













## Smart Farming with Sooty Tern Optimization based LS-HGNet Classification Model

V. Gokula Krishnan<sup>1\*</sup>, B. Vikranth<sup>2</sup>, M. Sumithra<sup>3</sup>, B. Prathusha Laxmi<sup>4</sup> and B. Shyamala Gowri<sup>5</sup>



<sup>1</sup>Department of CSE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Thandalam, Chennai, Tamil Nadu, India; <sup>2</sup>Department of CSE-Cyber Security, CVR College of Engineering, Hyderabad, Telangana, India; <sup>3</sup>Department of IT, Panimalar Engineering College, Poonamallee, Chennai, Tamil Nadu, India; <sup>4</sup>Department of AI & DS, R.M.K. College of Engineering and Technology, Gummidipoondi, Tamil Nadu, India; <sup>5</sup>Department of CSE, Easwari Engineering College, Ramapuram, Chennai, Tamil Nadu, India

### E-mail/Orcid Id:

**VGK**,  [gokul\\_kris143@yahoo.com](mailto:gokul_kris143@yahoo.com),  <https://orcid.org/0009-0005-6819-6729>; **BV**,  [b.vikranth@cvr.ac.in](mailto:b.vikranth@cvr.ac.in),  <https://orcid.org/0000-0002-1424-6584>; **MS**,  [msumithra@panimalar.ac.in](mailto:msumithra@panimalar.ac.in),  <https://orcid.org/0000-0003-4934-9062>; **BPL**,  [prathushalaxmi@gmail.com](mailto:prathushalaxmi@gmail.com),  <https://orcid.org/0000-0003-2248-5486>; **BSG**,  [shyamala.b@eec.srmmp.edu.in](mailto:shyamala.b@eec.srmmp.edu.in),  <https://orcid.org/0000-0003-1256-1836>

### Article History:

Received: 22<sup>nd</sup> Nov., 2023

Accepted: 10<sup>th</sup> Mar., 2024

Published: 30<sup>th</sup> Mar., 2024

### Keywords:

Maize tassel, drones, deep learning, stacked hourglass network, sooty tern optimization

### How to cite this Article:

V. Gokula Krishnan<sup>1\*</sup>, B. Vikranth<sup>2</sup>, M. Sumithra<sup>3</sup>, B. Prathusha Laxmi<sup>4</sup> and B. Shyamala Gowri<sup>5</sup> (2024). Smart Farming with Sooty Tern Optimization- Based LS-HGNet Classification Model. *International Journal of Experimental Research and Review*, 37(Spl.), 96-108.

### DOI:

<https://doi.org/10.52756/ijerr.2024.v37spl.008>

**Abstract:** Smart farming technologies enable farmers to use resources like water, fertilizer and pesticides as efficiently as possible. This paper discusses how Unmanned Aerial Vehicle (UAV) pictures can be used to automatically detect and count tassels, thereby advancing the advancement of strategic maize planting. The real state of affairs in cornfields is complicated, though, and the current algorithms struggle to provide the speed and accuracy required for real-time detection. This research employed a sizable, excellent dataset of maize tassels to solve this problem. This paper suggests using the bottom-hat-top-hat preprocessing technique to address the lighting irregularities and noise in maize photos taken by drones. The Lightweight weight-stacked hourglass Network (LS-HGNet) model is suggested for classification. The hourglass network structure of LS-HGNet, which is mostly utilised as a backbone network, has allowed significant advancements in the discovery of maize tassels. In light of this, the current work suggests a lighter variant of the hourglass network that also enhances the accuracy of tassel detection in maize plants. The additional skip connections used in the new hourglass network architecture allow minimal changes to the number of network parameters while improving performance. Consequently, the suggested LS-HGNet classifier lowers the computational burden and increases the convolutional receptive field. The hyperparameter tuning process is then carried out using the Sooty Tern Optimisation Algorithm (STOA), which helps increase tassel detection accuracy. Numerous tests were conducted to verify that the suggested approach is more accurate at 98.7% and more efficient than the most advanced techniques currently in use.

## Introduction

In order to increase farming operations' productivity, sustainability and efficiency, smart farming incorporates modern technologies (Karunathilake et al., 2023). Drones, also known as unmanned aerial vehicles, are cutting-edge techniques in contemporary agriculture that transform conventional farming methods through smart farming. These aerial vehicles, equipped with state-of-the-art sensors, cameras, and data analytics capabilities, enable farmers to manage their fields and crops in

entirely new ways. By employing UAV technology to deliver precise, real-time insights into crop health, soil conditions, and overall farm management, smart farming has advanced beyond conventional methods (Akkem et al., 2023; Dawn et al., 2023; Lachgar et al., 2023). Farmers now have access to valuable data that they can use to increase crop yields, make educated decisions, and promote sustainable agricultural practises thanks to the ability of these unmanned devices to collect high-resolution data across vast agricultural landscapes (El-



Ghamry et al., 2023). Smart farming with UAVs represents a significant advancement in agricultural productivity, efficiency, and environmental stewardship by combining technology with the time-honoured practise of cultivation. In the end, this approach shapes the future of farming (Dhruva et al., 2023).

What makes deep learning different from other machine learning branches is its ability to recognise patterns or representations automatically from raw data without explicit programming (Thirumalraj et al., 2024). Essentially, multiple layers of DL models are composed of systematically arranged networks of connected neurons or nodes (Meng et al., 2023). The advantages of deep learning have revolutionised numerous sectors, such as computer vision, natural language processing, healthcare, finance, and autonomous vehicles (Salehi et al., 2023). Due to deep learning, complicated problems that were previously challenging to resolve with traditional machine-learning techniques can now be resolved (Yuan et al., 2023).

CNNs process images of crops, fields, and plants from various sources. They have the capacity to identify diseases, assess crop health, assess growth stages, and monitor environmental variables (Zhang et al., 2023). CNNs analyse these visual cues and provide farmers with pertinent information that helps them make timely decisions. Conventional agricultural methods are transformed by CNNs used in smart farming through the use of computer vision (El-Ghamry et al., 2023). These networks enable farmers to make data-driven decisions, increase productivity, enhance sustainability, and optimise resource usage (Thorat et al., 2023; Aishwarya et al., 2023).

### Motivation

The suggested work offers a revolutionary method for transforming the planting of maize by leveraging cutting-edge smart farming technologies. This study is motivated by the recognition of precision agriculture's role in optimising the use of resources like water, fertilisers, and pesticides in contemporary farming practices. Using UAVs to automatically detect and count maize tassels is a crucial first step towards intelligent maize planting. However, because real-field scenarios are complex, current algorithms have difficulty achieving accurate real-time detection. This work presents a comprehensive strategy that includes a large, high-quality dataset and a novel preprocessing method to address image noises. The development of LS-HGNet, a more efficient and lightweight version of the Hourglass Network, significantly improves tassel detection accuracy while reducing processing load. Extended convolutional

receptive field with optimal performance is achieved by multiplying the number of skip connections. The accuracy of tassel detection is further enhanced by hyperparameter tuning with the SHOA. These developments culminate in a recommended method that achieves an amazing accuracy of 98.78%, surpassing the most recent techniques. This study makes a compelling case for the creation of clever methods for planting maize and emphasises how cutting-edge technologies have the revolutionary potential to revolutionise agricultural practices, increase crop yield, and promote sustainability.

### Main Contributions

- **Preprocessing Strategy:** explains how to fix problems like noise and erratic lighting in drone-captured photos of maize by using the "bottom-hat-top-hat strategy".
- **Proposed Model:** LS-HGNet is offered in the LS-HGNet Model as a more successful iteration of the hourglass network.
- **SHOA for Hyperparameter Tuning:** The accuracy of the proposed LS-HGNet classifier in tassel detection is increased by using the STOA for hyperparameter tuning.
- **Evaluation:** In this paper, performance metrics like Accuracy (ACC), Specificity (SP), F1-Score (F1), Recall (RC), and Precision (PR) are quantified to evaluate the overall results.

### Chapter Organisation

The format of the following is the paper: But Section 2 offers a much more in-depth examination of pertinent data. Section 3 briefly summarizes the suggested paradigm, while Section 4 explains the study's findings and validation process. Section 5 provides a summary of the findings to wrap up the investigation.

### Related Works

In order to improve the unmanned aerial vehicle (UAV) and the dataset data acquisition method, images of maize tassels gathered over various eras were first obtained, balancing picture quality and acquisition efficiency. Moreover, an attention mechanism was included to remove undesired elements and reduce noise (such as occlusions and overlaps) in the main features. Expanding upon YOLOX, this strong detection network has shown to be more dependable and suitable for use in intricate natural settings. The experiment's results showed 95.0% for the mean average precision (mAP@0.5), supporting the study's hypothesis. When the average values of the original model were compared to the increases were 1.7%, 1.8, 5.3, and 1.5 for the mAP@0.5, mAP@0.5–0.95, mAP@0.5–0.95 (area=small), and

mAP@0.5–0.95 (area=medium), respectively. The suggested technique successfully satisfied the vision system's requirements for resilience and precision in the detection of maize tassels.

The impact of the RetinaNet model on mapping variations in plant variety, planting density, brightness, and image resolution was investigated in the paper by Wang et al. (2013). As compared to the original RetinaNet model, the enhanced RetinaNet model demonstrated a significant improvement in identifying maize tassels. The average precision, recall rate, and precision for this study were, respectively, 0.9036, 0.9717 and 0.9802. Compared to RetinaNet, an enhanced version of the original model, increased recall rate, precision, and average precision by 4.6%, 1.57%, and 1.84%, respectively. The improved RetinaNet model identified smaller maize tassels more accurately than well-known target detection models like Faster R-CNN, YOLOX and SSD. Maize tassel detection deteriorated as the resolution decreased for equal-area images with varying resolutions. It also investigated how brightness affected detection in the different models. As the image's brightness rose, it became harder to identify maize tassels, especially for smaller ones. This study also examined the different models used to identify the tassels on five distinct types of maize. Zhengdan958, with R2 values of 0.9708, 0.9759 and 0.9545 on August 5, 9 and 20, 2021, respectively, was the most easily detected tassel. In the end, several models were used to recognize corn tassels planted at varying densities. Regarding Zhengdan958 tassel identification, the mean absolute errors at 29,985, 44,978, 67,466 and 89,955 plants/hm<sup>2</sup> were, respectively, 0.18, 0.26, 0.48 and 0.63. The planting density increased with a gradual increase in the detection error. Furthermore, this study offered a novel technique for small-scale maize tassel identification in farmland, enabling high-precision tassel identification. This technology would enable high throughput analysis of the phenotypic traits of maize.

Based on YOLOv7 as the original model, Pu et al. (2023) proposed an approach to maize tassel detection using a Tassel-YOLO model. The model used a VoVGSCSP module in the neck section in addition to the GSConv convolution, enhanced it to a SIOU loss function, and included a global attention mechanism. The model's computation cost and model parameters were 4.11 M and 11.4 G below, in that order, than those of YOLOv7. The counting accuracy went up to 97.55%. Experimental results show that Tassel-YOLO outperformed other widely used object detection algorithms in general. As a result, Tassel-YOLO

provided a unique method for detecting maize tassels using UAV aerial photos and successfully met the needs for real-time detection, thereby serving as an investigation of the YOLO network architecture.

Following the path of specialists in machine learning, the study by Lu et al. (2023) introduced Yolov8's cutting-edge technology to the study of plants. There were also a few simple yet effective adjustments made. The Path Aggregation Network (PANet) was designed to compensate for the resolution loss caused by the larger receptive field by integrating shallow-level information. In order to maximize the precision of up-sampled features, Content-Aware Re-assembly of Features (CARAFE), a lightweight up-sampling operator, was combined with the Multi-Efficient Channel Attention (Mlt-ECA) technique. The combined technique, dubbed Yolov8-UAV, greatly enhanced the ability to recognize small objects in images of unmanned aerial vehicles (UAVs). Four different plant species were included in the datasets that served as the basis for the analysis. Test results demonstrated that the proposed method had sufficient resilience and was highly competitive even against the most advanced counting techniques. In addition, a new dataset of cotton bolls with thorough bounding box annotations was made available to advance multidisciplinary computer vision and plant science research. New labels were supplied to correct previous errors in publicly available wheat ear datasets, which are in line with global research advances. All things considered, this research gave practitioners a solid approach to dealing with issues pertaining to practical implementation. Yolov8-UAV was advised to be used for UAV scenarios. Yolov8-N, on the other hand, was a good option for general scenes because of its generally good accuracy and speed. Two notable datasets with research value were supplied in order to promote the application of data sources to plant science. To summarise, the contribution entailed enhancing Yolov8's application in UAV scenarios and releasing two datasets that included bounding boxes.

A brand-new one-stage, single-level feature-based, Maize tassel detector without anchor (MT-Det) was proposed in the paper by Zeng et al. (2023). It was supposed to be simple but effective. Extensive analyses revealed that in terms of inference speed and detection accuracy, MT-Det performed better than feature pyramid detectors and one-level counterparts. In order to tackle the problem of notable accuracy decline when making direct inferences from high-resolution images, the proposed MT-Det improved mean average precision (mAP) by 13% and 38%, respectively, on proximal and

unmanned aerial vehicle (UAV) high-resolution images by incorporating a collection of technologies for hyper inference aided by slicing. MT-Det offered a practical, high-throughput method for precise and effective maize tassel detection and counting in real-world field settings.

Thermal and multispectral UAV remote sensing techniques were employed in the study by Jia et al. (2023) to monitor two different kinds of leaf spot diseases in maize: *Bipolaris maydis* is the cause of southern leaf blight, while *Curvularia lutana* causes curvularia leaf spot. Four cutting-edge classifiers were compared in order to create the best classification model to track the occurrence of these diseases: back propagation neural networks, support vector machines, random forest (RF), and extreme gradient boosting. The most helpful features for identifying four phases of the maize leaf spot illness including 4, 12, 19 and 30 days after inoculation—were identified using recursive feature elimination (RFE). The findings demonstrated that the multispectral indices most sensitive to the occurrence of maize leaf spots were those that comprised the red, red edge and near-infrared bands. It was also found that the two thermal parameters that were studied— Normalised canopy temperature and canopy temperature- were essential in identifying whether or not maize leaf spot was present. After 19 days of inoculation, healthy and leaf spot disease-affected maize could be identified using features filtered using the RF algorithm as well as the RF classifier, with precision >0.9 and recall >0.95. However, the accuracy was significantly lower in the early stages of the disease (precision = 0.4, recall = 0.53). It might be useful to monitor maize leaf spot disease in its early stages by using hyperspectral and oblique observations.

Tzutalin, D.L. suggested a unique lightweight neural network called Tassel LFANet to accurately and efficiently detect and count maize tassels in high spatiotemporal picture sequences (Tzutalin et al., 2023). The structure of this network was robust and efficient. The suggested method used a cross-stage fusion strategy to balance the variability of various layers, which enhanced Tassel LFANet's feature learning performance. Tassel LFANet further captured a variety of feature representations by utilising multiple receptive fields. It also included a novel visual channel attention module to increase the adaptability and precision of feature capture and detection. Tassel LFANet outperformed an updated version of lightweight networks in terms of performance, flexibility, and adaptability; it only needed 6.0M parameters and produced an F1 measure value of 94.4% and a mAP.@5 value of 96.8%, as demonstrated by a series of comparative experiments carried out on a newly

created, extremely informative dataset called MrMT. Additionally, the proposed model performed better in counting than the TasselNetV3-Seg† model, which is based on regression and has an R2 of 0.99, an RMSE of 2.68, and a mean absolute error (MAE) of 1.80. The suggested model satisfied the vision system's needs for speed and accuracy in tassel detection in maize. Moreover, the suggested approach was resilient and unaffected by regional variations, offering vital technical assistance for automated counting in the field.

### Research Gap

Regarding the robustness and adaptability of these models to various environmental conditions, there is a significant research gap in the field of maize tassel detection using different machine-learning models and techniques. To detect maize tassels using UAV images, a number of studies have proposed sophisticated models such as SEYOLOX-tiny, RetinaNet, Tassel-YOLO, Yolov8-UAV, MT-Det and others. However, a thorough comparison of these models under various environmental conditions is still lacking. The existing literature focuses on dataset variations, precision rates, and model performance metrics; however, little research has been done on how well these models adapt to various lighting conditions, weather variations, or geographic disparities. For precision agriculture to be used practically, it is imperative to comprehend how these models function in the face of real-world complexity, such as fluctuating brightness, planting densities, and distinct varieties of maize grown in various locations. This disparity impedes the comprehensive comprehension and implementation of maize tassel detection technologies in real-world farming situations, necessitating models with resilient performance in a range of environmental conditions. To enable the development of more flexible and reliable detection systems for real-world agricultural applications, future research should compare and assess how well these models perform in various environmental settings.

### Proposed Methodology

Figure 1 shows a schematic of the procedures needed to put into practice the recommended strategy. This section covers the bottom-hat-top-hat image preprocessing and classification procedure, which uses a STO-based LS-HGNet classifier and STO for hyperparameter tuning.

### Dataset Description

The tasselling, reproductive, and flowering stages are just a few of the phases that make up the maize tassel's growth stages. In the aerial image, the tasselling stage tassel is visible radially. The most noticeable and easiest

to manually label image features in maize fields with higher planting densities are those related to the tasselling stage. As a result, the two dataset data collections for the study were finished at the same time as the tasselling phase. The dataset for this study was provided by the maize field at the Sichuan Agricultural University's modern agricultural development and research base, which is situated in Chengdu, Sichuan Province, China. Using the DJI Mavic drone's onboard camera, RGB video frame data were recorded in June and July 2022 through two aerial surveys carried out at five and ten metres above sea level. The 12-megapixel camera on the drone required manual setting of the filming path. Table 1 provides a detailed breakdown of the video's specifications.

enhanced the dataset through data augmentation in order to raise the suggested model's training effectiveness.

### Data Augmentation

A deep learning technique called "data augmentation" creates new training data from the original dataset, thereby expanding it (Kuma et al., 2023). Data augmentation was applied to make the network learn more features by simulating the real-world environment on the initial dataset used in this study. The experiment employed conventional geometric transformations, such as scaling and rotation, as well as colour transformations, such as contrast enhancement and colour jittering (Jiang et al., 2023). Furthermore, two multi-image fusion techniques were used, namely Mix-up and Mosaic.

The following is the principle of mosaic: Initially, four



**Figure 1. Workflow of the proposed model.**

**Table 1. Conditions for Video Capture**

Date	Weather	Device	Resolution	FPS	Image Sensor
16 June 2022	Sunny	DJI Mavic drone	12 MP	24@1080P	1-inch CMOS
2 July 2022	Sunny	DJI Mavic drone	12 MP	24@1080P	1-inch CMOS

The OpenCV library was used to transform the picture frames from the RGB video frames. A single picture frame was taken every 48 frames, generating 960 unique datasets at a resolution of 1920 x 1080. The study obtained the original dataset using a variety of image preprocessing methods, such as contrast and brightness enhancement. Preprocessing images can bring out their features and help the model's precision and speed improve as the network picks up more precise features. Four employees made boundaries around the pictures of the maize tassels using the graphical annotation of images programme Labellmg, ensuring that the rectangular border enclosed every pixel in the tassel. Tassels made of maize with an occlusion area greater than 90% and visual indistinguishability were not labelled. At last, the research was able to acquire a raw dataset with 960 photos that included 41,232 maize tassels. The study

randomly chosen images are applied various data augmentation techniques, including rotation, scaling as well as colour space conversion. The final pictures are then positioned within a larger image of a predetermined size in the upper-left, lower-left, upper-right, and lower-right positions. Each image's labels receive a mapping that is applied following the transformation that is applied to it. In the end, the big picture is pieced together using the designated coordinates, and the final product is utilised to train the model. Augmenting mosaic data with more diverse training sets, lessening overfitting, and strengthening model robustness can all result in better model performance and overall capacity for generalisation.

In the Mix-up process, two samples are chosen at random from the training set, and their labels are likewise weighted, before they are subjected to a straightforward

random weighted sum. Considering a batch  $x_1$  as samples and a collection of  $y_y$  is the batch and matching labels  $x_2$  is an additional batch of specimens and a batch  $y_2$  is the matching labels.  $\lambda$  is determined by using parameters to calculate the distribution beta-derived mixing coefficient  $\alpha$  and  $\beta$ , and The fundamental Mixup formula is discovered in equations (1), (2) and (3) accordingly.

$$\lambda = \text{Beta}(\alpha, \beta) \quad (1)$$

$$\text{mixed\_batch}_x = \lambda \times \text{batch}_{x1} + (1 - \lambda) \times \text{batch}_{x2} \quad (2)$$

$$\text{mixed\_batch}_y = \lambda \times \text{batch}_{y1} + (1 - \lambda) \times \text{batch}_{y2} \quad (3)$$

The mixed-batch Beta distribution is referred to as "Beta."  $x_x$  speaks of the samples from mixed\_batch, as well as the mixed batch  $y_y$  alludes to the corresponding labels. By generating new training data through linear interpolation between mixup data augmentation with different images and labels, increases the training set's diversity.

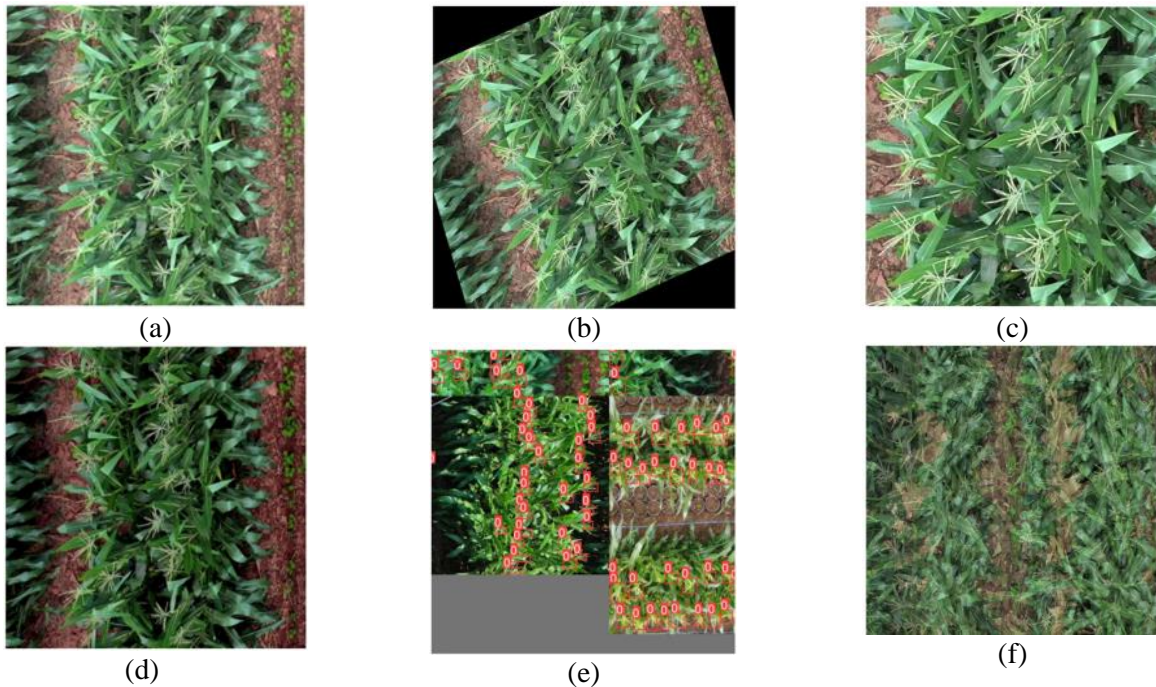
8:1:1 ratio. Figure 2 displays the results of the augmentation of pertinent data.

### Preprocessing using Bottom-Hat–Top-Hat method

First, to evaluate the influence of ambient sound, the bottom-hat top-hat (Bhutto et al., 2022) method is applied to each maize image. The bottom-hat transformation, also known as the white top-hat, is performed by taking the difference between the original image and its morphological opening. It enhances small, bright regions and can be used to highlight details in an image that are smaller than the structuring element used in the opening operation. It is widely acknowledged that images exhibit discrepancies in the intensity of background pixels due to non-uniform illumination, whereby grayscale pixels possess a lower intensity than background pixels. As a consequence, the primary objective is to eliminate the fluctuation in ambient illumination; this can be achieved by reducing the volume. The noise level is calculated using Equation (4), which is implemented via the bottom-hat operation.

$$W_b(f) = (f \cdot b) - f \quad (4)$$

The bullet sign represents the closing operation  $W_b(f)$



**Figure 2. A few outcomes of data augmentation techniques (a) original image, (b) rotation, (c) equal scaling, (d) color dithering, (e) mosaic and (f) mix-up.**

Through the use of offline augmentation, 1848 images were added to the dataset. Using an 8:1:1 ratio, the dataset was randomly divided into training, testing, and validation sets for the study. Figure 2 displays the results of the augmentation of pertinent data. Through the use of offline augmentation, 1848 images were added to the dataset. There were three sets of the dataset: training, testing, and validation at random by the study using an

on image  $f$ . Equation (4) facilitates the noise effect observation. Enhancing the images' varying contrast is the next step, which is calculated using the top-hat operation as indicated by equation (5).

$$W_w(f) = f - (f \circ b) \quad (5)$$

The circle, in this instance, stands for the opening procedure  $W_w(f)$  on image  $f$ . With equation (5), background noise is removed from an image by

deducting the upper hat from the bottom hat, producing an improved image. Following the image's noise reduction, classification is carried out using

### Classification using LS-HGNet

According to several studies, the network is lighter and performs better, thanks to the encoder-decoder architecture. With its ability to learn more complex features through the stacking of modules, The network of hourglasses employed in the study provides a fruitful solution to issues with maize tassel detection. A decoder reassembles features after increasing the image resolution, whereas an encoder lowers the image resolution to extract features first. The purpose of the skip connection in an hourglass network is to allow the decoder to properly restore features by connecting the encoder function to it. Recent research indicates that removing characteristics is an even more important step than just putting them back (Sun et al., 2019). An hourglass network's design allows data from a prior stack to be entered (n-1) and reflects what's in the stack right now (n), along with the results of the preceding stack (n-1) through the skip link. This structure's following stack only replicates relatively high-level elements that the decoder was capable of reconstructing. The suggested study improves feature extraction performance with the least modification necessary because the study's objective is to address this problem and lighten the network simultaneously.

Simple parallel skip connections are added to the next stack to transfer a feature extracted by the encoder. There isn't much more computing involved in this. The suggested structure transfers features to succeeding stack encoders, enhancing the encoder's extraction performance. Compared to the original hourglass network, this structure performs better. While the overall network size stays relatively constant, increased skip connections in an architecture lead to better encoder performance and improved performance.

### Residual Block Design

#### a) Dilated Convolution

To enable the network to learn to identify the characteristics of the entire plant, it is crucial to expand the receptive field in maize tassel detection. However, to broaden the field of receptivity, increasing the kernel size also increases the computational cost. Since the proposed study's objective was to create an effective hourglass network, dilated convolution was used to create a residual block. Equation (6) displays how many parameters are in the standard convolution, and the kernel size is  $K$ , The size of the input channel is  $C$ , and the size of the output channel is  $M$ . In the event that both output and input sizes

match  $H \times W$ , Equation (7) illustrates the necessary computational cost:

$$\# \text{ param} = K^2CM \quad (6)$$

$$\text{Computational Cost} = K^2CMHW \quad (7)$$

The quantity of parameters and the expense of computing the dilated convolution are the same facets of the conventional convolution if the kernel size remains fixed; the dilation size determines how wide the receptive field is  $D$ . In cases where Convolutions that dilate have a  $3 \times 3$  kernel size  $D_1 = 2$ , the computational cost and kernel size are identical to those of the  $3 \times 3$  typical convolution, however, the receiving field is identical to that of the  $5 \times 5$  standard convolution. Furthermore, as demonstrated when  $D_1 = 2$  and  $D_2 = 3$ , Because dilated convolution has no internal padding, its computational cost is marginally less than that of standard convolution with an equivalent kernel size.  $D = 1$  is equivalent to the conventional convolution.  $D = 2, 3$  are computed using the kernel's zero padding.

#### b) Depth-wise Separable Convolution

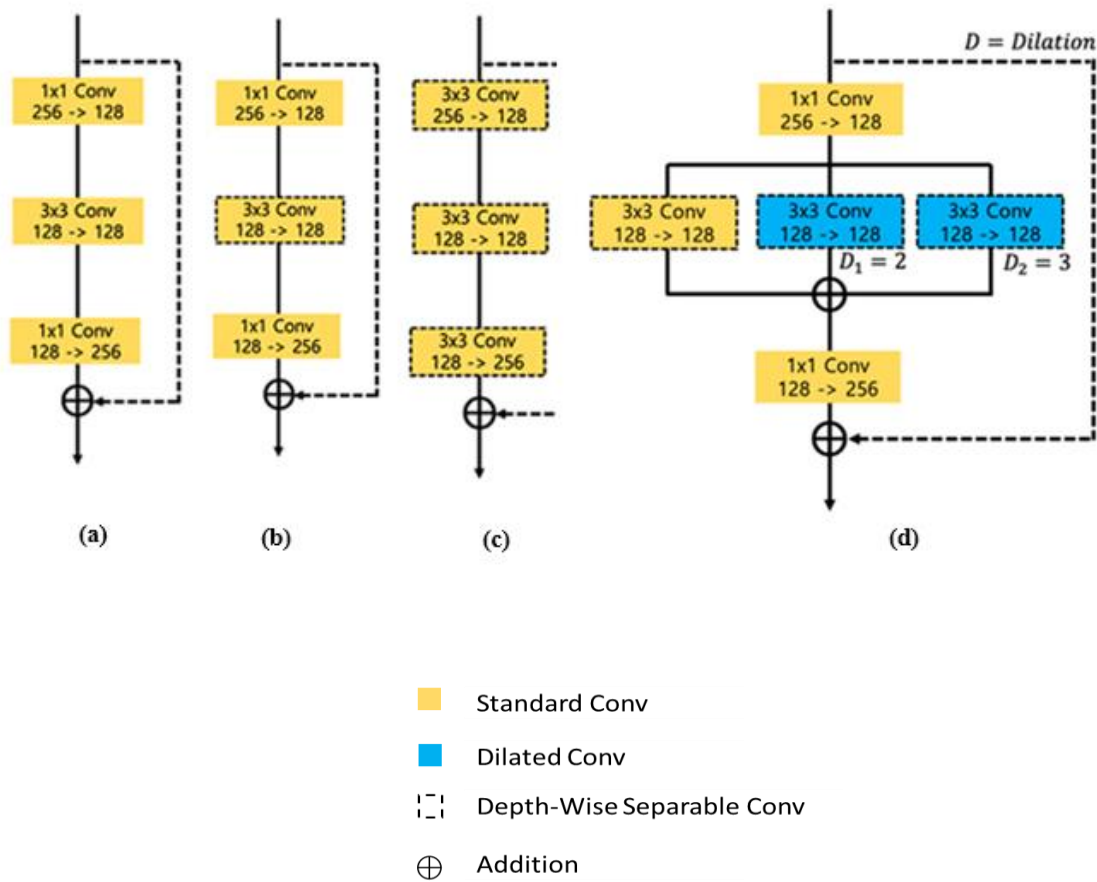
The depth-wise separable convolution was employed in the study, with dilated convolution serving as the residual block. Pointwise depth-wise separable convolution ( $1 \times 1$ ) convolution following depthwise convolution, with a different kernel for every channel. Although the amount of parameters was greatly decreased and the rate of computation was increased, this method performs worse than standard convolution. The investigation interpolated the diminished efficiency resulting from dilated convolution using depth-size separable convolutions to create a novel residual block.

### Proposed LS-HGNet

An hourglass module was produced by the initial stacked hourglass network utilizing a block of residual activation (Figure 3a). Residual blocks that have been preactivated have the structure [ReLU  $\rightarrow$  Batch Normalization  $\rightarrow$  Convolution], whereas conventional residual blocks are designed as [Convolution  $\rightarrow$  Batch Normalization  $\rightarrow$  ReLU]. This arrangement increases training speed and is useful for creating deep networks. Still, residual blocks were first intended for tasks involving object detection or image classification, where learning local features is crucial and where there is not much of a convolutional receptive field. Moreover, even though the leftover block featuring a bottleneck layout is employed in the deep network architecture to lower how many parameters and how much computing work goes into building a stacked multistage network, like an hourglass network, it is still large. As a result, a leftover block featuring an original structure is required in order

to decrease the network's size and increase the receptive field, hence improving maize tassel detection

has a structure like [ReLU→BatchNormalization→Depthwise Convolution



**Figure 3. (a) The vanilla hourglass network's pre-activation residual block. (b) Structure in which a depth-wise separable convolution is applied to the 3x3 convolution layer of (a). (c) Structure wherein (b)'s 1x1 layer is converted to a 3x3 layer. (d) The suggested multi-dilated light residual structure.**

performance.

Experiments were conducted on residual blocks with various structures in this study in order to develop a residual block with a creative layout that can address the problems mentioned above. In Figure 3b, a depth-wise separable convolution has been inserted into the preactivated residual block's middle layer. The study conducted experiments using this residual block to examine how the size and efficacy were impacted by the depth-wise separable convolution of the network in the detection of maize tassels. Changing with bottlenecked preactivated residual blocks made no sense, the layer to depth-separable convolution in sequence to minimize the number of parameters with 1 × 1 convolutions in the first and last layers. The performance is greatly decreased when between a 1×1 convolution and a depthwise convolution, there is a nonlinear function. Because of this, In this work, every depth-wise separable convolution

→ 1 × 1 Convolution] that between the depth-wise convolution doesn't employ a function for activation and 1 × 1 convolution.

The new module presented in Figure 3c was designed as part of the proposed work to assess the impact of the residual block bottleneck structure when using a depth-wise separable convolution. A modified version of Figure 3b is shown in Figure 3c, where the depth-wise separable convolutions of [256 → 128, 3 × 3] and [128 → 256, 3 × 3] were substituted for the first layer's conventional convolutions [256 → 128, 1 × 1] and the last layer [128 → 256, 1 × 1]. The suggested multidilated light residual block in Figure 3d was created using the remaining block of a fresh design to enhance performance and decrease the number of parameters. It displays the detailed structure of the suggested residual block in Table 2 below.



Using the multidilated light residual block in this study, the stack hourglass network was significantly lighter, and the immutability of scale was increased by employing multidilated convolution to increase the receptive field, which improved maize tassel detection performance).

**Hyperparameter Tuning using the STOA Algorithm**

Inspired by metallurgy, simulated annealing gradually decreases a temperature parameter to explore the hyperparameter space. Following the classification procedure, the suggested classifier's hyperparameters should be fine-tuned for impressive maize tassel detection results. The sooty tern optimisation algorithm (STOA), described below, inspires the hyperparameter tuning process (Javeed et al., 2023). It is primarily used to divide up cancerous nodules by determining the best characteristics, which contributes to improving the diagnosis' accuracy. The phases of attack and migration, which correspond to discovery and exploitation, are used to implement STOA towards feature optimisation. The STOA for feature optimisation is implemented by Algorithm 1.

**a) Migration (Exploration)**

When migrating, a ST must meet the following requirements.

**Collision evasion:**  $M_{SA}$  provides the search agent's (SA's) new location in equation (8), which deals with preventing collisions between nearby SAs (STs).

$$\vec{C}_{ST}^L = M_{SA} \times \vec{P}_{ST}^L \quad (8)$$

where,

$\vec{C}_{ST}^L$  – SA's location that is independent of other SAs;

$\vec{P}_{ST}^L$  – Present location of SA;

$M_{SA}$  – SA movement in the presumptive search area.

$$M_{SA} = C_{fac} - \left( i \times \frac{C_{fac}}{MaxIter} \right) \quad (9)$$

where in equation (9),

$i$  – Present iteration,  $i = 0,1,2, \dots Max Iter$ ;

$C_{fac}$  – regulating factor (configured at 2) that alters ' $M_{SA}$ ' declined linearly to zero.

Proceed towards your best neighbor: Following the resolution of a collision, SAs take the route of the neighbor who offers the greatest advantage.

$$\vec{M}_{ST}^L = C_{Best} \times (\vec{P}_{BST}^L(i) - \vec{P}_{ST}^L(i)) \quad (10)$$

where in equation (10),

$\vec{M}_{ST}^L$ -Various places in SA ( $\vec{P}_{ST}^L$ ) towards the fittest and best SA ( $\vec{P}_{BST}^L$ );

$C_{Best}$  - Using a random variable to enhance exploration.

$$C_{Best} = 0.5 \times Ran \quad (11)$$

where Ran is a random number between 0 and 1 as shown in equation (11).

Last but not least, SA or ST updates its location in accordance with the best SA.

$$\vec{G}_{ST}^L = \vec{C}_{ST}^L + \vec{M}_{ST}^L \quad (12)$$

where in equation (12),

$\vec{G}_{ST}^L$  – gap between the fittest SA and the SA. L

**b) Attacking (Exploitation)**

STs adjust their angle of attack and velocity as they migrate. They use their wings to increase their altitude. When they strike their prey, they fly in spirals.

$$X' = Rad \times Sin(a) \quad (13)$$

$$Y' = Rad \times Cos(a) \quad (14)$$

$$Z' = Rad \times a \quad (15)$$

$$r = u \times e^{kv} \quad (16)$$

whereas mentioned in equations (13) and (14),

Rad- Each spiral turn's radius;

A-Range of  $[0 \leq k \leq 2\pi]$ ;

u, v- Constants that are thought to represent a spiral ' 1

;

e- The foundation of a natural algorithm.

Equations (15)–(17) are utilized to determine the altered position of SA.

$$\vec{P}_{ST}^L(i) = (\vec{G}_{ST}^L \times (X' + Y' + Z')) \times \vec{P}_{BST}^L(i) \quad (17)$$

where  $\vec{P}_{ST}^L(i)$  - adjusts other SAs' locations while maintaining the best possible outcome.

Algorithm 1: STOA
Input Population ( $\vec{P}_{SI}^L(i)$ )
Output Best SA ( $\vec{P}_{BSI}^L(i)$ )
Initialize 'MSA' and ' $C_B$ '
Determine the fitness of every SA
while ( $i < MaxIter$ ) do
for every SA, do
Modify the locations of SAs using Equation (10)
end // for
Update ' $S_A$ ' and ' $C_{Best}$ '
Find the fitness of every SA
Modify ' $\vec{P}_{BSI}^L(i)$ ' if a better solution than the previously perfect one exists $i = i + 1$
return ( $\vec{P}_{BSI}^L(i)$ )
End

**Results and Discussion**

**Experimental Setup**

A desktop workstation with a 3.30 GHz Intel(R) Core i9-7900X CPU and 64 GB of RAM was used for the trials. The computer was running Ubuntu 16.04, a Linux-based operating system. The proposed study used scikit-learn and PyTorch, two well-known deep learning frameworks, to execute the simulations.

## Performance Metrics

ACC is a widely used statistic to evaluate the performance of segmentation models. Equation (33), when applied to all samples, yields the percentage of correctly recognised samples.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (33)$$

Equation (34), which introduces the PR rate, assesses how well a model can predict positive samples among those that it considers to be positive.

$$Precision = \frac{TP}{TP+FP} \quad (34)$$

Equation (35), likewise referred to as the true positive rate, or ST, assesses how well a prediction model detects actual positive data. The true positive to total true positive and negative result ratio is used to calculate it.

$$Recall = \frac{TP}{TP+FN} \quad (35)$$

Equation (36), which measures ACC by averaging PR and RC, is weighted and defines the F1 score. It provides an evaluation of the test's ability to distinguish between favourable and unfavourable outcomes.

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \quad (36)$$

**Table 2. Classification analysis of various models**

Models	Accuracy	Precision	Recall	F1-Score
ResNet	92.9	92.4	92.5	92.2
ImageNet	93.7	93.3	93.5	93.4
GoogleNet	94.8	94.4	94.3	94.1
PolyNet	95.8	95.4	95.7	95.2
Proposed LS-HGNet	97.5	97.3	97.6	97.1

The accuracy, precision, recall, and F1-score performance metrics of several image recognition models are displayed in table 2 and figure 4. Among the models compared is ResNet, which has 92.9% accuracy, 92.4% precision, 92.5% recall, and 92.2% F1-score. With 93.7% accuracy, 93.3% precision, 93.5% recall, and 93.4% F1-score, ImageNet comes in second. With scores of 94.8% accuracy, 94.4% precision, 94.3% recall, and a 94.1% F1-score, GoogleNet demonstrates superior metrics. With 95.8% accuracy, 95.4% precision, 95.7% recall, and an F1-score of 95.2%, PolyNet performs better than the prior models. Out of all the models, the suggested LS-HGNet performs the best, with 97.5% accuracy, 97.3% precision, 97.6% recall, and 97.1% F1-score. The LS-HGNet model has the best overall performance across all assessed metrics in this comparison, which shows the gradual advances in image recognition models.

**Table 3. Accuracy analysis with STO**

Models	without STO	With STO
ResNet	92.9	95.4
ImageNet	92.7	94.3
GoogleNet	93.8	95.6
PolyNet	94.8	95.8
Proposed LS-HGNet	95.5	97.7

In Table 3 and figure 5, an accuracy analysis comparing various models both without and with the incorporation of STO is presented. The models evaluated include ResNet, ImageNet, GoogleNet, PolyNet, and the proposed LS-HGNet. Without the integration of STO, ResNet achieves an accuracy of 92.9%, ImageNet at 92.7%, GoogleNet at 93.8%, PolyNet at 94.8%, and the Proposed LS-HGNet at 95.5%. Upon introducing STO into the models, there is a noticeable improvement in accuracy across the board. ResNet's accuracy increases to 95.4%, ImageNet to 94.3%, GoogleNet to 95.6%, PolyNet to 95.8%, and the Proposed LS-HGNet experiences a substantial boost, reaching an impressive accuracy of 97.7%. These results highlight the positive impact of STO on enhancing the performance of these models, with the LS-HGNet particularly demonstrating its efficacy in leveraging the spatial transformation operator for improved accuracy.

## Conclusion

This study used deep learning to identify and count maize tassels. A high-quality data-set of aerial photos of maize tassels while in the tasselling stage was first produced using pre-processed aerial video footage taken by unmanned aerial vehicles. The study presents the STOA-based LS-HGNet model as a solution to the problems of poor tassel detection accuracy and sluggish inference speeds. This work applies the top-hat- bottom-hat preprocessing method to eliminate noise and uneven lighting from maize photos. Next, a light-weight stacked hourglass network is suggested for the classification process in order to detect maize tassels. The proposed LS-HGNet classifier uses the STOA for hyperparameter tuning, which helps to increase the accuracy of tassel detection. A range of tests were carried out to objectively assess the suggested methods' performance, and the results, which showed an accuracy rate of 98.7%, verified that the suggested method represents a significant advancement in the detection of maize tassels. Additionally, expanding the dataset to incorporate additional varieties, growth stages, and environmental conditions of maize would enhance the model's adaptability and future applicability. Proposed LS-HGNet

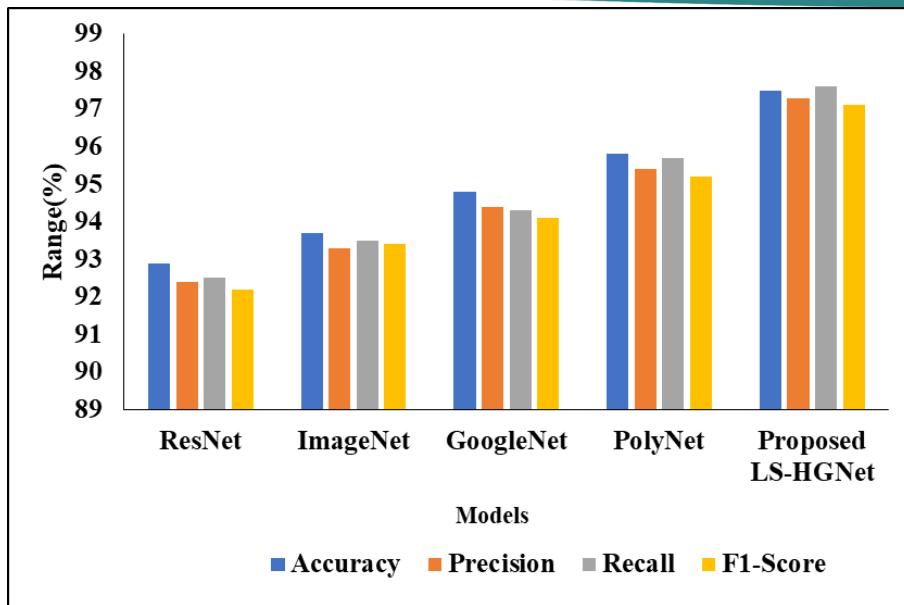


Figure 4. Classification analysis on existing models with proposed model.

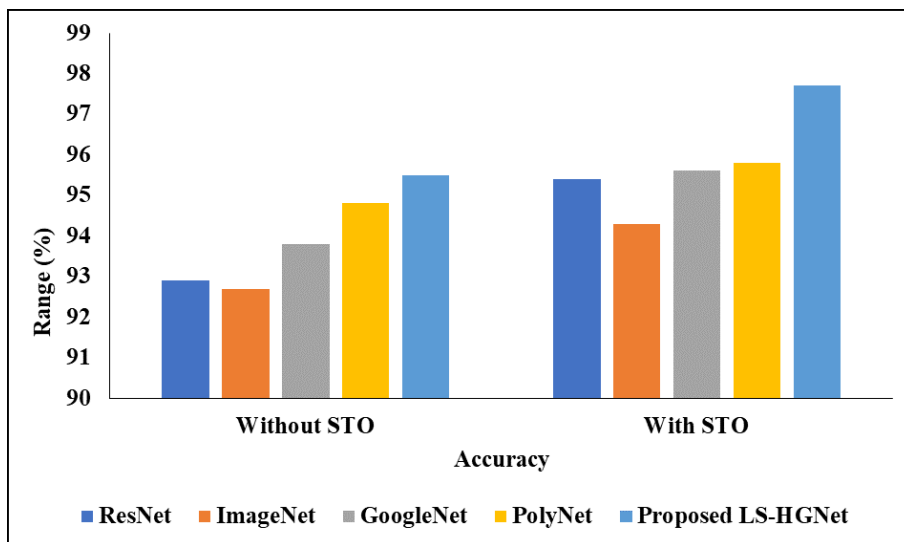


Figure 5. Accuracy analysis with STO.

experiences a substantial boost, reaching an impressive accuracy of 97.7%. These results highlight the positive impact of STO on enhancing the performance of these models, with the LS-HGNet particularly demonstrating its efficacy in leveraging the spatial transformation operator for improved accuracy.

#### Acknowledgements

The authors would like to acknowledge Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India, for providing the various facilities to carry out this research.

#### Conflict of Interest

The authors declare no conflict of interest.

#### References

- Aishwarya, N., Praveena, N. G., Priyanka, S., & Pramod, J. (2023). Smart farming for detection and identification of tomato plant diseases using light weight deep neural network. *Multimedia Tools and Applications*, 82(12), 18799-18810. <https://doi.org/10.1007/s11042-022-1312-5>
- Akkem, Y., Biswas, S. K., & Varanasi, A. (2023). Smart farming using artificial intelligence: A review. *Engineering Applications of Artificial Intelligence*, 120, 105899. <https://doi.org/10.1016/j.engappai.2022.105899>
- Bhutto, J. A., Tian, L., Du, Q., Sun, Z., Yu, L., & Tahir, M. F. (2022). CT and MRI medical image fusion using noise-removal and contrast enhancement

- scheme with convolutional neural network. *Entropy*, 24(3), 393. <https://doi.org/10.3390/e24030393>
- Dawn, N., Ghosh, T., Ghosh, S., Saha, A., Mukherjee, P., Sarkar, S., Guha, S., & Sanyal, T. (2023). Implementation of Artificial Intelligence, Machine Learning, and Internet of Things (IoT) in revolutionizing Agriculture: A review on recent trends and challenges. *Int. J. Exp. Res. Rev.*, 30, 190-218. <https://doi.org/10.52756/ijerr.2023.v30.018>
- Dhruva, A. D., Prasad, B., Kamepalli, S., & Kuniseti, S. (2023). An efficient mechanism using IoT and wireless communication for smart farming. *Materials Today: Proceedings*, 80, 3691-3696. <https://doi.org/10.1016/j.matpr.2022.12.688>
- El-Ghamry, A., Darwish, A., & Hassanien, A. E. (2023). An optimized CNN-based intrusion detection system for reducing risks in smart farming. *Internet of Things*, 22, 100709. <https://doi.org/10.1016/j.iot.2022.100709>
- El-Ghamry, A., Darwish, A., & Hassanien, A. E. (2023). An optimized CNN-based intrusion detection system for reducing risks in smart farming. *Internet of Things*, 22, 100709. <https://doi.org/10.1016/j.iot.2022.100709>
- Javeed, A., Dallora, A. L., Berglund, J. S., Ali, A., Ali, L., & Anderberg, P. (2023). Machine Learning for Dementia Prediction: A Systematic Review and Future Research Directions. *Journal of Medical Systems*, 47(17). <https://doi.org/10.1007/s10916-022-01753-8>
- Jia, X., Yin, D., Bai, Y., Yu, X., Song, Y., Cheng, M., ... & Jin, X. (2023). Monitoring Maize Leaf Spot Disease Using Multi-Source UAV Imagery. *Drones*, 7(11), 650. <https://doi.org/10.3390/drones7110650>
- Jiang, W.; Zhang, K.; Wang, N.; Yu, M. MeshCut data augmentation for deep learning in computer vision. *PLOS ONE*, 15(1), e0243613. <https://doi.org/10.1371/journal.pone.0243613>
- Karunathilake, E. M. B. M., Le, A. T., Heo, S., Chung, Y. S., & Mansoor, S. (2023). The path to smart farming: Innovations and opportunities in precision agriculture. *Agriculture*, 13(8), 1593. <https://doi.org/10.3390/agriculture13081593>
- Kumar, T., Turab, M., Raj, K., Mileo, A., Brennan, R., & Bendeche, M. (2023). Advanced Data Augmentation Approaches: A Comprehensive Survey and Future directions. arXiv, 2023, arXiv:2301.02830. <https://doi.org/arXiv:2301.02830>
- Lachgar, M., Hrimech, H., & Kartit, A. (2023). Unmanned aerial vehicle-based applications in smart farming: A systematic review. *International Journal of Advanced Computer Science and Applications*, 14(6). <https://doi.org/10.14569/IJACSA.2023.0140616>
- Lu, D., Ye, J., Wang, Y., & Yu, Z. (2023). Plant Detection and Counting: Enhancing Precision Agriculture in UAV and General Scenes. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2023.3147529>
- Meng, K., Wu, Q., Xu, J., Chen, W., Feng, Z., Schober, R., & Swindlehurst, A. L. (2023). UAV-enabled integrated sensing and communication: Opportunities and challenges. *IEEE Wireless Communications*. <https://doi.org/10.1109/MWC.2023.3146074>
- Pu, H., Chen, X., Yang, Y., Tang, R., Luo, J., Wang, Y., & Mu, J. (2023). Tassel-YOLO: A New High-Precision and Real-Time Method for Maize Tassel Detection and Counting Based on UAV Aerial Images. *Drones*, 7(8), 492. <https://doi.org/10.3390/drones7080492>
- Salehi, A. W., Khan, S., Gupta, G., Alabdullah, B. I., Almjally, A., Alsolai, H., ... & Mellit, A. (2023). A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope. *Sustainability*, 15(7), 5930. <https://doi.org/10.3390/su15075930>
- Song, C. Y., Zhang, F., Li, J. S., Xie, J. Y., Chen, Y. A. N. G., Hang, Z. H. O. U., & Zhang, J. X. (2023). Detection of maize tassels for UAV remote sensing image with an improved YOLOX model. *Journal of Integrative Agriculture*, 22(6), 1671-1683. [https://doi.org/10.1016/S2095-3119\(23\)66481-7](https://doi.org/10.1016/S2095-3119(23)66481-7)
- Sun, K., Xiao, B., Liu, D., & Wang, J. (2019). Deep High-Resolution Representation Learning for Human Pose Estimation. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5686–5696. Long Beach, CA, USA.
- Thirumalraj, A., Anusuya, V. S., & Manjunatha, B. (2024). Detection of Ephemeral Sand River Flow Using Hybrid Sandpiper Optimization-Based CNN Model. In A. Kumar, A. Srivastav, A. Dubey, V. Dutt, & N. Vyas (Eds.), *Innovations in Machine Learning and IoT for Water Management*. pp. 195-214. IGI Global. <https://doi.org/10.4018/979-8-3693-1194-3.ch010>
- Thorat, T., Patle, B. K., & Kashyap, S. K. (2023). Intelligent insecticide and fertilizer recommendation system based on TPF-CNN for smart farming. *Smart Agricultural Technology*, 3, 100114. <https://doi.org/10.1016/j.smag.2022.100114>

- Tzutalin, D.L. Git Code. 2015. Available online: <https://github.com/tzutalin/labelImg> (accessed on 10 April 2023).
- Wang, B., Yang, G., Yang, H., Gu, J., Xu, S., Zhao, D., & Xu, B. (2023). Multiscale Maize Tassel Identification Based on Improved RetinaNet Model and UAV Images. *Remote Sensing*, 15(10), 2530. <https://doi.org/10.3390/rs15102530>
- Yuan, J., Zhou, F., Guo, Z., Li, X., & Yu, H. (2023). HCformer: hybrid CNN-transformer for LDCT image denoising. *Journal of Digital Imaging*, 36(5), 2290-2305. <https://doi.org/10.1007/s10278-022-00657-w>
- Zeng, F., Ding, Z., Song, Q., Qiu, G., Liu, Y., & Yue, X. (2023). MT-Det: A novel fast object detector of maize tassel from high-resolution imagery using single level feature. *Computers and Electronics in Agriculture*, 214, 108305. <https://doi.org/10.1016/j.compag.2022.108305>
- Zhang, Q., Xiao, J., Tian, C., Chun-Wei Lin, J., & Zhang, S. (2023). A robust deformed convolutional neural network (CNN) for image denoising. *CAAI Transactions on Intelligence Technology*, 8(2), 331-342. <https://doi.org/10.1049/caai.2022.0467>

#### How to cite this Article:

V. Gokula Krishnan<sup>1\*</sup>, B. Vikranth<sup>2</sup>, M. Sumithra<sup>3</sup>, B. Prathusha Laxmi<sup>4</sup> and B. Shyamala Gowri<sup>5</sup> (2024). Smart Farming with Sooty Tern Optimization based LS-HGNet Classification Model. *International Journal of Experimental Research and Review*, 37(Spl.), 96-108.

**DOI :** <https://doi.org/10.52756/ijerr.2024.v37spl.008>



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.