

**AN IMPROVED SWIN TRANSFORMER BASED STACKED
CAPSULE AUTOENCODER MODEL FOR WEED
DETECTION IN RICE FIELDS**

Sourabh Sagar¹, Mahantesh N.Birje²

¹Research Associate, Computer Science and Engineering Department,
Visvesvaraya Technological University, Belagavi -590018, Karnataka ,
India.

²Professor,
Computer Science and Engineering Department, Visvesvaraya
Technological University, Belagavi -590018, Karnataka , India.

Abstract

Weeds that grow on cultivated land compete with crop plants for essential resources such as nutrients, water, space, and sunlight. Consequently, controlling weeds is crucial for achieving successful agricultural production. One primary cause of yield decline is weed competition, which hinders the growth of paddy by competing with cultivated plants for vital resources like moisture and solar energy. Conventional methods for detecting weeds, such as UAVs and satellite imagery, are expensive to implement, while manual detection is often inefficient. To address these limitations, an enhanced Swin Transformer model for weed detection in rice fields has been proposed. Initially, the inputs are sourced from the Mendeley repository, and preprocessing is performed using image resizing and Modified Gaussian Wiener filtering (MGWF) techniques to ensure uniformity and reduce unwanted noise. The processed images are then fed into the proposed Improved Sequential Head Swin Transformer with Stacked Capsule Autoencoder (ISH-ST-SCAE) model, which detects weeds by extracting significant features. Furthermore, the model's parameters are refined using the opposition-based Orangutan Optimization Approach (O3A), enhancing its training performance. As a result, this approach efficiently detects weeds in rice fields, contributing to increased crop yields. Various evaluation metrics, including correctness, specificity, Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the Receiver Operating Characteristic (ROC) curve, are employed to assess the effectiveness of this framework. The proposed system achieved an accuracy of 99.45%, precision of 99.50%, recall of 99.69%, F1-score of 99.59%, specificity of 99.94%, MAE of 0.0138, and RMSE of 0.1964.

Keywords: Improved swin transformer, Stacked capsule autoencoder, Orangutan optimization, Weed detection.

1. Introduction

The most important staple crop in the world, rice, is essential to ensuring global nutritional security. Its worldwide distribution has positioned rice as the primary dietary source for humans. In the process of rice farming, unwanted plants are a key element influencing rice yield [1]. To enhance the weed control procedure, cultivators unintentionally shifted to chemical methods. However, this frequently results in other problems, such as decreased soil fertility [2]. Unfavorable areas, such as lawns, farmlands, and other growing regions, are ideal for the rapid growth of undesirable plants. They absorb vital resources like solar energy, water, and ground moisture from cultivated crops [3]. In India, weeds are typically managed using traditional methods to ignore pests such as hand weeding, mechanical tools, and the application of chemical herbicides [4]. The process of manually weeding entails uprooting undesirable plants. This process is less preferable because it takes a lot of time and effort. Mechanical weed eradication uses machinery such as a rotary cultivator; however, it is ineffective in rice fields due to the difficult and uneven terrain. In addition to, herbicides also have their ill effects on the environment [5]. However, many individuals understand how to cultivate paddy, but they lack effective ways for weed eradication in rice fields [6].

This work proposes implementing deep learning (DL) techniques for weed detection in rice fields in order to overcome these problems. DL is essential for distinguishing between rice and weeds. So that the rice field's weeds can be easily removed, this technology minimizes human effort and improves crop management [7]. Since DL has been effectively applied across multiple domains, it has lately expanded into the farming industry too [8]. It is an advanced and modern method for data analysis and image processing that has shown encouraging outcomes and substantial advantages in a range of applications [9]. DL is a subset of ML; it expands traditional ML by introducing increased depth and complexity, employing deep neural architectures or multiple interconnected layers of nodes to gradually derive and refine features from the dataset [10]. The dataset is obtained from the rice field images, and it was taken from a mobile phone at a 360-degree angle above a rice field [11]. The RGB dataset has been produced, and the purpose of this dataset is to record a variety of weeds. This dataset's diversity guarantees that the machine learning models trained on it can effectively identify and differentiate between undesirable and advantageous plant species [12]. DL has made detecting weeds in rice fields simpler due to its capability to automatically derive features from images can learn directly from large volumes of annotated weed images [13]. DL has data in an end-to-end manner, automatically extracting the visual features required for weed classification without relying on manual feature design or selection [14]. A DL approach for identifying weeds in rice field images, delivering precise real time recognition and being extensively adopted due to its low cost hardware requirements [15].

The objective is to design a rapid, automated framework utilizing computer vision and DL to identify weeds in rice fields [16]. CNNs were widely recognized as a unique and crucial part of DL because of their tremendous capabilities [17]. CNNs are a type of deep

learning technique used to identify undesirable plants in rice fields [18]. As a result, the laborious and time-consuming process of manually visualizing crops should be eliminated [19]. Additionally, compared to other approaches, it offers superior performance. But the result can be enhanced with more datasets. The results show that this model can effectively identify weeds in rice fields and assist in making well-informed crop management decisions [20].

1.1 Motivation

Food security and crop production face serious threats from the spread of weeds in rice fields; therefore, prompt and accurate weed detection is crucial for effective management. Conventional machine learning techniques often depend on handcrafted features and exhibit limited generalization across various field conditions. Conversely, traditional manual weed detection methods are time-consuming, labor-intensive, and prone to human errors. Field weed detection through computer vision typically relies on conventional image processing. In these traditional methods, features such as color, texture, and shape are manually extracted and paired with classifiers based on machine learning approaches, which tend to be less adaptable, have difficulty in complex backgrounds, and require extensive feature engineering, consequently limiting their accuracy in diverse field conditions. Recent advancements in deep learning and artificial intelligence have started to overcome these challenges by facilitating automated feature extraction and enhancing model adaptability to different environments. This transition not only improves the accuracy of weed detection but also simplifies the process, making it more efficient for agricultural practices. Consequently, farmers can depend on these advanced systems to make timely decisions regarding weed management, ultimately resulting in higher crop yields and decreased reliance on chemical herbicides. This technological advancement marks a significant move toward sustainable agriculture, where precision and efficiency are vital.

1.2 Objectives

- ❖ To enhance the quality of images and ensure uniformity, image denoising and resizing are performed in the initial stage.
- ❖ To propose a new modified swin transformer with a capsule autoencoder for extracting important features and detecting weeds accurately.
- ❖ To improve the model's efficiency and avoid complexity issues, the parameters of the detection model are fine-tuned by applying a new meta-heuristic optimization approach.

This proposed methodology is structured as follows: module 1 illustrates the introduction, motivation, and objectives, module 2 summarizes the recent studies of weed detection in rice fields using the swin transformer model. The proposed methodology is described in module 3. Further, module 4 discusses the performance evaluation and outcomes of the proposed methodology. Module 5 summarizes the entire work in the form of a conclusion.

2. Related works

Chen et al. [21] introduced a PIS-NET: Optimized, minimally supervised per-object segmentation framework utilizing labeled points for unwanted plant detection in paddy fields. It used the PIS-NET method for analyzing the weed detection. This research annotates a small set of randomly selected points for each occurrence and utilizes these point annotations as weak guidance to efficiently accomplish weed segmentation in rice paddies. This method addresses the challenges of labeling weed data and improves upon the low accuracy seen in weed detection algorithms. The method was evaluated on a self-constructed rice image dataset. This method's drawback is that it cannot be used on a mobile platform.

Duong et al. [22] developed a system for automated weed detection, leveraging the synergy between EfficientNet and transfer learning to improve prediction. This research introduces a weed classification method that employs EfficientNet-B4 and EfficientNet-Lite4 as primary architectures. The method's shortcomings include its inability to be implemented in robotics and the fact that Android phones do not support this network topology.

Habib et al. [23] introduced an innovative segmentation method, DWUNet, aimed at improving weed control in agriculture. The model delivers exceptional accuracy, offering an optimal solution for real-time use. Visual inspection of results validates DWUNet's capability to precisely identify weeds, correct detection errors, and generalize effectively to unseen data and limitation of this study is acquiring the dataset.

Liao et al. [24] introduced SC-Net, a strip convolutional network designed for segmenting rice seedlings and weeds in paddy fields. The architecture employs U-Net as its backbone. The approach was tested on both a custom rice seedling dataset and publicly available agricultural datasets, achieving mIoU scores of 87.48% and 89% respectively.

Guo et al. [25] present a UAV-based approach for weed detection in rice fields using image segmentation with a CNN-Transformer framework. CTFFNet integrates CNN and Transformer to segment weeds in remote sensing images. It incorporates two feature extraction modules, one using a transformer for global semantic information and another using a Residual CNN for local visual details. Table 1 illustrates the Prior research analysis of weed detection in rice fields.

Table 1: Prior research analysis of weed detection in rice fields

References	Methods	Performance	Limitations
Chen et al. [21]	PIS-NET	This system attained an accuracy of 88%.	It can't be implemented in a mobile platform.
Duong et al. [22]	EfficientNet and transfer learning	This system attained an accuracy of 90% respectively.	It can't deploy in robotics, and this network model is not used by Android phones.
Habib et al. [23]	DWUNet	This system attained an	This system cannot have a

		accuracy of 92.34%.	proper dataset for classification.
Liao et al. [24]	SC-Net	This system attained an accuracy of 87.44%.	Less suitable for real-time, low cost deployment,
Guo et al. [25]	Transformer and Residual CNN	This system attained an accuracy of 91.42%.	Less suitable for real time, low cost deployment.

Sugarcane is a widely used crop in agriculture, serving as a raw material for the production of sugar, ethanol, and bagasse. Machines are commonly employed for sugarcane harvesting. Currently, billets are utilized for cultivating sugarcane, and healthy billets are essential for optimal growth and maximum yield. However, the harvesting equipment can damage the billets, which may result in lower quality and the potential transmission of diseases. To effectively detect damage in sugarcane billets within dynamic environments, innovative solutions are necessary for harvesters. Unfortunately, conventional methods often lack accuracy due to complex backgrounds. This paper introduces a Fractional Deer Hunting Optimization Algorithm-based Deep Maxout Network (FDHOA-based DMN) designed to effectively categorize damage to sugarcane billets. The process begins with a median filter to pre-process the images, followed by using the Pyramid Scene Parsing network (PSPNet) for segmentation. Subsequently, the Deep Maxout Network (DMN) classifies the damaged sugarcane billets, having been trained with the FDHOA. The DMN is chosen for classification due to its efficient learning capabilities and ability to recognize inherent data properties. The FDHOA is developed by integrating fractional calculus with the deer hunting optimization method. The model's performance is evaluated through accuracy, sensitivity, specificity, and F1-score metrics, resulting in values of 0.940, 0.930, 0.959, and 0.935, respectively [31].

The Internet of Things (IoT) plays a vital role across various industries, including agriculture, due to its extensive and straightforward integration capabilities. IoT technologies enhance agricultural systems, driving growth and efficiency. They assist farmers in optimizing resource usage and making informed decisions through advanced field monitoring methods. However, IoT-based agricultural data is vulnerable to attacks, necessitating a robust authentication scheme. This paper presents the Auth Key_Deep Convolutional Neural Network (Auth Key_DCNN) as a solution for secure data sharing within IoT-enabled agriculture systems. The system randomly initializes the initial parameters for the sensors, the Private Key Generator (PKG), the controller, and the data user entities. Through the use of DCNN, these entities are registered, and a secret key is generated by the PKG. Data protection is achieved by encrypting the information transferred between the controller and the user. Experimental results indicate that the Auth Key_DCNN model demonstrates strong performance, with a computational cost of 142.56, memory usage of 49.5 MB, and a computational time of 1.34 seconds [32].

Innovative IoT services have facilitated the development of complex systems that utilize a variety of technologies. The Internet of Things (IoT) encompasses the integration of networks

to deliver seamless services to users. In regions facing water scarcity, irrigation practices are being enhanced to maintain sugarcane crop production. Efficient water usage in agriculture presents a challenge due to escalating demand and the need for increased output. To address this issue, the paper proposes a technique based on the sunflower-atom optimization and Deep Convolutional Neural Network (SFAO-DeepCNN) for effective water management in sugarcane cultivation. Initially, the Artificial Bee Colony (ABC) technique is employed to gather data within the IoT network. Subsequently, a push-based method and log transformation prepare the data for analysis. The Apriori algorithm is utilized to select significant features, adapting a feature score from a broad array of options, thereby enhancing classifier accuracy. We have developed a new feature score specifically aimed at identifying key features for determining the optimal watering quantity. For water control, DeepCNN is trained using the proposed Sunflower Atom Optimization (SFAO) method. The SFAO-DeepCNN achieved an accuracy of 92%, specificity of 91.2%, and sensitivity of 94.1%, surpassing existing methodologies [33].

With IoT, networks can be integrated to provide seamless services to consumers. Agriculture is essential for the sustained growth of sugarcane in areas with limited water resources. Effective water utilization presents a significant challenge in agriculture due to increasing water demands and high output requirements. This research proposes using the Sunflower Atom Optimization Deep Convolutional Neural Network (DCNN) for optimal water management in sugarcane plantations. In this approach, a push-based strategy and log change are employed to prepare the data for processing. The Apriori algorithm is utilized to select the most valuable features from a large pool of available options. This algorithm enhances classifier accuracy. A new feature scoring method was developed to identify useful traits for estimating water requirements. The water control phase is completed using the proposed Sunflower Atom Optimization (SFAO) algorithm with DeepCNN. The suggested SFAO-Deep CNN achieved 91.2% specificity, 94.1% sensitivity, and 91.2% accuracy, outperforming other techniques [34].

2.1 Problem Statement

CNNs, a recent advancement in deep learning, have demonstrated superior performance; however, they frequently struggle to capture contextual linkages and long-range relationships in complex field images. In order to address these drawbacks, this study proposes a brand-novel Transformer built on a deep learning framework that makes use of the self-attention mechanism to efficiently model global spatial dependencies and improve weed identification precision in a variety of crowded rice field settings. By offering a more reliable, scalable, and automated solution for real-time weed detection, the suggested method seeks to address the drawbacks of current techniques and eventually support precision agriculture and sustainable crop management.

3. Proposed Methodology

The agricultural industry greatly contributes to economic development by supplying food, generating essential raw materials for various industries, and creating employment for a large portion of the population. However, weeds are a common issue in farmlands worldwide. They rival important crops for light, moisture, and soil resources while also spreading infections while attracting destructive insects and pests, ultimately leading to reduced crop yields. This proposed methodology developed a novel deep learning framework to detect the weeds in rice fields. Figure 1 illustrates the workflow of weed detection in rice fields. Initially, the data is collected from the publicly available dataset, and then the gathered data is pre-processed through image resizing and MGWF methods to remove the unwanted noises. Further, the pre-processed data is given to the ISH-ST-SCAE model for extracting the features to detect the weeds in rice fields. Training of the model is improved by the optimization technique called Opposition-based Orangutan Optimization (O^3 or OOO).

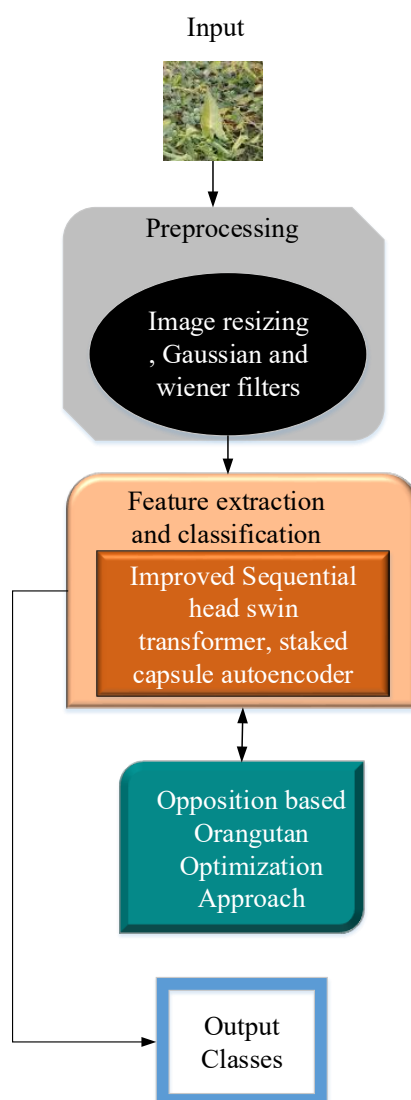


Figure 1: Workflow of weed detection in rice fields

3.1 Pre-processing

Initially, the data is resized to remove partial pixels and ensure all the images are in a suitable size to feed as input into the models. A Gaussian Wiener filter is applied to resized images to enhance the quality of the image and to denoising the images. A Gaussian filter is widely used for image smoothing because of its ability to reduce noise while preserving important details. The Gaussian filter reduces the contribution of pixels at wider distances while giving priority to core pixels by giving them more influence than the mean filter, which treats all nearby pixels equally. This weighting is based on the standard deviation σ , which also helps to define the size of the neighbourhood filters. Additionally, a Gaussian filter produces smooth transitions and reduces harsh edges, making the image appear more natural. In contrast, mean and median filters do not consider pixel distance from the centre, which can cause unwanted visual distortions when used on high detail images with larger filters. The mathematical representation of the Gaussian filter reflects this distance based approach.

$$N_{R,C} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{R^2 + C^2}{2\sigma^2}\right) \quad (1)$$

here, R refers to rows of that kernel pixel, C is the columns from the centre. Gaussian filters are cut off around 4π , as the values beyond this point are nearly zero.

Additionally, a Wiener filter is integrated with the Gaussian filter for image restoration, where the goal is to recover the original image from a blurred version. This method is particularly effective for inverse filtering when some information about the blurring process is available. However, even if the details of the blur are unknown, Gaussian filtering can still be applied to improve the image. The primary objective of this method is to reduce the MSE between the recovers image and the original. It considers both the effects of blurring and the statistical properties of noise present in the image. The Wiener filter seeks to rebuild the original image while also suppressing noise through smoothing operations. In inverse filtering, when a degraded image is transformed into the frequency domain $J(s,t)$ and divided by the degradation function $L(s,t)$, the resulting output yields the reconstructed image $\hat{J}(s,t)$.

$$\frac{1}{L(s,t)} J(s,t) = \hat{J}(s,t) \quad (2)$$

If the module is predetermined L and the transform $J(s,t)$ is obtained, the outcome is $\hat{J}(s,t)$. In inverse filtering, the error term is not predetermined. If the value $L(s,t)$ becomes small, then the error effect is magnified. Such errors are disregarded in the Wiener filter. Further, the equation is transformed as follows,

$$\hat{J}(s,t) = \left[\frac{L^*(s,t) V_f(s,t)}{V_f(s,t) (s,t)^2 + V_m(s,t)} \right] J(s,t) \quad (3)$$

here, $V_f(s, t)$ and $V_m(s, t)$ are the original image spectra and additive noise, $W(s, t)$ is the blur filter. In a certain region of the image, if there is a blur motion, the degradation function $L(s, t)$ is formulated. Certain types of noise $P(s, t)$ can be introduced. Additionally, direct inverse filtering can be applied, where the first image is filtered, and the outcome is processed using Wiener filtering

3.2 Feature extraction and classification

The pre-processed features are fed into the sequential head swin transformer based stacked capsule auto encoder model for detecting the weeds in rice fields. The novel sequential swin transformer consists of Window based Multi-head Self-Attention (W_MSA) and shifted Window Multi-head Self-Attention (SW_MSA). SW_MSA builds upon M_SA by shifting the window positions across the input sequence. Figure 2 illustrates the swin transformer model.

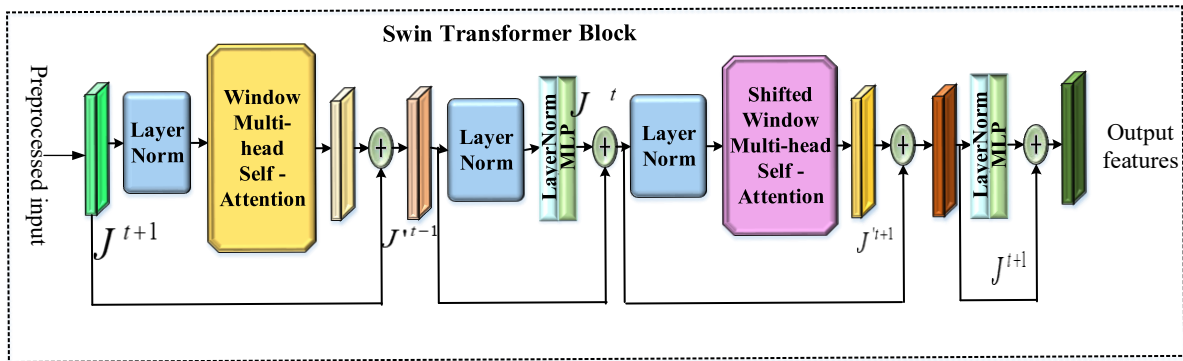


Figure 2: Workflow of the Swin transformer model

This shift allows different windows to interact with each other, enhancing the model to get global context and making it more powerful in feature representation. Every block within the swin transformer comprises two normalization layers, such as the multi-head attention mechanism and the Multi-Layer Perceptron (MLP). Figure 2 shows the sequential arrangement of two swin transformer blocks, one using W_MSA and the other using SW_MSA. To ensure efficient training and avoid issues like vanishing gradients, residual connections are used throughout the architecture. These connections help to stabilize deep transformer models and support their training. By utilizing the window portioning approach, a series of swin transformer blocks can be mathematically represented as follows:

$$J_n^t = W - M_SA(L_N(J^t - 1)) + J^{t-1} \quad (4)$$

$$J^t = M_LP(L_N(J_n^t)) + J_n^t \quad (5)$$

$$J_n^{t+1} = SW_M_SA(L_N(J^t - 1)) + J^t \quad (6)$$

$$J^{t+1} = M_LP(L_N(J_n^{t+1})) + J_n^{t+1} \quad (7)$$

here J_n^t represents the output of (S)_W_MSA module, J^t refers to the M_LP module of the t^{th} block, and the LN denotes normalization layer. Self-attention is formulated below:

$$Atten(q, k, v) = Soft \max \left(\frac{qk^L}{\sqrt{g_p}} + C \right) v \quad (8)$$

where q, k and $V \in S^{N^2 \times g}$ are query, key, and value to measure the similarity, N^2 indicates the count of patches within a window, g represents the *query/key* dimension, C refers to the bias matrix.

SCAE is a modern capsule network that replaces traditional routing mechanisms with an autoencoder. It performs both supervised and unsupervised classification simultaneously. The model contains two main parts: Part Capsule Autoencoder (PCAE) and the Object Capsule Autoencoder (OCAE). When an image is fed into the system, the PCAE identifies and extracts individual parts of objects. The OCAE then assembles these components into finished items. Recognizing various components and how they are arranged to form objects is the foundation of the prediction process. Each capsule in PCAE carries a 6D pose vector, a 1D probability of presence, and an m -dimensional feature vector. The OCAE applies a set transformer to convert part capsules into object level capsules. For classification, SCAE uses linear classifiers in supervised tasks and K-means clustering in unsupervised scenarios. This approach enhances interpretability and part-based object understanding.

PCAE models images by treating them as compositions of geometric parts. To achieve this, initially identify the individual parts in an image and determine their spatial arrangement relative to the viewer. In PCAE, each component capsule m possesses a six dimensional pose Z_m comprising two rotation parameters, two translation parameters, a scaling factor, and a shear value, along with a state variable $e_m \in [0,1]$ and a distinct identity. To detect diseased parts, PCAE uses an auto encoder structure, where the encoder estimates the pose and presence of part capsules and the decoder learns a visual template R_m for each part. If a part is likely to exist based on its presence value, its learned template is transformed using the estimated pose $\hat{R}_m$. These transformed templates are then assembled to reconstruct the input image. After this, the output is passed to the OCAE, which is described below. Further the model treats the image $U \in [0,1]^{r \times n \times b}$ as a grid of pixel values and restrict the count of part capsules to a minimum M threshold and the encoder is utilized to extract each capsule's pose z_m , presence probability e_m and a set of special features denoted as $e_m \in \mathbb{R}^{b_x}$, these essential characteristics can modify the appearance of part patterns depending on the entry of data for instance, to ass colour or apply more complex changes. Additionally, they assist OCAE in learning the distinctive features of parts, like whether they happened or how they relate to other parts. Each template $R_m \in [0,1]^{t_h \times n_h \times (b+1)}$ represents a part and is smaller in size than

the full image U , but includes an extra alpha channel, which likely controls the transparency of the pattern during composition, here, R_m^g utilized to refer alpha channel and R_m^b refer to its colours. Collaborating possibilities of various components are directly proportional to the product of the existence probabilities of part capsules and the learned alpha channel values corresponding to each pattern.

$$Z_{1:M}, e_{1:M}, X_{1:M} = t^{enc}(U) \quad (9)$$

$$b_m = MLP(Z_m) \quad (10)$$

$$\hat{R}_m = TransformImage(R_m, Z_m) \quad (11)$$

$$K_{m,s,c} \alpha e_m \hat{R}_{m,s,c}^g \quad (12)$$

$$K(U) \pi \sum_{s,c}^M K_{m,s,c}^U f(U_{s,c} | b_m \cdot \hat{R}_{m,s,c}^b; \sigma_U^2) \quad (13)$$

When PCAE is trained, it learns visual patterns or templates that represent different parts of objects.

The OCAE aims to discover full objects using the parts previously identified by the PCAE. To do this, it combines pose vectors Z_m , special features X_m , and flattened templates R_m of each part capsule as inputs. Initially, feed part capsule probabilities e_m , using them to guide the set transformer's attention so that absent parts are ignored during object formation. Further, e_m it is also used to weight the log-likelihood of each part, ensuring that the influence of absence parts is reduced. This is mathematically done by raising the part's likelihood to the power of e_m , bn to stabilize training and avoid the collapse of learned variables; gradients are stopped for all inputs to the OCAE except the special features. Additionally, the discovered object capsules do not directly modify part presence or pose during inference, but their gradients flow backward to influence the lower-level PCAE encoder. Each part capsule is treated as having an independent identity, defined by its template and distinctive characteristics instead of only 2D coordinates. The pose of each component is represented as a combination of predictions from object capsules. For each part, OCAE produces multiple candidate predictions $Y_{p,1:M}$ with a single one chosen as the final output.

$$k(Z_{1:M}, e_{1:M}) = \pi \pi \left[\sum_{p=1}^M \frac{g_p g_{p,m}}{\sum_s g_s \sum_c g_{s,c}} K(Z_m | p, m) \right]^{e_m} \quad (14)$$

Here, $g_p g_{p,m}$ are weights for object and part capsules, and $K(Z_m | p, m)$ this are the probability of observing pose. Training the OCAE involves maximizing this likelihood to help the model

learn object structures. Figure 3 illustrates the swin transformer based auto-encoder architecture.

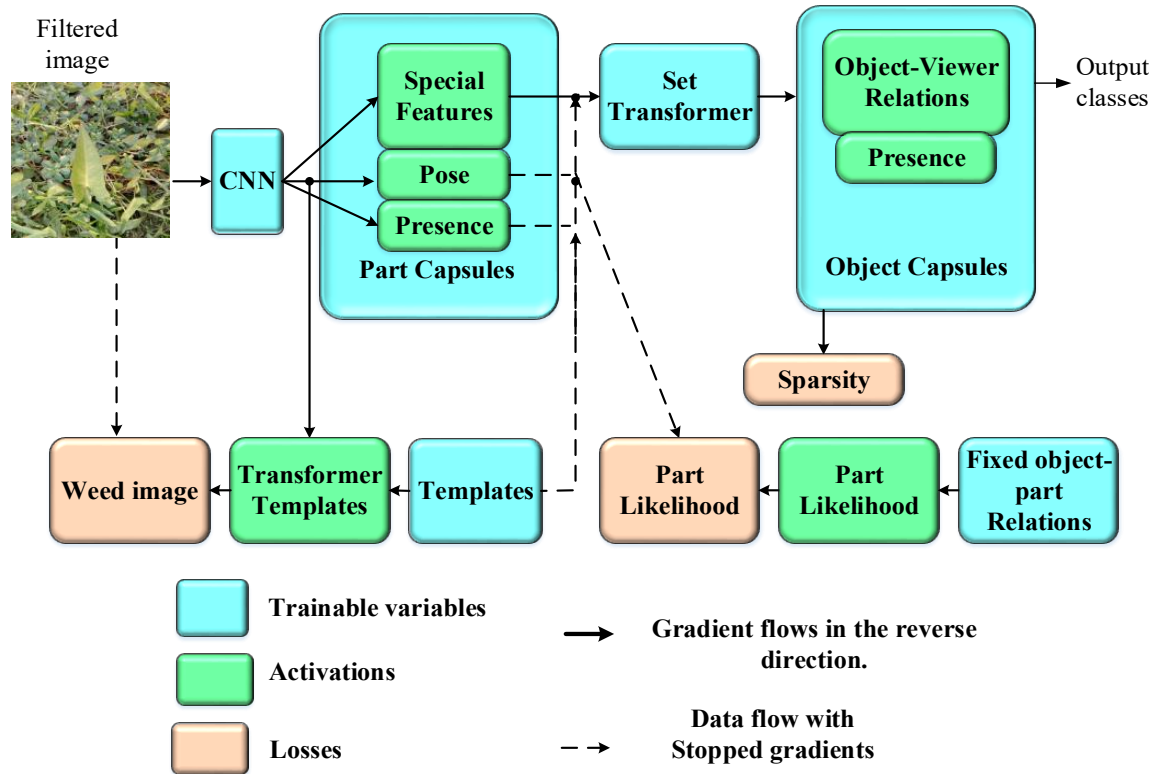


Figure 3: Swin transformer based auto-encoder Architecture

Each class would be assigned P/B capsules in prior sparsity, where P the total number of object is capsules and B is the number of classes. This is assuming that the training data contains objects from different categories that appear with equal probability and that the goal is to assign an equal number of object capsules to each category. Furthermore, only P/B capsules should be active for each training sample in order to assume that each image contains a single object. This implies that on average, the sum of the presence probabilities across all samples should be C/B for each object capsule, where C is the batch size.

$$\mathcal{L}_{prior} = \frac{1}{C} \sum_{c=1}^C (\hat{y}_c - P/B)^2 + \frac{1}{P} \sum_{p=1}^P (\hat{y}_p - C/B)^2 \quad (15)$$

In posterior sparsity, minimizing the capsule posterior presence $\mathcal{H}(\bar{w}_p)$ and maximizing between entropy $\mathcal{H}(\hat{w}_c)$, here $\mathcal{H}$ refers to the entropy, further $\bar{w}_p$ and $\hat{w}_c$ are the standardized form of $\sum_{p,m} g_{c,p,m}^{posterior}$ and $\sum_{c,m} g_{b,p,m}^{posterior}$ respectively. The overall loss is expressed as,

$$\mathcal{L}_{posterior} = \frac{1}{P} \sum_{p=1}^P \mathcal{H}(\bar{w}_p) - \frac{1}{C} \sum_{c=1}^C \mathcal{H}(\hat{w}_c) \quad (16)$$

Finally, the classifier utilized the capsule output encoded by OCAE to detect the label of the image.

3.3 Opposition based Orangutan Optimization Approach (O based OOA)

The O^3 approach boosts model training and refines the parameters for better results. OOA is a nature inspired optimization technique that mimics the social and food searching behaviours of orangutans. In this algorithm, each orangutan symbolizes a possible solution, and their locations in the search space are updated step by step to effectively search through and make use of the solution space. The algorithm works through two main phases foraging phase, which focuses on exploring new areas, and the nesting phase, which focuses on refining and improving existing solutions.

3.3.1 Initialization

The population of candidate solutions, representing the orangutans, which are initialized randomly across the search space. The number of orangutans and the number of choice factors are represented by the vector $J \times k$ matrix that encodes each orangutan's position.

Here, J denotes the quantity of orangutans and k is the count of decision variables. This initialization is formulated as,

$$Y = \begin{bmatrix} y_{1,1} & \cdots & y_{1,s} & \cdots & y_k \\ \vdots & & \ddots & & \vdots \\ y_{r,1} & \cdots & y_{r,s} & \cdots & y_{r,k} \\ \vdots & & \ddots & & \vdots \\ y_{J,1} & \cdots & y_{J,s} & \cdots & y_{J,k} \end{bmatrix}_{J \times k} \quad (17)$$

$$y_{r,s} = mc_s + l \times (rc_s - mc_s) \quad (18)$$

Here, rc_s and mc_s are lower and upper bounds of s^{th} the decision variable, and l is a random number in $[0, 1]$. The objective function of the optimization problem is used to evaluate each option and determine how effectively the solution solves the problem. These fitness scores are crucial in steering the algorithm's iterative improvement process. They help to identify the best performing solution, known as the current leader, as well as the worst. The leader is crucial for guiding the search in subsequent iterations, enabling the algorithm to concentrate on more promising regions of the solution space

3.3.2 Exploration phase

During the exploration phase, the algorithm imitates how orangutans search wide regions to locate food sources. Better solutions, or orangutans with higher fitness scores, have an impact on the other orangutans movements throughout this phase. This enables the algorithm to examine the search domain in a more effective manner. The updated position of each

orangutan is computed using a specific mathematical formula, which is designed to help the population spread out and investigate new potential solutions in the problem space.

$$y_{r,s}^{Po1} = y_{r,s} + l \cdot (SD_{r,s} - F \cdot y_{r,s}) \quad (19)$$

Here, $SD_{r,s}$ refers the selected weed source, F is a random value in $\{1, 2\}$. If the updated location increases the objective function score, then the updated position will be accepted.

$$Y_r = \begin{cases} Y_r^{Po1}, & T_r^{Po1} < T_r \\ Y_r, & else \end{cases} \quad (20)$$

This phase helps to maintain variety among potential solutions by encouraging the exploration of unvisited areas in the solution space, thus avoiding premature convergence.

3.3.3 Exploitation phase

In this phase, the focus shifts to refining existing solutions, similar to how orangutans construct their nests close to where they currently are. This stage emphasizes local search and fine tuning, aiming to improve the current solutions by making smaller, more precise adjustments. The new position for each orangutan is calculated using a specific formula designed to enhance performance by exploiting promising areas already found in the search space.

$$y_{r,s}^{Po2} = y_{r,s} + (1 - 2 u_{r,s}) \cdot \frac{lc_s - mc_r}{f} \quad (21)$$

Here, f denotes the current iteration count and $u_{r,s}$ refers to the random value in $[0,1]$. The position is changed if the objective function improves:

$$Y_r = \begin{cases} y_{r,s}^{Po2}, & T_r^{Po2} < T_r \\ y_r, & else \end{cases} \quad (22)$$

This phase focuses on refining the candidate results, allowing the model to gradually move closer to an optimal outcome. It improves the precision of the search by making small adjustments to already promising solutions. The complete procedure for applying the OOA is outlined in Algorithm 1. This algorithm outlines the O³A approach for solving optimization problems. It begins by defining the problem, initialization parameters, and generating a random initial population. Orangutans use particular equations to update their positions and look for possible weed sources during the exploration phase of the process, which alternates with the exploitation phase, where solutions are improved for greater precision. The optimal solution is monitored and updated during iterations. Finally, the model results in an efficient result after completing a number of iterations.

Algorithm 1
Step 1: Start Step 2: provide problem details: variables, objective function to optimize, and any applicable constraints Step 3: Initialize: population size J, maximum number of iterations k. Step 4: Randomly create an initial solution using equation (17) Step 5: Evaluate fitness Step 6: for $f = 1$ to k do : Step 7: For $r = 1$ to J do : Step 8: exploration phase Step 9: Identify weed sources Step 10: Randomly choose a weed for the r^{th} orangutan Step 11: Compute and update the new position using equations (19) and (20) Step 12: exploitation phase Step 13: Compute refined and update position using equations (21) and (22) Step 14: end for Step 15: Track best solution Step 16: Return the best solution Step 17: end

4. Simulation results

This module illustrates the benefits and innovation of the proposed rice weed detection approach. This part initiates with an experimental setup, a description of the dataset, and is demonstrated with experimental results with detailed performance analysis of various weed detection approaches. Tables 2 and 3 illustrate the system configuration and Hyperparameter details.

4.1 Experimental setup

Table 2: System Configuration

Components	Details
Processor	Intel(R) Core (TM) i7-4770 CPU @ 3.40GHz 3.40Hz
Installed RAM	16.0GB
System type	64-bit operating system, x64-based processor

Table 3: Hyperparameter details

Parameter	Description	Typical range
N	Population size	10-50
Max_iter	Maximum number of iterations	50-200
Ib_val	Lower bound value for each parameter being optimized	[1e-5,8]
Epochs	Number of full passes over the training	300
Metrics	Evaluates classification accuracy	['accuracy']
Optimizer	Uses tuned learning rate from 03A	Adam
Model architecture	Defines CBB/Transformer/Hybrid architecture	Build_model()from ISH_ST_SCAE.py

4.1.1 Dataset description

This proposed system utilized rice field weed detection from the mendeley repository [26]. This dataset contains 3632 images, and it contains 11 classes, which were captured through a smartphone camera. Initially, all the images are resized from the original image into a 224×224 dimension, and then the resized image is denoised using a Gaussian Wiener filter.

4.2 Performance metrics

This portion illustrates the evaluation metrics that are employed to evaluate the performance of the proposed swin transformer capsule auto-encoder model. Performance metrics such as accuracy, precision, recall, fl-score, specificity, and so on. Mendeley repository is used to demonstrate the performance of the designed model.

$$Accuracy = \frac{(PT + NT)}{(PT + NT + PF + FN)} \quad (23)$$

$$Precision = \frac{PT}{PT + PF} \quad (24)$$

$$Recall = \frac{PT}{PT + NF} \quad (25)$$

$$Specificity = \frac{NT}{NT + PF} \quad (26)$$

$$F - score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \quad (27)$$

$$MCC = \frac{(PT \times NT - PF \times NF)}{\sqrt{(PT + PF) \times (PT + NF) \times (NT + PF) \times (NT + NF)}} \quad (28)$$

4.2.1 Performance Analysis

Performance of the improved Swin transformer classification model is evaluated against the proposed model with four conventional models, such as VGG_16 [27], ResNet_50 [28], EfficientNet_B0 [29], and CNN [30]. Figure 4 illustrates the accuracy, which is compared with the proposed with existing models.

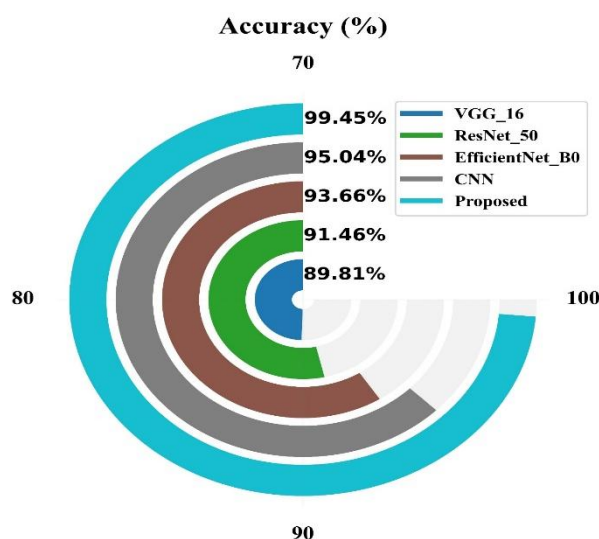


Figure 4: Accuracy analysis

The accuracy graph shows the assessment of the developed model, which is high compared to the existing models such as VGG-16, ResNet_50, EfficientNet_B0, and CNN. The proposed model attained an accuracy of 99.45% and it performs more effectively than existing models. Compared to the other models, CNN displays an accuracy of 95.04%, EfficientNet 93.66%, ResNet_50 91.46%, and VGG_16 89.81%, which is significantly lower. Figure 5 illustrates the precision analysis.

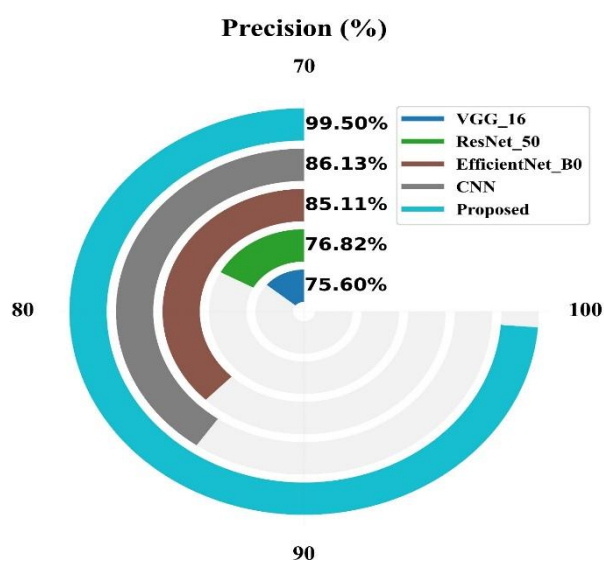


Figure 5: Precision analysis

The above graph illustrates the comparison of precision against four existing models with the proposed model. The improved Swin transformer classification performs well, which gives a high precision of 99.50% for the proposed system, while the precision of CNN and EfficientNet B0 is 86.13% and 85.11% respectively. Precision of VGG_16 is 75.60% which is comparatively very low, as compared to the Precision of ResNet_50 is 86.13%. It demonstrates that this proposed model outperforms all the existing models and performs well in classification. Figure 6 demonstrates the recall evaluation of the developed model.

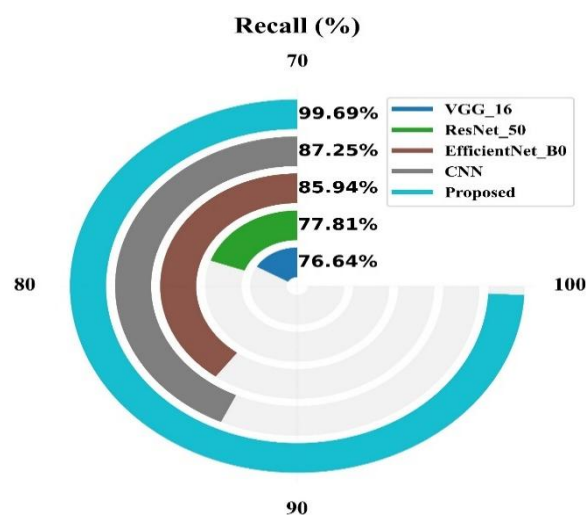


Figure 6: Recall analysis

The above graph shows the recall of the proposed method with four existing models in detecting rice field weeds. Recall of the proposed model is 99.69% which is very high among all the existing methods, indicating that the proposed model performs well in feature extraction and classification. Where the recall of existing models, such as CNN attained 87.25%, EfficientNet_B0 attained 85.94%, ResNet_50 attained 77.81% and the VGG performed very low recall of 76.64% respectively. Figure 7 demonstrates the F1-score evaluation.

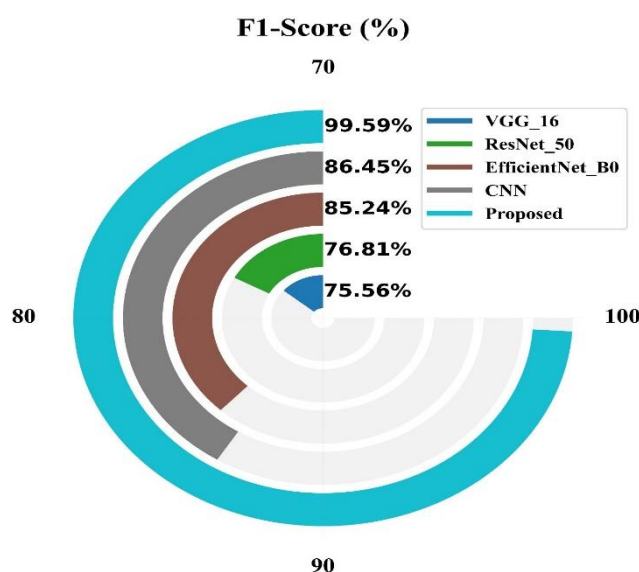


Figure 7: F1-score analysis

The F1-score of the proposed model is 99.59% which outperforms the existing model performance. This metric shows that this proposed system can handle the imbalanced data and maintain accuracy under uncertainty. F1-score of CNN is 86.45%, EfficientNet_B0 of 85.24%, ResNet_50 of 76.81% and VGG_16 of 75.56% respectively. This illustrates, the proposed model is more efficient than the existing models. Figure 8 illustrates the specificity analysis of the proposed model.

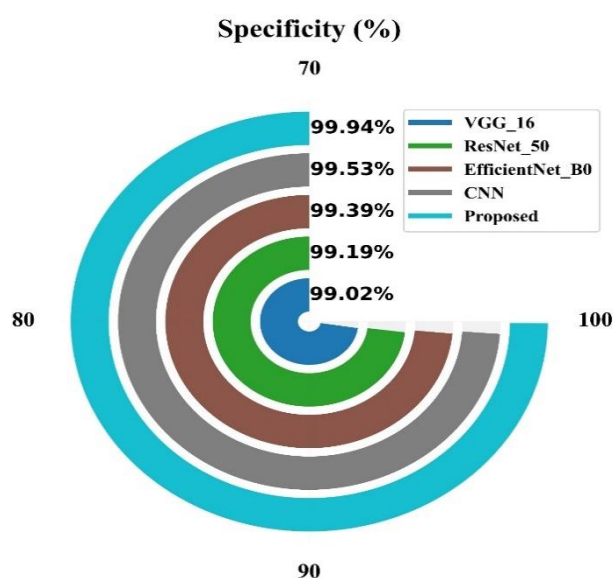


Figure 8: Specificity analysis

The specificity graph shows the developed system performs high than the conventional models such as CNN, EfficientNet_B0, ResNet_50, and VGG_16. The proposed system attained a specificity of 99.94% which demonstrates that this model can efficiently identify the weeds in rice fields. Specificity of the existing CNN model is 99.53%, EfficientNet_B0 of

99.39%, ResNet_50 of 99.19% and VGG_16 of 99.02% which performs very low than the other existing models. Figure 9 illustrates the RMSE analysis.

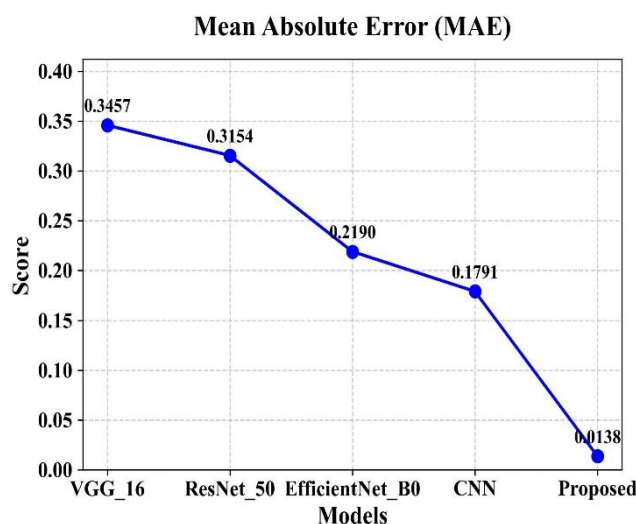


Figure 9: MAE analysis

The graph compares the MAE values of the proposed model with four deep learning models. The VGG_16 shows the highest MAE of 0.3457, indicating lower accuracy. ResNet_50 and EfficientNet B0 perform moderately with MAE of 0.3154 and 0.2190, respectively. CNN performs better while acquiring a lower error of 0.1791. The proposed model achieves a lowest error of 0.0138, showing the most accurate predictions. It shows that this proposed methodology performs better than the existing models. Figure 10 illustrates the RMSE analysis.

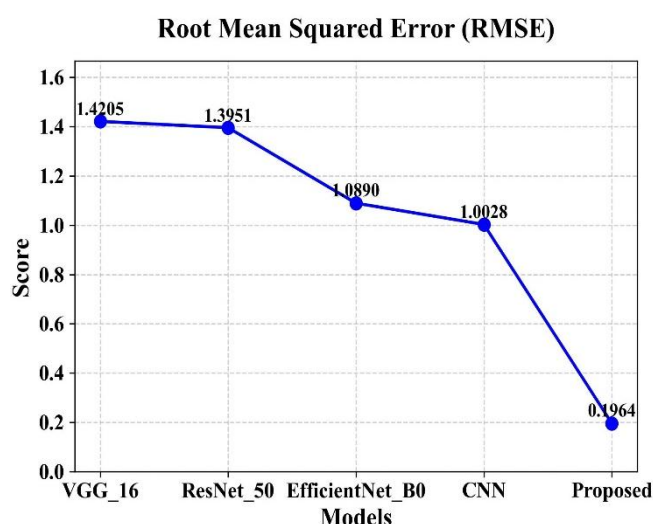


Figure 10: RMSE analysis

The graph illustrates the RMSE values of various existing models with the proposed model, measuring prediction error. VGG_16 and ResNet_50 show high RMSE values of 1.4025 and 1.3951, respectively. EfficientNet_B0 and CNN demonstrate better performance than the

alternatives of two existing models of 1.089 and 1.0028. The proposed model outperforms all the existing models with an RMSE of 0.1964, respectively, which indicates that this proposed model makes better predictions with minimal error rates. Figure 11 illustrates the training and testing accuracy analysis of the developed model.

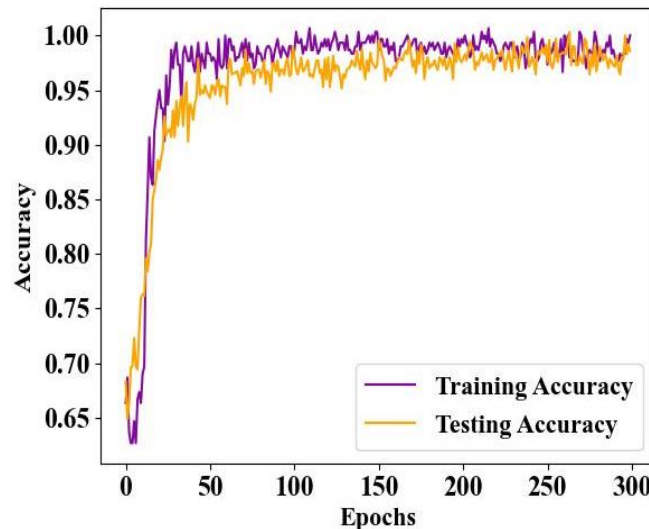


Figure 11: Training and testing accuracy analysis

The above graph shows the training and testing accuracy over 300 epochs. The peak training and testing accuracy show that the model performs well in classification and feature representation. It confirms the superior generalization capacity during training. The robustness during testing shows the model is less prone to overfitting. It confirms the developed methods in real world weed detection in rice fields. Figure 12 illustrates the training and testing loss analysis of the developed model.

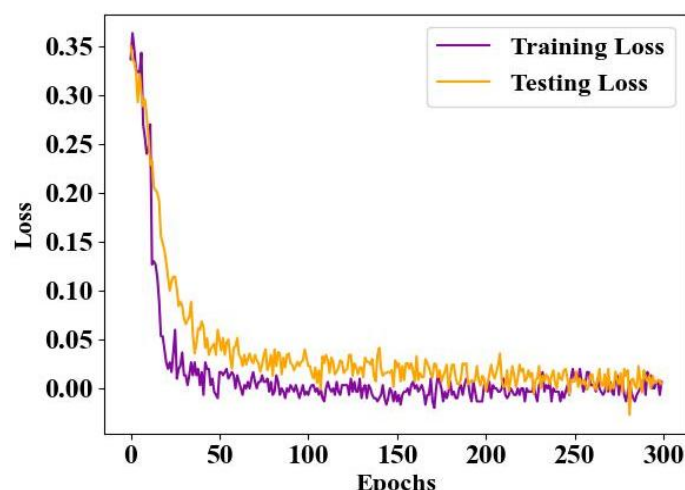


Figure 12: Training and testing loss analysis

The training loss curve illustrates how well each model minimizes error over time. The proposed model achieves the lowest and most stable training loss, indicating fast and effective convergence. Compared to others, it avoids overfitting and maintains a smoother descent. The

lowest training loss indicates better prediction reliability on unseen data. This proves that the proposed model balanced between accuracy and error reduction. ROC curve analysis is illustrated in Figure 13.

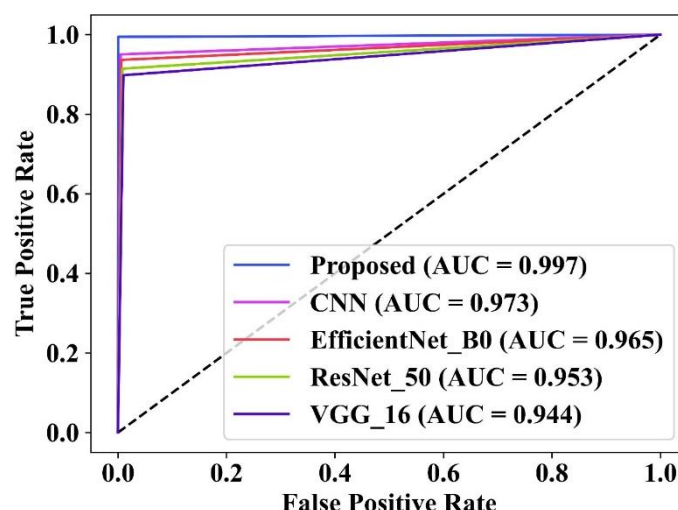


Figure 13: ROC curve evaluation

The ROC curve depicts the effectiveness of various models in terms of their classification ability. The x-axis shows the false positive rates, whereas the y-axis represents the true positive rates. Among all models, the developed model attains the highest AUC of 0.997, showing superior accuracy. CNN follows with an AUC of 0.973, while VGG_16 lags behind at 0.944. AUC of EfficientNet_B0 and ResNet_50 is 0.965 and 0.953, respectively. A higher AUC indicates a better ability to distinguish between classes. Therefore proposed system gives better accuracy than the existing models. Figure 14 illustrates the confusion matrix evaluation of the developed model.

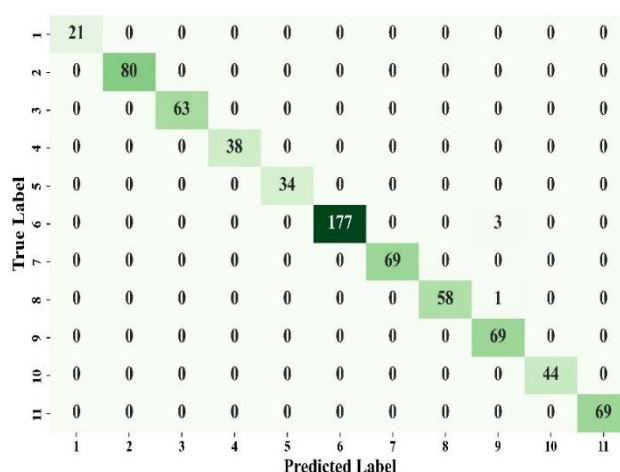


Figure 14: Confusion matrix analysis

The confusion matrix shows the prediction accuracy across 11 classes by showing the relationship between true and predicted labels. The 11 classes are *Alternanthera philoxeroides*, *Centella asiatica*, *Commelina benghalensis*, *Cyperus ochraceus*, *Fimbristylis littoralis*, *Pomoea aquatic*, *Masilea minuta*, *Panicum repens*, *Paspalum scrobiculatum*, *Pteris vittata*, and *Synedrella nodiflora*. Most of the predictions lie along the diagonal, indicating a high correct classification rate. Misclassifications are minimal, such as 3 instances of class 6 predicted as class 7 and 1 instance of class 8 as class 9. The darkest green squares represent the highest number of correct predictions, especially for class 6 with 177 correct outputs. This matrix demonstrates that the model performs strongly across all classes with minimal error.

4.3 Discussion

Weeds are a major factor in lowering rice yield, leading to substantial economic losses and problems with the overall quality of the harvest. They provide a constant threat to crops throughout the whole cultivation process since they can damage them at any point in their growth, from early development to maturity. Therefore, an efficient swin transformer was developed to detect the weeds in the rice field. This stage looks at different weed detection technologies and how the proposed system works better than others. For example, in order to address the challenges of weed detection in rice fields, [21] suggested the PIS-NET model. However, this model has high computational complexity, which this proposed system reduces by using an optimized, lightweight architecture. Similarly, [22] developed a model using EfficientNet and transfer learning to improve forecasting accuracy, but this system is still too expensive and cannot be deployed on robotics. This proposed model uses a lightweight architecture optimized for real-time operation to overcome these limitations. The DWUNet model was proposed by [23], which detects the weeds in fields. However, this system lacks a suitable dataset. To achieve high prediction accuracy, this proposed model has a Mendeley repository with a dataset on weed identification in rice fields. Similarly, [24] suggested a convolutional network, SC-Net, that is computationally demanding and less appropriate for low-cost, real-time deployment. However, the capsule net design of the proposed model is built on Swin Transformers, allowing for quicker and more economical deployment. The CTFFNet model, developed by [25], combines CNN and Transformer for rice field remote sensing image segmentation. However, because it mixes CNN and Transformer, it is computationally complex and less appropriate for low-cost, real-time deployment. However, the proposed model allows for faster and more effective weed detection while maintaining high accuracy and a lower computing cost. Table 4 illustrates the comparative analysis of previous research with the proposed system.

Table 4: Comparative analysis

References	Methods	Accuracy (%)
Chen et al. [21]	PIS-NET	88%
Duong et al. [22]	EfficientNet and transfer learning	90%

Habib et al. [23]	DWUNet	92.34%
Liao et al. [24]	SC-Net	87.44%
Guo et al. [25]	Transformer and Residual CNN	91.42%
Proposed model	Improved swin transformer	99.45%

5. Conclusion

Rice serves as a primary food source for more than half of the global population. Weeds are a major contributor to reduced rice yields, causing substantial economic losses as well as quality degradation of the harvest. They can impact the crop during any phase of its growth cycle. To overcome the above limitations, this proposed system developed an improved swin transformer based capsule auto encoder model to effectively detect the weeds in rice fields. Initially, the data is gathered from the mendeley repository, where the raw images are resized, and the MGWF filter is applied to remove the unwanted noise from the images. Further, the image is sent to pre-processing, ISH-ST-SCAE model is applied to the weeds by extracting the essential features. Additionally, the parameters are fine-tuned through the OOA approach to improve the prediction accuracy. Performance of this system was evaluated using several performance metrics such as accuracy, precision, recall, specificity, and so on. The proposed system attained an accuracy of 99.45%, precision of 99.50%, and recall of 99.69% respectively. Future scope can be extended for real time deployment in autonomous agricultural robots and drones to enhance large scale weed detection efficiency.

References

- [1] Mohapatra, Pravat K., Binod Bihari Sahu, Pravat K. Mohapatra, and Binod Bihari Sahu. "Importance of rice as human food." Panicle architecture of rice and its relationship with grain filling (2022): 1-25.
- [2] Kumar, JR Dinesh, C. Ganesh Babu, and K. Priyadharsini. "An experimental investigation to spotting the weeds in rice field using deepnet." Materials Today: Proceedings 45 (2021): 8041-8053.
- [3] Adhinata, Faisal Dharma, and Raden Sumiharto. "A comprehensive survey on weed and crop classification using machine learning and deep learning." Artificial intelligence in agriculture (2024).
- [4] Rao, A. N., Ravi Gopal Singh, Gulshan Mahajan, and S. P. Wani. "Weed research issues, challenges, and opportunities in India." Crop Protection 134 (2020): 104451.
- [5] Kamath, Radhika, Mamatha Balachandra, Amodini Vardhan, and Ujjwal Maheshwari. "Classification of paddy crop and weeds using semantic segmentation." Cogent engineering 9, no. 1 (2022): 2018791.
- [6] Hollaus, Alexander, Christoph Schunko, Rainer Weissshaidinger, Poline Bala, and Christian R. Vogl. "Indigenous farmers' perceptions of problems in the rice field

- agroecosystems in the upper Baram, Malaysia." *Journal of Ethnobiology and Ethnomedicine* 18, no. 1 (2022): 26.
- [7] Dadashzadeh, Mojtaba, Yousef Abbaspour-Gilandeh, Tarahom Mesri-Gundoshmian, Sajad Sabzi, and Juan Ignacio Arribas. "A stereoscopic video computer vision system for weed discrimination in rice field under both natural and controlled light conditions by machine learning." *Measurement* 237 (2024): 115072.
- [8] Kamilaris, Andreas, and Francesc X. Prenafeta-Boldú. "Deep learning in agriculture: A survey." *Computers and electronics in agriculture* 147 (2018): 70-90.
- [9] Dargan, Shaveta, Munish Kumar, Maruthi Rohit Ayyagari, and Gulshan Kumar. "A survey of deep learning and its applications: a new paradigm to machine learning." *Archives of computational methods in engineering* 27 (2020): 1071-1092.
- [10] Indrakumari, R., T. Poongodi, and Kiran Singh. "Introduction to deep learning." *Advanced deep learning for engineers and scientists: a practical approach* (2021): 1-22.
- [11] Ashraf, Taskeen, and Yasir Niaz Khan. "Weed density classification in rice crop using computer vision." *Computers and Electronics in Agriculture* 175 (2020): 105590.
- [12] Ali, Md Sawkat, Mohammad Rifat Ahmmad Rashid, Tasnim Hossain, Md Ahsan Kabir, Md Kamrul, Sayam Hossain Bhuiyan Aumy, Mehedi Hasan Mridha, Imam Hossain Sajeeb, Mohammad Manzurul Islam, and Taskeed Jabid. "A comprehensive dataset of rice field weed detection from Bangladesh." *Data in Brief* 57 (2024): 110981.
- [13] Asad, Muhammad Hamza, and Abdul Bais. "Weed detection in canola fields using maximum likelihood classification and deep convolutional neural network." *Information Processing in Agriculture* 7, no. 4 (2020): 535-545.
- [14] Guo, Zhonghui, Dongdong Cai, Yunyi Zhou, Tongyu Xu, and Fenghua Yu. "Identifying rice field weeds from unmanned aerial vehicle remote sensing imagery using deep learning." *Plant Methods* 20, no. 1 (2024): 105.
- [15] Peng, Hongxing, Zihe Li, Zhiyan Zhou, and Yuanyuan Shao. "Weed detection in paddy field using an improved RetinaNet network." *Computers and Electronics in Agriculture* 199 (2022): 107179.
- [16] Dhanya, V. G., A. Subeesh, N. L. Kushwaha, Dinesh Kumar Vishwakarma, T. Nagesh Kumar, G. Ritika, and A. N. Singh. "Deep learning based computer vision approaches for smart agricultural applications." *Artificial Intelligence in Agriculture* 6 (2022): 211-229.
- [17] Udayananda, G. K. V. L., Chathurangi Shyalika, and P. P. N. V. Kumara. "Rice plant disease diagnosing using machine learning techniques: a comprehensive review." *SN Applied Sciences* 4, no. 11 (2022): 311.
- [18] Razfar, Najmeh, Julian True, Rodina Bassiouny, Vishaal Venkatesh, and Rasha Kashef. "Weed detection in soybean crops using custom lightweight deep learning models." *Journal of Agriculture and Food Research* 8 (2022): 100308.

- [19] Ngugi, Lawrence C., Moataz Abelwahab, and Mohammed Abo-Zahhad. "Recent advances in image processing techniques for automated leaf pest and disease recognition—A review." *Information processing in agriculture* 8, no. 1 (2021): 27-51.
- [20] Attri, Ishana, Lalit Kumar Awasthi, and Teek Parval Sharma. "Machine learning in agriculture: a review of crop management applications." *Multimedia Tools and Applications* 83, no. 5 (2024): 12875-12915.
- [21] Chen, Hao, Youliu Zhang, Caijie He, Chuting Chen, Yaoxuan Zhang, Zhiying Chen, Yu Jiang, Caixia Lin, Ruijun Ma, and Long Qi. "PIS-Net: Efficient weakly supervised instance segmentation network based on annotated points for rice field weed identification." *Smart Agricultural Technology* 9 (2024): 100557.
- [22] Duong, Linh T., Toan B. Tran, Nhi H. Le, Vuong M. Ngo, and Phuong T. Nguyen. "Automatic detection of weeds: synergy between EfficientNet and transfer learning to enhance the prediction accuracy." *Soft Computing* 28, no. 6 (2024): 5029-5044.
- [23] Habib, Mohammed, Salma Sekhra, Adil Tannouche, and Youssef Ounejjar. "New segmentation approach for effective weed management in agriculture." *Smart Agricultural Technology* 8 (2024): 100505.
- [24] Liao, Juan, Minhui Chen, Kai Zhang, Huiyu Zhou, Yu Zou, Wei Xiong, Shun Zhang, Fuming Kuang, and Dequan Zhu. "SC-Net: A new strip convolutional network model for rice seedling and weed segmentation in paddy field." *Computers and Electronics in Agriculture* 220 (2024): 108862.
- [25] Guo, Zhonghui, Dongdong Cai, Zhongyu Jin, Tongyu Xu, and Fenghua Yu. "Research on unmanned aerial vehicle (UAV) rice field weed sensing image segmentation method based on CNN-transformer." *Computers and Electronics in Agriculture* 229 (2025): 109719.
- [26] <https://data.mendeley.com/datasets/mt72bmzx73/4>
- [27] Mannepalli, Praveen Kumar, Ayonija Pathre, Gunjan Chhabra, Priyanka Anup Ujjainkar, and Shrutika Wanjari. "Diagnosis of bacterial leaf blight, leaf smut, and brown spot in rice leafs using VGG16." *Procedia Computer Science* 235 (2024): 193-200.
- [28] Al-Gaashani, Mehdhar SAM, Nagwan Abdel Samee, Rana Alnashwan, Mashael Khayyat, and Mohammed Saleh Ali Muthanna. "Using a Resnet50 with a kernel attention mechanism for rice disease diagnosis." *Life* 13, no. 6 (2023): 1277.
- [29] Sathya, K., K. Kanmani, M. Revathy Meenal, D. Suganthi, and T. S. Lakshmi. "Plant Disease Identification Using Extended-EfficientNet Deep Learning Model in Smart Farming." *Automated Machine Learning and Industrial Applications* (2025): 207-241.
- [30] Ilangoan, P., K. Meena, and M. Sharmila Begum. "Weedbot: Weed Easy Extraction Using Deep Learning with a Robotic System." In *Journal of Physics: Conference Series*, vol. 2923, no. 1, p. 012003. IOP Publishing, 2024.
- [31] Kumbi, Arun A. et al. "Fractional deer hunting optimization algorithm enabled deep learning framework for sugarcane billet damage classification." *Int. J. Syst. Assur. Eng. Manag.* 16 (2025): 2176-2192.

- [32] Kumbi, Arun A. and Mahantesh N. Birje. "A novel authentication scheme for secure data sharing in IoT enabled agriculture." *Web Intelligence* 22 (2023): 381 - 400.
- [33] Kumbi, Arun A. and Mahantesh N. Birje. "Deep CNN based Sunflower Atom Optimization Method for Optimal Water Control in IoT." *Wireless Personal Communications* 122 (2021): 1221 - 1246.
- [34] Kumbi, Arun A. et al. "Sunflower Optimization Technique for Ideal Water Control using Deep CNN." 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS) (2024): 1-5.