood. The *P* is fixed as 8 to lower the dimension and execution time. The *h* parameter determines the radius of filters. It is selected to detect statistical details in boundaries. The image is blurred and loses pixel brightness and textural variety when the *h* parameter is quite high. The filtered image is prone to noise if it is selected too small. In this paper, multi-scale texture analysis and investigation of optimal *h* parameter for grasshopper species classification have been carried out. Neutrosophic set noise suppression and enhancement of grasshopper images have positive impact on the classification performance. The method is analyzed with various *h* and *R* values ranging from 1 to 5.

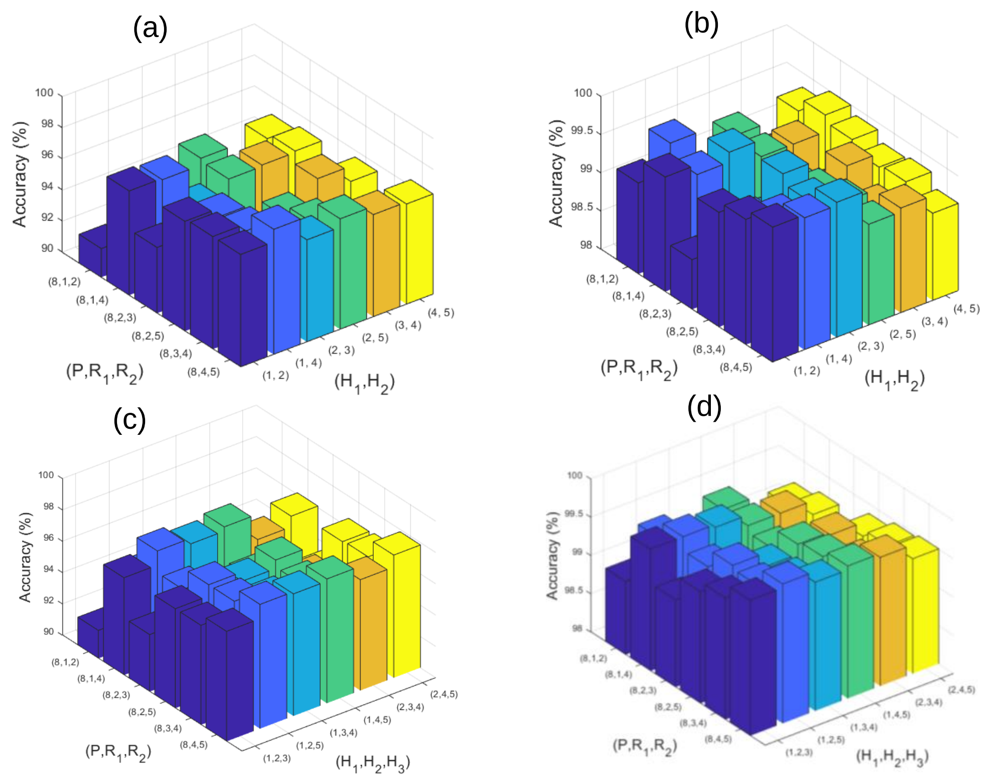


Figure 6- Experimental results (%) of $neutroCLBP^{h_1+h_2}_{(P_1,R_1)+(P_1,R_2)}$ for (a) kNN and (b) SVM classifiers and $neutroCLBP^{h_1+h_2+h_3}_{(P_1,R_1)+(P_1,R_2)}$ for (a) kNN and (b) SVM

The impact of *h* and *R* parameters were studied on classification performance using GHCD11 dataset. Firstly, the effects of these parameters on the $neutroCLBP^{h}_{(P,R)}$ were investigated. The most favorable outcomes were achieved with *R*=5 as shown in Figure 5. Additionally, it has been seen that the accuracy increases with *R* value. The best results were obtained with *R*=5 and *h*=1 parameters in $neutroCLBP^{h}_{(P,R)}$ for both 1-NN and SVM classifiers. The most favorable outcomes were achieved with *R*=5 and $h_1 = 1,\ h_2 = 4$ parameters in $neutroCLBP^{h_1+h_2}_{(P_1,R_1)}$ for 1-NN classifier. The best results were obtained with $h_1 = 1,\ h_2 = 4$ parameters in $neutroCLBP^{h_1+h_2}_{(P_1,R_1)}$ for SVM classifier. It is seen from Figure 5 that, the use of more than one *h* parameter increases the classification accuracy by 3.41% for 1-NN classifier and by 2.06% for 1-NN classifier on average. It is also clear from Figure 5 that the SVM classifier gives 7.77% better results than the 1-NN classifier on average.

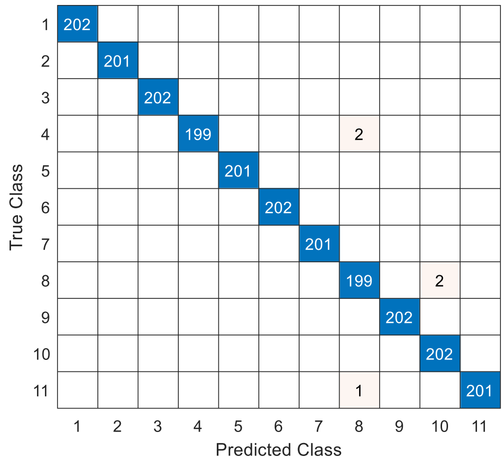


Figure 7- Confusion matrix for $neutroCLBP^{h_1+h_2+h_3}_{(P_1,R_1)+(P_1,R_2)}$

A confusion matrix is provided to demonstrate the multi-resolution $neutroCLBP^{h_1+h_2+h_3}_{(P_1,R_1)+(P_1,R_2)}$'s accuracy in a more thorough description in classifying grasshopper species as shown in in Figure 7.

Table 2- The comparison of the neutroCLBP with state-of-the-art deep learning models

Model	*Reference*	*Feature Size*	*Accuracy*
VGG16	Simonyan & Zisserman 2014	25.088	97.97
VGG19	Simonyan & Zisserman 2014	25.088	97.83
ResNet50	He et al. 2015	100.352	91.38
ResNet152	He et al. 2015	100.352	98.44
MobileNet	Sandler et al. 2018	50.176	97.78
DenseNet121	Huang et al. 2016	50.176	97.78
DenseNet169	Huang et al. 2016	81.536	99.34
DenseNet201	Huang et al. 2016	94.080	99.24
neutroCLBP	This paper	600	**99.77**

The proposed framework was evaluated with hand-crafted and deep learning-based approaches. Table 2 shows the comparison of the *neutroCLBP* method with current deep learning models. The *neutroCLBP* gives 99.77% classification accuracy on GHCD11 dataset. It is seen from Table 2 and 3 that, deep learning-based approaches yield better outcomes than other hand-crafted methods. However, the proposed framework was outperformed deep learning methods with reasonable computational cost. All models were trained utilizing stochastic gradient descent (SGD) method employing a batch size of 8, 50 epochs, 0.0001 learning rate and Adam optimization. DenseNet169 model gave the closest results. It is clear from Figure 8 that, the *neutroCLBP* shows high performance with low time complexity. The training times of *neutroCLBP* and different deep learning models were compared. VGG16 (Simonyan & Zisserman 2014), VGG19 (Simonyan & Zisserman 2014), ResNet50 (He et al. 2015), ResNet152 (He et al. 2015), MobileNet (Sandler et al. 2018), DenseNet121 (Huang et al. 2016), DenseNet169 (Huang et al. 2016), DenseNet201 (Huang et al. 2016) models were used for time comparison. The proposed methodology is about 40 times faster than the MobileNet, the fastest model, and about 152 times faster than the ResNet152, the slowest model. Besides, deep learning models require graphics cards with a significant amount of computing power. The proposed methodology has both higher classification accuracy and lower time complexity than deep learning models. Experimental results show that the proposed methodology is suitable for real-time applications. The experiments in Table 2 and Figure 8 were carried out with Intel i9 Intel Core i9-11900K CPU, GeForce RTX™ 3080 Ti graphics cards, and 64 GB of RAM.

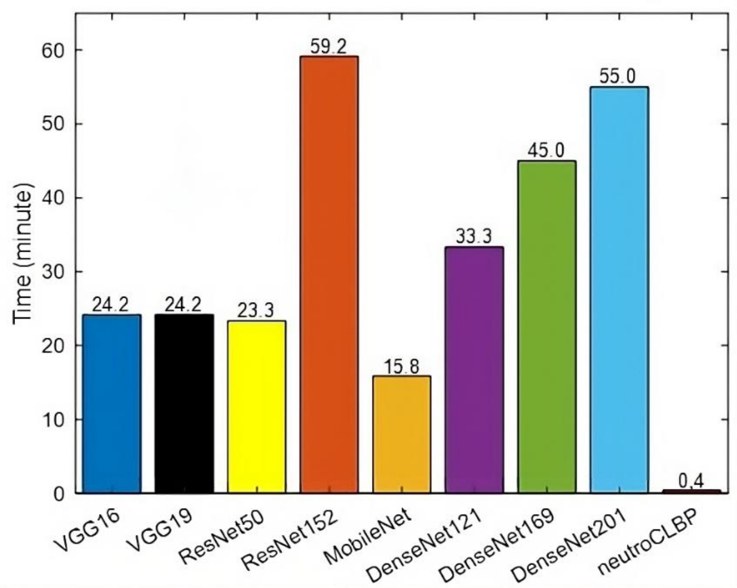


Figure 8- Comparison of models training time

Table 3- The comparison of *neutroCLBP* using current approaches on the GHCD11 dataset

Method	*Reference*	*Accuracy*
LBP	Ojala et al. 1996	89.35
CLBP_M	Guo et al. 2010	89.94
CLBP_S	Guo et al. 2010	90.44
LCvMSP	El Merabet & Ruichek 2018	95.67
LCxMSP	El Merabet & Ruichek 2018	93.59
RLBGC	El Khadiri et al. 2018	95.26
ALBGC	El Khadiri et al. 2018	96.62
ACS-LBP	El Merabet et al. 2019	98.92
LDTP	El khadiri et al. 2018	79.97
neutroCLBP	This paper	**99.77**

The classification performance of $neutroCLBP^{h_1+h_2}_{(P_1,R_1)+(P_1,R_2)}$ and $neutroCLBP^{(h_1+h_2)+(h_2+h_3)}_{(P_1,R_1)+(P_1,R_2)}$ were investigated from many aspects in the second experiment. The $neutroCLBP^{(h_1+h_2)+(h_2+h_3)}_{(P_1,R_1)+(P_1,R_2)}$ was evaluated with three *h* and two *R* values. The $neutroCLBP^{h_1+h_2}_{(P_1,R_1)+(P_1,R_2)}$ was evaluated with two *h* and two *R* values. A multi-scale analysis was conducted since the neutrosophic components were obtained for multiple *h* values. The GHCD11 database contains different types of grasshoppers. Figure 6 shows the classification accuracy results of $neutroCLBP^{h_1+h_2}_{(P_1,R_1)+(P_1,R_2)}$ and $neutroCLBP^{h_1+h_2+h_3}_{(P_1,R_1)+(P_1,R_2)}$ features for both kNN and SVM classifiers. The best results were obtained with $R_1=4, R_2=5$ and $h_1=1,\ h_2=2$ parameters in $neutroCLBP^{h_1+h_2}_{(P_1,R_1)+(P_1,R_2)}$ for the SVM classifier. The best results were obtained with $R_1=4, R_2=5$ and $h_1=1,\ h_2=2,\ h_3=3$ parameters in $neutroCLBP^{(h_1+h_2)+(h_2+h_3)}_{(P_1,R_1)+(P_1,R_2)}$ for the SVM classifier. It is seen from Figure 6 that, the use of more than one parameter increases the classification accuracy. Besides, the $neutroCLBP^{(h_1+h_2)+(h_2+h_3)}_{(P_1,R_1)+(P_1,R_2)}$ gives slightly better classification results than $neutroCLBP^{h_1+h_2}_{(P_1,R_1)+(P_1,R_2)}$. However, its computational complexity is high. Experimental results show that multi-scale features yielded superior results.

4. Discussion

Diagnosis and identification processes for insects include many features; It could consist of very complex and controversial applications carried out within the framework of the examination and interpretation of morphological, physiological, genital, and other molecular markers. Although the margin of error increases in the diagnostics and definitions that are made sometimes, the misuse of the diagnostic keys could prolong the process and misdiagnoses could lead to an inevitable situation. Traditional predictive diagnostic methods are only limited by the data they have. Therefore, collecting and obtaining data is difficult. However, for new researchers or young researchers interested in insects, the use of more professional diagnostic tools could provide solutions to many important problems. With this study, it will be possible to carry out species-diagnostic applications of high accuracy and clear information in the determination of many agricultural pests, especially insects. In addition, it will prevent the emergence of erroneous applications with time and wrong results in diagnosis. With the created digital application, one of the main results is to prevent the damage they will cause by identifying and learning about many plant pests, especially grasshoppers. Some of the other very important results are; It seems possible to provide guidance for teaching purposes, data-forecasting, agricultural practices, and pesticide use.

The following are the paper's main contributions: A large dataset containing 11 different types of grasshopper species, named the GHCD11. It consists of mix of images with different scale, size, rotation, shapes, backgrounds, and views. In this way, any classification algorithm could be strong enough to deal with these challenges. A novel methodology that could classify the grasshopper species with high accuracy is proposed. The proposed methodology consists of various stages. The grasshopper images are first converted to neutrosophic field in the first stage. Thus, stronger features are obtained. Additionally, a cross-scale joint approach is used to integrate these discriminatory features with rotation invariant LBP features. The proposed methodology has both higher classification accuracy and lower time complexity than deep learning models. The proposed methodology could achieve up to 99.7% classification accuracy even while working with challenging dataset of wide range in image scale, rotation, and size. It does not require high-powered expensive graphics cards. Besides, grayscale grasshopper images exhibit greater resilience to noise in the neutrosophic domain. The presence of neutrosophic components plays a role in noise suppression, leading to a more accurate detection of edges. The proposed approach guarantees the extraction of more robust features. This marks the inaugural integration of neutrosophic set theory and the LBP method into an insect species identification system through extensive state-of-the-art research. The method put forth demonstrates resilience in the face of challenges such as noise, rotation, and variations in lighting conditions.

5. Conclusions

Species identification in grasshoppers has traditionally been based on body organs (cephalo, thorax, and abdomen) and genital organs (genital plate). There is no automatic grasshopper species identification system available yet. In this study, many plant pests and other insect species (grasshoppers, etc.) have been detected by image processing and machine learning methods. The CLBP method undergoes reinterpretation through the lens of the neutrosophic set, leading to the development of a novel feature extraction approach that leverages neutrosophic set components. The outcome is an image imbued with enhanced significance, achieved through the mitigation of noise effects. Besides, the *neutroCLBP* based grasshopper classification framework is proposed. In this framework, the neutrosophic membership components of the grasshopper image were utilized. In this way, noise effects were minimized, and a more informative image was generated. The proposed framework has remarkable performance with acceptable computational expense. The experiments represent that the proposed framework could accurately classify grasshopper species. The comparative experiments show that the *neutroCLBP* improves the classification accuracy of LBP-based current handcrafted features at least about 1% and at most about 11% thanks to the hybrid approach. Besides, the multi-scale strategy positively affected discrimination of the grasshopper images with high intra-class variation and inter-class similarity. The average accuracy is up to 99.77%, which is higher than VGG16, VGG19, ResNet50, ResNet152, MobileNet, DenseNet121, DenseNet169, DenseNet201 models with reasonable computational cost. The proposed framework could also used for education in plant protection.

Declarations

Conflicts of Interest: No conflict of interest was declared by the authors.

Declaration of Ethical Standards: The authors of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

Author Contributions: Study conception and design: NA, Mİ; data collection: Mİ; software, methodology: NA; analysis and interpretation of results: NA, Mİ; draft manuscript preparation: NA, Mİ. All authors reviewed and approved the final version of the manuscript.

References

Alpaslan N (2022). Neutrosophic set based local binary pattern for texture classification. Expert Systems with Applications 209: 118350. - doi: 10.1016/J.ESWA.2022.118350

Cheng X, Zhang Y, Chen Y, Wu Y & Yue Y (2017). Pest identification via deep residual learning in complex background. Computers and Electronics in Agriculture 141: 351–356

Chudzik P, Mitchell A, Alkaseem M, Wu Y, Fang S, Hudaib T, Pearson S & Al-Diri B (2020). Mobile Real-Time Grasshopper Detection and Data Aggregation Framework. Scientific Reports 2020 10:1. 10: 1–10. - doi: 10.1038/s41598-020-57674-8

Collett R A & Fisher D O (2017). Time-lapse camera trapping as an alternative to pitfall trapping for estimating activity of leaf litter arthropods. Ecology and Evolution 7: 7527–7533

Ding W & Taylor G (2016). Automatic moth detection from trap images for pest management. Computers and Electronics in Agriculture 123: 17–28. - doi: 10.1016/J.COMPAG.2016.02.003

El Khadiri I, Chahi A, El Merabet Y, Ruichek Y& Touahni R (2018). Local directional ternary pattern: A New texture descriptor for texture classification. Computer Vision and Image Understanding 169: 14–27

El Khadiri I, Kas M, El Merabet Y, Ruichek Y& Touahni R (2018). Repulsive-and-attractive local binary gradient contours: New and efficient feature descriptors for texture classification. Information Sciences 467: 634–653. - doi: 10.1016/J.INS.2018.02.009

El Merabet Y& Ruichek Y (2018). Local Concave-and-Convex Micro-Structure Patterns for texture classification. Pattern Recognition 76: 303–322. - doi: 10.1016/J.PATCOG.2017.11.005

El Merabet Y, Ruichek Y& el idrissi A (2019). Attractive-and-repulsive center-symmetric local binary patterns for texture classification. Engineering Applications of Artificial Intelligence 78: 158–172

Engel J, Hertzog L, Tiede J, Wagg C, Ebeling A, Briesen H, Weisser W W, Engel C, Hertzog L, Tiede J, Wagg C, Ebeling A, Briesen H& Weisser W W (2017). Pitfall trap sampling bias depends on body mass, temperature, and trap number: insights from an individual-based model. Ecosphere 8: e01790.

FAO | Food and Agriculture Organization of the United Nations (2020). Desert Locust Upsurge.

Gul E, Alpaslan N & Emiroglu M E (2021). Robust optimization of SVM hyper-parameters for spillway type selection. *Ain Shams Engineering Journal* 12: 2413–2423. - doi: 10.1016/J.ASEJ.2020.10.022

Gullan P & Cranston P (2014). The insects: an outline of entomology. John Wiley & Sons.

Guo Z, Zhang L& Zhang D (2010). A completed modeling of local binary pattern operator for texture classification. IEEE Transactions on Image Processing 19: 1657–1663

Hansen O L P, Svenning J C, Olsen K, Dupont S, Garner B H, Iosifidis A, Price B W & Høye T T (2020). Species-level image classification with convolutional neural network enables insect identification from habitus images. Ecology and Evolution 10: 737–747

He K, Zhang X, Ren S& Sun J (2015). Deep Residual Learning for Image Recognition. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, Vol. 2016-December

Huang G, Liu Z, van der Maaten L & Weinberger K Q (2016). Densely Connected Convolutional Networks. In: 30th IEEE Conference on Computer Vision and Pattern Recognition. Institute of Electrical and Electronics Engineers Inc., Vol. 2017-January

İlçin M (2019). Investigation of Orthoptera: Insecta Fauna of Useful, Harmful and Predator Species in the Batman Region (Turkey). Science Stays True Here" Biological and Chemical Research 6: 30–40

İlçin M & Satar A (2018). On the Orthopteran Fauna (Insecta: Orthoptera) of Agricultural Regions of Batman Province (Turkey). Boletín de la Sociedad Entomológica Aragonesa (SEA) 62: 163-166

İlçin M & Satar A (2020). Dociostaurus (Dociostaurus) maroccanus Thunberg, 1815 (Acrididae:Orthoptera) Türünün Sürü Oluşturma ve Bitkilere Zarar Durumunun Araştırılması. *Turkish Journal of Nature and Science* 9: 80–83

İlçin M, Satar A & Balkaya A (2021). Remarks on the outbreak of Calliptamus italicus Linnaeus, 1758 (Acrididae: Orthoptera) in Bingöl province, Turkey., pp. 259–261. Retrieved from https://www.researchgate.net/publication/357480371

Kasinathan T, Singaraju D & Uyyala SR (2021). Insect classification and detection in field crops using modern machine learning techniques. Information Processing in Agriculture 8: 446–457

Liu L, Wang R, Xie C, Yang P, Wang F, Sudirman S & Liu W (2019). PestNet: An End-to-End Deep Learning Approach for Large-Scale Multi-Class Pest Detection and Classification. IEEE Access 7: 45301–45312

Liu Z, Gao J, Yang G, Zhang H, He Y (2016). Localization and Classification of Paddy Field Pests using a Saliency Map and Deep Convolutional Neural Network. Scientific Reports 2016 6:1. 6: 1–12

Martineau M, Conte D, Raveaux R, Arnault I, Munier D & Venturini G (2017). A survey on image-based insect classification. Pattern Recognition 65: 273–284

Nanni L, Manfè A, Maguolo G, Lumini A & Brahnam S (2022). High performing ensemble of convolutional neural networks for insect pest image detection. Ecological Informatics 67: 101515

Norouzzadeh M S, Nguyen A, Kosmala M, Swanson A, Palmer M S, Packer C & Clune J (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proceedings of the National Academy of Sciences of the United States of America 115: E5716–E5725.

Ojala T, Pietikäinen M & Harwood D (1996). A comparative study of texture measures with classification based on featured distributions. Pattern Recognition 29: 51–59

Ojala T, Pietikainen M & Maenpaa T (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 24: 971–987

Qin Y, Li Z, Zhao L, Fowler G, Fang Y (2013). The current and future potential geographical distribution of the Italian Locust, Calliptamus italicus (Linnaeus) (Orthoptera: Acrididae) in China. IFIP Advances in Information and Communication Technology 393 AICT: 290–298

Sandler M, Howard A, Zhu M, Zhmoginov A & Chen L C (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society

Sengur A, Budak U, Akbulut Y, Karabatak M & Tanyildizi E (2019). A survey on neutrosophic medical image segmentation. In: Neutrosophic Set in Medical Image Analysis

Simonyan K & Zisserman A (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. In: 3rd International Conference on Learning Representations. International Conference on Learning Representations, ICLR

Sinev S Y (2012). Coordinating the Traditional and Modern Approaches in the Systematics of Insects. Original Russian Text © S.Yu. Sinev. 92: 821–832

Skvarla M J, Larson J L & Dowling A P G (2014). Pitfalls and preservatives: a review. *The Journal of the Entomological Society of Ontario*. 145 pp

Sreedevi K, Meshram N & Shashank P R (2015). Insect Taxonomy—Basics to Barcoding. New Horizons in Insect Science: Towards Sustainable Pest Management pp. 3–12

Valan M, Makonyi K, Maki A, Vondráček D, Vondráček V & Ronquist F (2019). Automated taxonomic identification of insects with expert-level accuracy using effective feature transfer from convolutional networks. Systematic Biology 68(6): 876-895

Xia D, Chen P, Wang B, Zhang J & Xie C (2018). Insect Detection and Classification Based on an Improved Convolutional Neural Network. Sensors 2018, Vol. 18, Page 4169. 18: 4169

Xie C, Wang R, Zhang J, Chen P, Dong W, Li R, Chen T & Chen H (2018). Multi-level learning features for automatic classification of field crop pests. Computers and Electronics in Agriculture 152: 233–241

Zhang L, Lecoq M, Latchininsky A & Hunter D (2019). Locust and Grasshopper Management. https://doi.org/10.1146/annurev-ento-011118-112500 64: 15–34

Zhang Z (2011). Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. Magnolia press