https://doi.org/10.17221/77/2022-RAE

# A novel hybrid feature method for weeds identification in the agriculture sector

*Sheeraz Arif[1,2]*, Rashid Hussain[3], Nadia Mustaqim Ansari[4], Waseem Rauf[1]*

[1]*Department of Computer Science, Faculty of Information Technology, Salim Habib University, Karachi, Pakistan*
[2]*Department of Information and Communication Engineering, Beijing Institute of Technology, Beijing, China*
[3]*Department of Electrical Engineering, Faculty of Engineering Science and Technology, Hamdard University, Karachi, Pakistan*
[4]*Department of Electronic Engineering, Dawood University of Engineering and Technology, Karachi, Pakistan*

**Corresponding author: Sheeraz.arif@shu.edu.pk*

**Citation:** Arif S., Hussain R., Ansari N.M., Rauf W. (2023): A novel hybrid feature method for weeds identification in the agriculture sector. Res. Agr. Eng., 69: 132–142.

**Abstract:** Weed identification and controlling systems are gaining great attention and are very effective for large productivity in the agriculture sector. Currently, farmers are facing a weed control and management problem, and to tackle this challenge precision agriculture in the form of selective spraying is much-needed practice. In this article, we introduce a novel framework for a weed identification system that leverages (hybrid) the robust and relevant features of deep learning models, such as convolutional neural network (CNN) and handcrafted features. First, we apply the image preprocessing and augmentation techniques for image quality and dataset size enhancement. Then, we apply handcrafted feature extraction techniques, such as local binary pattern (LBP) and histogram of oriented gradients (HOG) to extract texture and shape features from the input. We also apply the deep learning model, such as CNN, to capture the relevant semantic features. Lastly, we concatenate the features extracted from a different domain and explore the performance using different classifiers. We achieved better performance and classification accuracy in the presence of the extreme gradient boosting (XGBoost) classifier. The achieved results witnessed the effectiveness and applicability of the given method and the importance of concatenated features.

**Keywords:** convolutional neural network; deep learning; handcrafted features; weed detection; XGBoost classifier

In farming, weeds identification and their proper elimination are the major challenges. Weeds are unwanted harmful plants that consume water, soil fertility, minerals, sunlight, soil space, and other natural resources of wanted plants and cause the reduction of crop yield. In many developing countries, weed detection is a costly and time-consuming process. According to an Australian study, around AUD 1.5 billion is spent to control weeds activities in terms of herbicides and machinery, and lose AUD 2.5 on weeds-affected cultivation production Olsen et al. (2019).

Excessive usage of herbicides can be hazardous to the environment and unhealthy for crops. To limit the usage of chemical herbicides European Union (EU) formulated a strategy in 2020 known

https://doi.org/10.17221/77/2022-RAE

as the Farm to Fork strategy that aims to limit the usage of chemical herbicides by 50% by 2030 European Commision (2021). In many developed countries, intelligent weed control systems have been employed for precision agriculture, also known as spot spraying. The main purpose of this technique is to reduce labour costs and potentially limit the usage of chemical herbicides. In addition, we can preserve the crop's health and protect the environment by using a reduced amount of chemicals.

According to the research studies, about 250 000 various species of plants exist, and there are approximately 260 species related to the weed plants (Lowry and Smith 2018; Zimdahl 2018a; Zimdahl 2018b). However, different varieties of weeds and plants with respect to their shapes, colour, size, presence of overlapping, environmental changes, plant characteristics, and types of soil are the factors that make the intelligent spraying system more challenging. Due to the advancement in Artificial Intelligence (AI), many farmers adopted the intelligent spraying system to detect and remove weeds from wanted plants.

Many researchers have successfully applied the concepts of image processing and computer vision to tackle the problems of weed detection and proposed a variety of AI-based algorithms. The important steps in image processing and computer vision are image capturing, pre-processing of images, augmentation, features extraction, feature selection, and classification. Among these steps, feature extraction and classification are the most important steps, and many researchers proposed effective methods based on machine learning (ML, sub-domain of AI). ML algorithms are further categorised into two types: (*i*) Extracting handcrafted features from the given raw data termed as conventional ML methods and other methods automatically extracting features that help to classify data points known and *(ii)* deep learning (DL) methods (Janiesch et al. 2021). Both ML and DL-based methods have their own strengths and weaknesses. Representations of hand-crafted local features are considered very effective in dealing with problems, such as background clutter, noise and illumination changes, and orientation changes. In addition, expert knowledge can also be embedded during the procedure of feature extraction and require a small dataset to train the model. However, feature extraction/selection is a very cumbersome process and needs expert domain knowledge (Aversano Arif et al. 2019). Deep-learning-based methods automatically learn a high level of semantic information from training data without applying any heuristic rules. The ability of diverse feature generation at multiple hierarchal levels can be applied to perform different tasks, and the detection/identification of weeds is one of them. However, deep-learning-based models require huge data and computational power for training.

Early research works mostly related to hand-crafted features, and many researchers used these techniques to classify weeds/crops based on different characteristics, such as shape, colour, texture, spectrum, edges, and spatial and geometrical features. Yang et al. (2017) and Sabzi et al. (2018) proposed a method based on colour and texture features. Le et al. 2019; Le et al. 2020 applied a local binary pattern (LBP) for the extraction of texture features and successfully differentiated between crop and weed. Some research studies (Hamuda et al. 2017, Bakhshipour and Jafari 2018; Chen et al. 2021) combine the different shapes and textures-based features and obtain remarkable results. Ma et al. (2016) identified the grape leaves by combining a histogram of oriented gradients (HOG) features with a support vector machine (SVM). He et al. (2013) combined the multi-source information of different features like texture, shape, and fractal dimension with SVM.

Due to the remarkable success of DL, its application has also been extended to the agricultural field (Aversano et al. 2020; Fu et al. 2020). Recently, convolutional neural networks (CNNs) have been the most widely applied networks to solve weed classification problems. Researchers adopted different variants of CNNs, such as AlexNet (Krizhevsky et al. 2012), GoogLeNet (Szegedy et al. 2015), ResNet-50 (He et al. 2016), Inception-v3 (Szegedy et al. 2016), VGG-Nets (Simonyan and Zisserman 2015), and achieved high performance. Olsen et al. (2019) developed a weeds' image dataset and trained ResNet-50 architecture to get real-time performance. Tang et al. (2017) proposed a clustering-based method using the K-means) feature with multilayered CNN architecture. Zou et al. (2021) proposed a segmentation algorithm to cluster semantic features using a simplified U-net to separate weeds and plants in images. You et al. (2020) introduced another segmentation model for weeds and crop identification using deep neural network models. These DL-based methods have shown impressive improvement in solving the problem of weed classification and achieved promising performance.

https://doi.org/10.17221/77/2022-RAE

However, more accurate and effective classification models are required to enhance weeds and plants' identification/classification accuracy. In this research work, we introduce a model in the context of intelligent selective spraying based on hybrid features (dense and handcrafted features) for the classification of weeds and plants and argue that the leveraging of features extracted from different channels can increase the generalisation and recognition ability (Arif et al. 2019; Wu et al. 2021). Recently, Wang et al. (2022) developed a new dataset for weeds and utilised YOLOv3, YOLOv5, and Faster R-CNN for weed identification. Razfar et al. (2022) proposed a new 5-layer CNN architecture with few training parameters and achieved better results. Yang et al. (2022) utilised a combination of object detection networks (Faster R-CNN and YOLOv3) and successfully discriminated the grass weed from plants.

This research work has emphasised the extraction and leveraging of handcrafted and CNN-based features on the basis of the argument that a combination of features captured from different channels can really increase the generalisation ability of weed and plant classification. Handcrafted features are well capable of extracting low-level features, such as edges, shape, and texture-based features, and the DL method provides us with dense hierarchal features. Firstly, we perform the image pre-processing step such as noise suppression using a median filter and image quality enhancement (pixel-wise) using contrast-limited adaptive histogram equalisation (CLAHE). Then, to increase the size of the dataset, we perform some augmentation techniques, such as flipping and rotation. Next, handcrafted features are extracted from RGB images using LBP (Ojala et al. 1996) and HOG (Dalal and Triggs 2005) for texture and shape features, respectively. High-level dense features are extracted using CNN. After concatenation of features from both domains (handcrafted and DL), we apply the extreme gradient boosting (XGBoost) classifier (Chen and Guestrin 2016) and obtain binary classification results in the form of weed or plant. Research contributions of our work are listed as follows:

(*i*) We successfully combine the handcrafted features (HOG and LBP) with CNN-based features and obtain superior performance.

(*ii*) We apply the image processing and augmentation method for cleaning and image enhancement and also to enlarge the dataset size.

The aim of this research study is to explore and verify the effectiveness of hybrid features (extracted from different domains) to improve the identification accuracy between weeds and plants to benefit the agriculture sector.

## MATERIAL AND METHODS

To train our model, we used the Deep Weed dataset collected by an Australian research group Olsen et al. (2019). However, we have performed some pre-processing steps to increase the size, generality, and variability of the datasets. We use Raspberry Pi-3 with a Pi camera of version 2.1 (Raspberry, Taiwan) for image acquisition with a video resolution of 1 280 × 720.

To increase the size of the dataset (new weeds images) synthetically, we perform the data augmentation method. It is an ideal approach to avoid the overfitting effect during the learning process. In addition, the augmentation methods bring variability to the existing dataset and increase the classification accuracy, and address the issues of unbalanced data. We utilise re-sizing, rotation, flipping, and colourisation as augmentation processes in our experiments.

In our experiment, each original image from the training subset is re-sized from 1 280 × 720 (original image size) to 224 × 224 as our proposed model accepts a fixed size input of 224 × 224 pixels. Each image is rotated with angles 45, 90, 180, and 270° to enlarge the datasets with different orientations. Each image was vertically and horizontally flipped with 50% probability, and in the case of the RGB image, each colour channel was randomly shifted about 10% of the maximum available 8-bit colour encoding range [0, 255]. To prepare the training and testing set, the target classes are marked, and labels of each sample image are obtained. We adopt the ($k$)-fold cross-validation, and all the original images are divided into five sub-sample sets. These sub-sets with original images (weeds and plants images) are partitioned into 70, 10, and 20% split of training, validation, and training sets, respectively. The validation subset is used to monitor the training process and overfitting reduction. The validation subset also ensures that image samples are random and not repeated in a test subset. The 20% of images in our testing subset were not used during the training process. This process is essential to minimise bias errors.

**Proposed methodology.** This section discusses our proposed framework, and its working flow is given

https://doi.org/10.17221/77/2022-RAE

in Figure 1. The main objective of this article is to show the effectiveness of hybrid (combined/concatenated) features extracted from handcrafted and CNN-based methods to improve the classification accuracy between weeds and plants. We separately extracted the features HOG (for shape features), LBP (for texture features), and CNN (discriminative) features, and then we fuse/concatenate all of them to make combined feature descriptors before feeding them to our classifier. We provide a detailed description of each step, such as pre-processing and augmentation methods, extraction of handcrafted features using LBP and HOG techniques, extraction of high-level features using CNN, merging of features, and classification.

To obtain a better classification performance, we apply image pre-processing (noise suppression and image enhancement) and augmentation techniques (to increase and balance the size of data). This step is very important as raw data may distract the classification model and result in a high misclassification rate. During the acquisition process, images may be damaged by several factors, such as poor resolution, noise, improper lighting, etc. First, we apply the median filter (non-linear) on raw images to preserve the edges and noise suppression for a clear visual appearance. In median filtering, the 3 × 3 tensor of an image is arranged in ascending order to take the mean, and the middle value of a selected tensor is replaced by the mean value. Next, for image contrast enhancement and noise amplification reduction, we apply CLAHE (Pizer et al. 1990), which focuses on all areas (pixels) of the given image and solves the uneven greyscale distribution in an image. In addition, it limits the over-amplification of contrast. In CLAHE, images are segmented into non-overlapping regions of size 8 × 8, and a histogram of each region is computed. A threshold value is selected, and if the grey levels in the histogram exceed the value of the threshold, then we clip the value, and the excesses value against the threshold value is evenly distributed to all neighbouring grey levels. In this way, noise or over-amplification of the contrast can be controlled. Finally, the neighbouring regions combine with each other with the interpolation method (to enhance the mapping function) to avoid the boundary effect. CLAHE does this by setting a threshold. If some grey levels in the image exceed the threshold, the excess is evenly distributed to all grey levels.

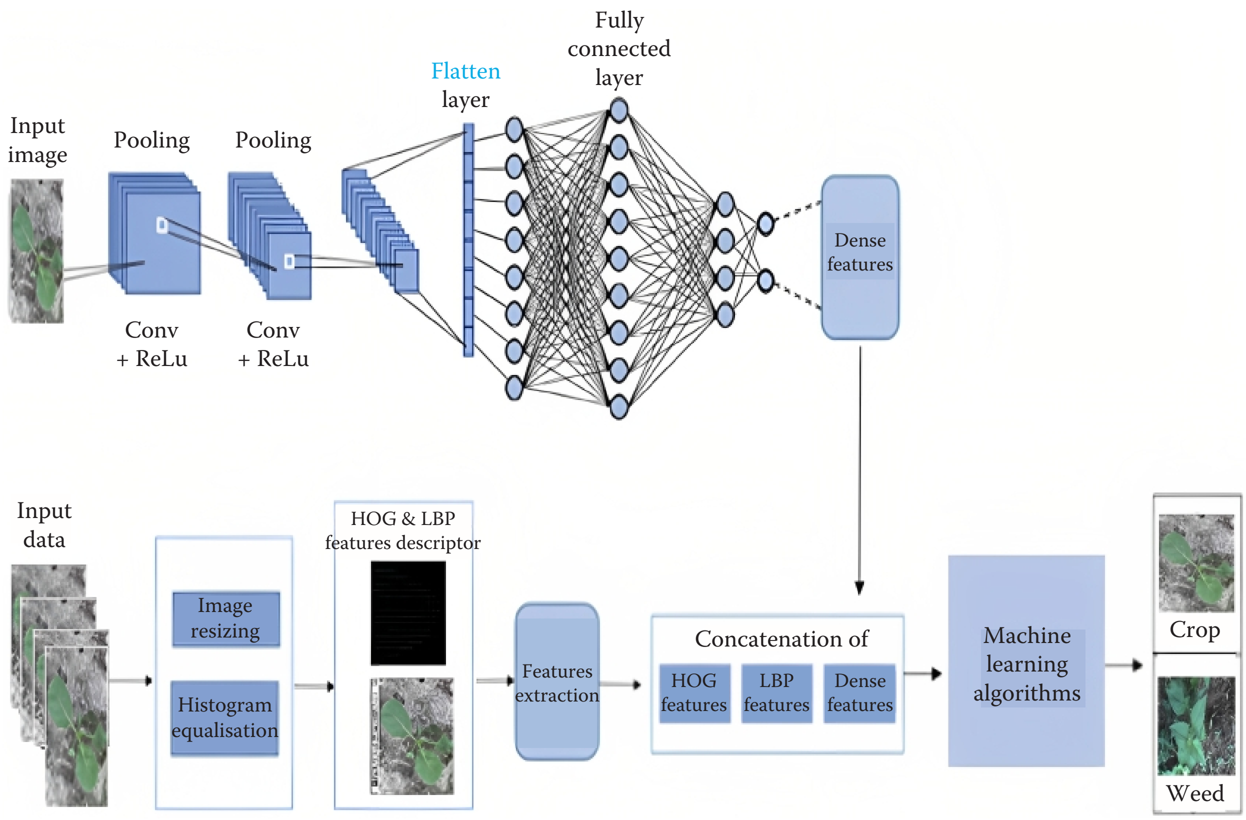


Figure 1. Illustration of the overall workflow of the proposed model

Conv – convolution; ReLu – rectified linear unit; HOG – histogram of oriented gradients; LBP – local binary pattern

https://doi.org/10.17221/77/2022-RAE

In our model, to increase the data size and tackle the unbalance data structure (in case of any dominating class to increasing any inferior class), we adopt data augmentation techniques, such as flipping, rotation, and colourisation. This process generates the data/images synthetically with different orientations, scales, and colours, which is very important to avoid overfitting during the training and validation process of deep learning models. In our experiments, each image is rotated in the range of [+360°, –360°] and zooms up to 20% of the original image's dimension. Each image was vertically and horizontally flipped with 50% probability, and in the case of the RGB image, each colour channel was randomly shifted to about 10% of the maximum 8-bit colour encoding range [0, 255]. Some samples of the augmentation process are given in Figure 2.

In the next step, we present effective ways to extract relevant handcrafted features at minimal computing cost, which can efficiently satisfy the current weed detection/classification requirements. Handcrafted features mean manually extracted features (not automatically extracted) using conventional machine learning algorithms, such as LBP and HOG in our case. Handcrafted features allow embedding domain/expert knowledge in the process of feature extraction. Handcrafted features have optimal interclass variance. Their space is well-portioned. Most of these handcrafted features can be categorised into (*i*) geometric features, (*ii*) texture features, (*iii*) shape features, (*iv*) gradient features, and (*v*) appearance features. In our case, shape and texture are more relevant and significant features.

The LBP is an effective technique to extract the texture and grey-scale contrast in an image. Other important properties of LBP features include computational efficiency and robustness against illumination conditions. LBP descriptor works in a 3 × 3 window which uses the central pixel as the threshold of the neighbouring pixels. In a 3 × 3 win-

(A)
(B)
(C)
(D)
(E)

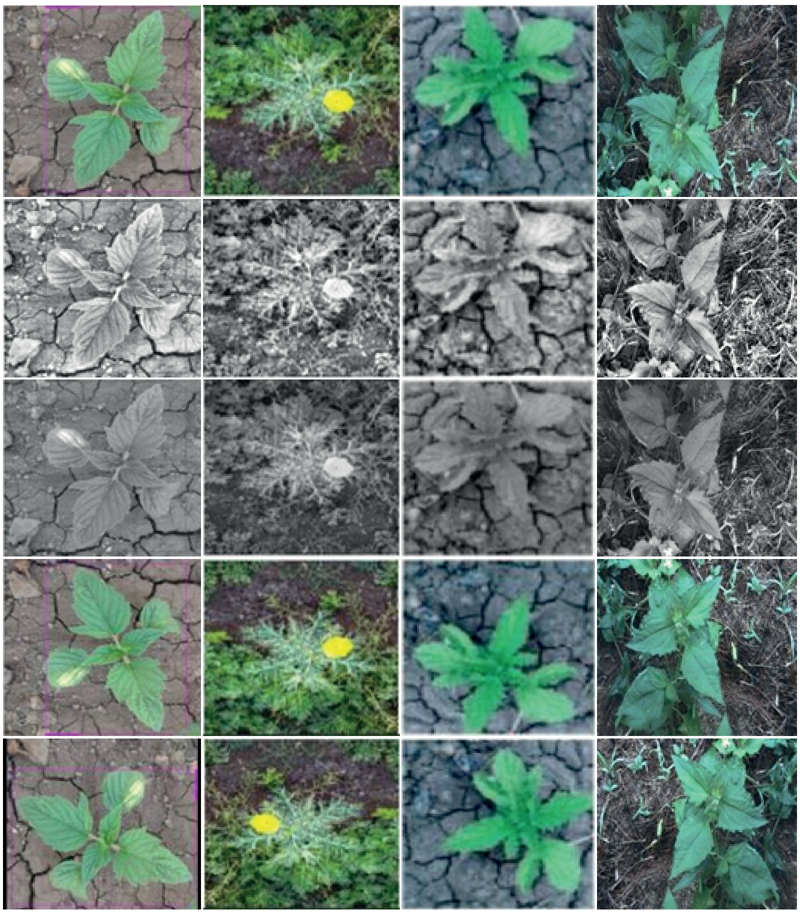

Figure 2. Results of different augmentation methods with sample frames: (A) original frames of plants and weeds, (B) results of CLAHE, (C) results of colour shifting, (D) results of flipping, and (E) results of rotation

CLAHE – contrast-limited adaptive histogram equalisation

https://doi.org/10.17221/77/2022-RAE

dow, for every pixel ($x$, $y$) with $N$ neighbouring pixels at $R$ radius, it computes the difference in the intensities of the central pixel ($x$, $y$) with the neighbouring $N$ pixels, refer to Equation (1). If the resultant difference is negative, it assigns 0 to the pixel ($x$, $y$) or 1 otherwise [Equation (2)]. Lastly, it converts the central pixel value with its corresponding decimal value.

$$\text{LBP}(N, R) = \sum_{n=0}^{N-1} s\left(i_n - i_c\right) 2^p \tag{1}$$

where: $i_n$ – neighbouring pixels ; $i_c$ – current pixels.

$$f(x) = \begin{cases} 1 \text{ ; if } x > 0 \\ 0 \text{ ; otherwise} \end{cases} \tag{2}$$

The HOG feature descriptor emphasises the structure/shape of the objects. It computes the gradient and orientation of edge pixels. These orientations give information related to the global shape of an object presents in the image. The image of size 224 × 224 is passed through the HOG feature descriptor. The descriptor divided the image into several blocks. Each block consists of 16 × 16 cells and each cell contains a 16 × 16 number of pixels for each block, the gradient with its orientation is computed. Lastly, a histogram is created using the gradients and orientations for each region separately, given the name histogram of oriented gradients.

In the next step, we extract complex and dense features using the CNN model. CNN is the type of feed-forward network comprised of end-to-end trainable multi-layer architecture that follows the hierarchal neural mechanism of the human brain that transfers information from low-level to high-level and represents discriminative information.

CNN network usually adopts two techniques for training, i.e. transfer learning (TL) and de novo. In de novo techniques the CNN model is trained from scratch and optimally learns features from the given dataset. TL adopts pre-trained freely available CNNs, such as AlexNet, GoogLeNET, VGG-Net, ResNet, and Generative adversarial network (GAN) and deals with small data. TL technique allows to re-train of only a few layers (usually the last layers are re-trained) of the pre-trained CNN model in order to adapt it to a given problem and dataset. This technique is helpful when the given dataset is not large enough to train robustly all layers of CNN (Pan and Yang 2010).

In our method, we use the TL approach and deep-learned features are extracted from VGG-16. CNN mainly consists of five parts: a convolution layer (with channels 64, 128, 256, and 512), max-pooling, a fully connected layer, an output layer, and an activation function. Our CNN network comprises 18 layers that include 12 convolutional layers, 5 max-pooling layers, one fully connected (FC) layer, and one output layer with the softmax function to calculate prediction probabilities. It takes an image of the size 224 × 224 and 3 × 3 convolutions with a 1-pixel stride and 1 padding. Rectified linear unit (ReLU) serves as an activation function, with Softmax classifier at the last layer. The details of our CNN architecture are depicted in Figure 3.

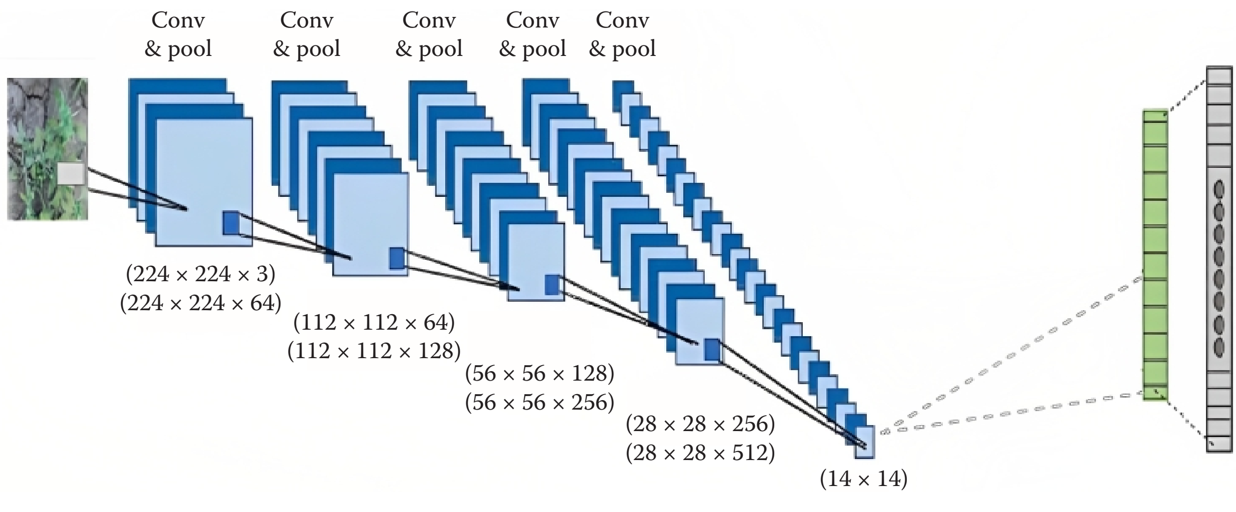


Figure 3. Illustration of CNN architecture used in our model

CNN – convolutional neural network; conv – convolution; pool – pooling

https://doi.org/10.17221/77/2022-RAE

Next, we concatenate all (HOG, LBP, and CNN-based) features in the form of feature descriptor, and then we train different classifiers, such as XGBoost, K–nearest neighbour (KNN) (Cover and Hart 1967), SVM (Vapnik et al. 1999), and attained reasonable results. However, XGBoost achieved superior performance when using hybrid features. It is a tree-based ensemble classifier in which new learners are added to minimize the errors made by the prior learners. XGBoost works similarly to a gradient-boosting algorithm that predicts the residuals of prior learners and adds them to the final predicted result. It minimizes the loss of each learner by using a gradient descent algorithm (Kiefer and Wolfowitz 1952).

## RESULTS AND DISCUSSION

The progress of the two models, i.e. VGG-16 and our introduced hybrid model in terms of accuracy and loss (cross-entropy) after successive epochs during the training process is depicted in Figure 4. The top two figures (4A–B) show the improvement in training accuracy of the two models for every mini-batch at 125 epochs. The accuracy achieved for training and validation of our proposed model (hybrid) is 95.5% as compared to the state-of-the-art VGG-16 model of 94.4%. We can also analyse from the figure that the training process is quite smooth, parallel, and consistent for training and validation of the proposed model. The two bottom figures (4C–D) illustrate the training and validation loss of the two models. Our proposed model shows the lowest training (0.07) and validation (0.09) loss as compared to the VGG-16 architecture of training (0.10) and validation (0.23) with a number of 125 epochs. The underlying reason for better scores for training and validation accuracy of our introduced model is an equal and effective contribution of the features achieved from two different domains, i.e. handcrafted and deep learning, and provide the intrinsic capabilities to obtain better performance.

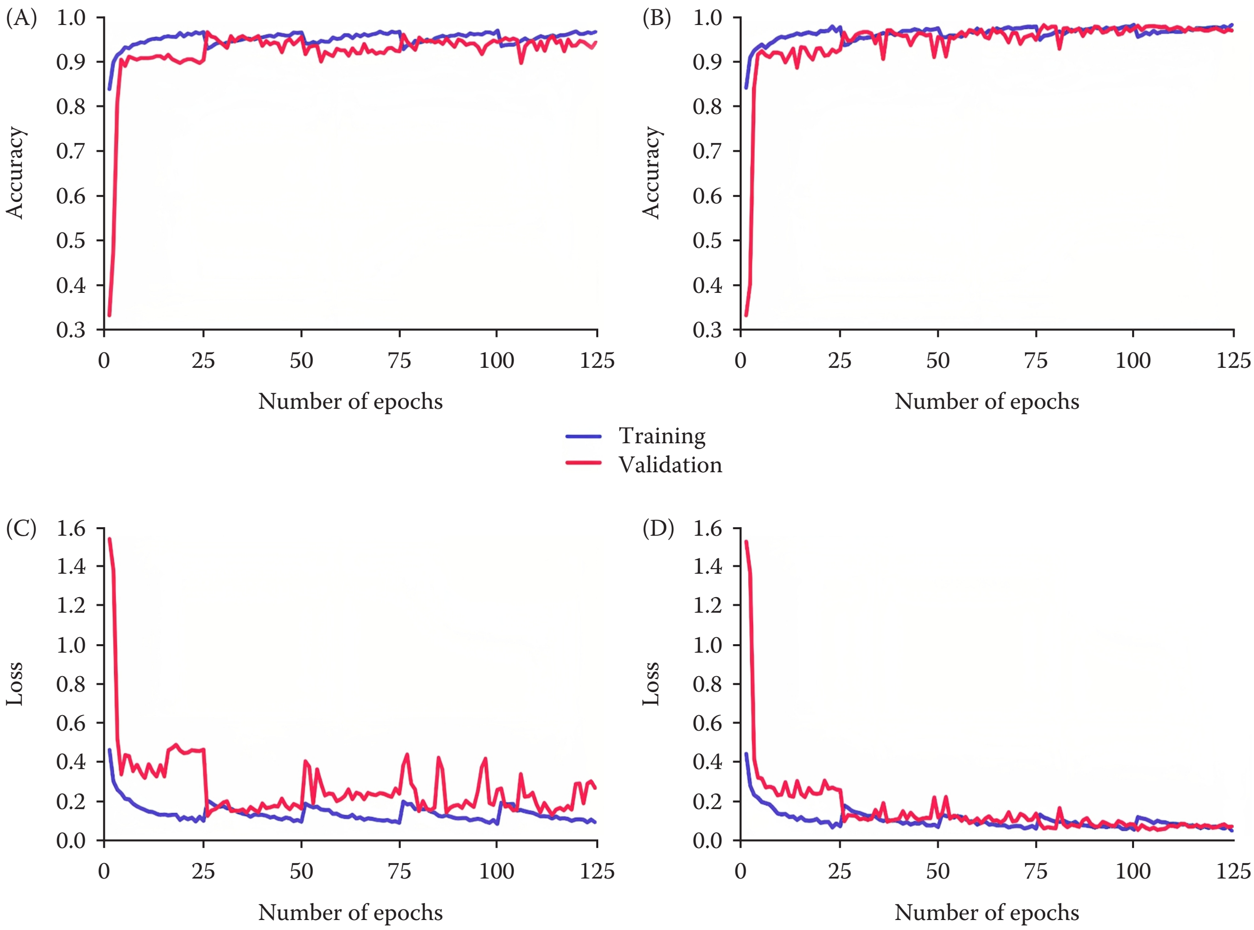


Figure 4. Graph visualisation of performance evaluation of VGG-16 model and proposed model. The accuracy and loss representation of (A), (C) VGG-16 model; (B), (D) proposed model

https://doi.org/10.17221/77/2022-RAE

Table 1. The performance comparison under different light conditions

| Experiment ~~No.~~ | Accuracy (%) | | |
|---|---|---|---|
| | morning period | noon period | afternoon period |
| Beetroot | 94.5 | 94.1 | 94.0 |
| Rice | 93.5 | 93.9 | 94.1 |
| Siam weed/snakeweed | 94.0 | 94.2 | 94.5 |
| Parkinsonia/Chinese apple | 93.1 | 94.0 | 93.8 |
| Mean accuracy | 93.7 | 94.0 | 94.1 |

We validate the performance of our model under different light conditions to show the effectiveness and consistency of our proposed model under different light intensities (conditions). This is one of our exploration experiments and the main reason for the conduction of this experiment is that under certain light conditions, some sample weeds look similar, such as Snakeweed and Chinese apple, so the recognition accuracy may be affected due to shadow size variations and illumination changes. We have captured some additional images using the Pi camera with the same resolution as our adopted dataset. We selected morning time (6:30 to 7:00 a.m.), afternoon time (12:00 to 12:30 p.m.), and evening time (4:00 to 4:30 p.m.) mostly in the summer season of Pakistan and there are variations in light intensity. We capture the images in the morning, noon, and afternoon time. We have captured a total of 240 images including 2 plants (beetroot and rice) and 4 (Parkinsonia, Chinese apple, Siam weed, and snakeweed) that are available in a dataset and can be found in different regions of Pakistan. We captured 40 samples of each with different orientations and light conditions (morning, afternoon, and evening). We utilized 70% for the training of the model and 30% for testing. We observe that the average performance was almost similar and consistent, demonstrating that our model was not affected under different light conditions. Table 1 illustrates the performance of our framework and obtained consistent results with superior classification accuracy.

Figure 5 depicts the confusion matrix of the competitive VGG-16 test phase and the proposed framework (hybrid features) for weeds and plants (beetroot and rice in this experiment) classification. The $x$-axis represents the predicted labels and the $y$-axis denotes the ground truth labels. Our model obtains 95.5% for 2 classes and it is worth noticing that classes (weeds and plants) with similar colour, pattern, texture, and shape are more easily confused with each other. The main reason for misclassification is the similarities of the features and representations among the leaf structure of weeds and plants. In addition, the number of training samples is very small, so the achieved results are a bit confusing and some level of misclassification occurs. In our experiment, among 1 500 images our introduced model has consistent and better true positive and true negative values and lesser test samples are misclassified as compared to the competitive VGG-16 net-

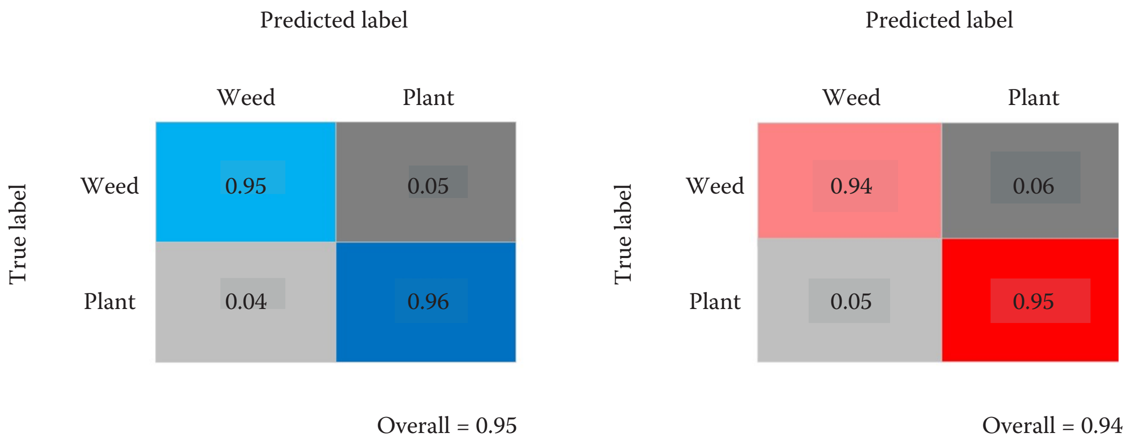


Figure 5. (A) Confusion matrices comparison, (B) proposed method VGG-16

https://doi.org/10.17221/77/2022-RAE

Table 2. Performance comparison of different feature descriptors with classifiers in terms of classification accuracy (%)

| Descriptor | Classifiers | | |
|---|---|---|---|
| | SVM | KNN | XGBoost |
| HOG (Dalal and Triggs 2005) | 69.0 | 66.0 | 70.1 |
| LBP (Ojala et al. 1996) | 71.0 | 68.4 | 74.0 |
| HOG + LBP | 75.4 | 72.0 | 77.8 |
| VGG-16 (Simonyan and Zisserman 2015) | 83.9 | 78.4 | 84.6 |
| HOG + LBP + VGG-16 | 87.7 | 79.7 | 94.5* |

* XGBoost classifier obtained the highest classification accuracy in the presence of all combined features (HOG+LBP+VGG-16); SVM – support vector machine; KNN – K–nearest neighbour; HOG – histogram of oriented gradients; LBP – local binary pattern

work. Our proposed model achieved a classification accuracy of 95.5% as compared to the competitive CNN network with an accuracy of 94.4%.

Next, we compare the performance of different feature descriptors, such as HOG, LBP, and VGG-16, and their possible combinations using three classifiers (on an individual basis) SVM, KNN, and XGBoost. Table 2 reported the average recognition accuracy (%) for each case. We trained the proposed model using individual feature descriptors (HOG, LBP, and VGG-16) and also with their overall possible combinations (as given in the table). Next, we compute the recognition accuracy using three classifiers on an individual basis as exploration results and reported the attained results. According to the obtained results, we attained better results in the case of only VGG-16 as compared to the individual case of HOG and LBP, and also with their combination. However, the combination of all featured descriptors (HOG, LBP, and VGG-16) that contributes to the shape, texture, and discriminative features respectively shows the highest recognition accuracy when applied to XGBoost which is higher than VGG-16 and improved the recognition accuracy.

We further analyse the validation and effectiveness of our proposed model and compare classification accuracy against the models and their results reported by Olsen et al. (2019). We adopted the same distribution of datasets in terms of training, validation, and testing. Results in terms of average classification accuracy (%) are listed in Table 3. We select 8 weeds from the dataset according to their interclass variations and similarities as given in the table. The classification accuracy is observed to vary from weed to weed due to the weeds having less visible features and images given in the dataset overlapped with plants. ResNet-50 obtained better results than Inception-v3. We can also observe that GoogLeNet (Szegedy et al. 2015) achieved almost similar results, however, the proposed model achieved results of 95.5 average classification accuracy exceeding state-of-the-art which is due to the complimentary information from the concatenation of hybrid features.

Table 3. Performance comparison of the proposed model in terms of average classification accuracy (%)

| Name of weed | Inception-v3 (Olsen et al. 2019) | ResNet-50 (Olsen et al. 2019) | GoogLeNet (Szegedy et al. 2015) | Proposed method |
|---|---|---|---|---|
| Chinee apple | 85.3 | 88.5 | 91.7 | 93.5 |
| Lantana | 94.4 | 95.0 | 96.0 | 96.0 |
| Parkinsonia | 96.8 | 97.2 | 95.9 | 97.0 |
| Parthenium | 94.9 | 95.8 | 95.6 | 96.1 |
| Prickly acacia | 92.8 | 95.5 | 94.7 | 95.9 |
| Rubber vine | 93.1 | 92.5 | 94.8 | 94.5 |
| Siam weed | 97.6 | 96.5 | 96.9 | 97.0 |
| Snake weed | 88.0 | 88.8 | 94.6 | 94.4 |
| Average accuracy | 92.8 | 93.7 | 95.0 | 95.5* |

* The proposed method achieved the highest average accuracy as compared to state-of-the-art deep learning-based model

https://doi.org/10.17221/77/2022-RAE

## CONCLUSION

The proposed automated method addresses the problem of inaccurate recognition of weeds in agriculture. In the weed detection process, the extraction of texture, shape, and complex patterns is very crucial as weeds are very similar to plants. Very few research studies are available in the existing literature that successfully combines the features from a different channel with better results. This research work intelligently combines the features extracted from different domains, such as handcrafted and deep learning, and extracts the valid and relevant features that lead to high weed recognition accuracy. The introduced method achieved 97% training accuracy and 95.5% testing accuracy in predicting weeds. We performed different exploration experiments using concatenated features in the presence of different state-of-the-art classifiers and the XGBoost classifier secured the highest weed classification accuracy of over 95%. Our attained results witnessed that hybrid features extracted from the different channels/domains can really boost weed classification accuracy as illustrated by obtained results. In future work, we will examine the performance of our method using a large dataset. We will also enhance the capabilities of our proposed models in case of weeds and plants overlapping and occlusion. We can also consider the detection rate (speed) to analyse the computational speed (frames processed/s) of our proposed model and compare it with other state-of-the-art existing models.

## REFERENCES

Arif S., Wang J., Hussain F., Fei Z. (2019): Trajectory-based 3D convolutional descriptors for human action recognition. Journal of Information Science and Engineering, 35: 851–870.

Aversano L., Bernardi M.L., Cimitile M., Iammarino M., Rondinella S. (2020): Tomato diseases classification based on VGG and transfer learning. In: Proceedings of the 2020 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), Trento, Nov 4–6, 2020: 129–133.

Bakhshipour A., Jafari A. (2018): Evaluation of support vector machine and artificial neural networks in weed detection using shape features. Computers and Electronics in Agriculture, 145: 153–160.

Chen T., Guestrin C. (2016): XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, Aug 13–17, 2016: 785–794.

Chen Y., Wu Z., Zhao B., Fan C., Shi S. (2021): Weed and corn seedling detection in field based on multi feature fusion and support vector machine. Sensors, 21: 212–230.

Cover T., Hart P. (1967): Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13: 21–27.

Dalal N., Triggs B. (2005): Histograms of oriented gradients for human detection. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, June 20–25, 2005: 886–893.

European Commission (2021): Farm to Fork targets – Progress. Available at https://food.ec.europa.eu/plants/pesticides/sustainable-use-pesticides/farm-fork-targets-progress_en (accessed Sept 6, 2021).

Fu L., Gao F., Wu J., Li R., Karkee M., Zhang Q. (2020): Application of consumer RGB-D cameras for fruit detection and localization in the field: A critical review. Computers and Electronics in Agriculture, 177: 105687.

Hamuda E., Mc Ginley B., Glavin M., Jones E. (2017): Automatic crop detection under field conditions using the HSV colour space and morphological operations. Computer and Electronics in Agriculture, 133: 97–107.

He D., Qiao Y., Li P., Gao Z., Li H., Tang J. (2013): Weed recognition based on SVM-DS multi-feature fusion. Transactions of the Chinese Society of Agricultural Machinery, 44: 182–187.

He K., Zhang X., Ren S., Sun J. (2016): Deep Residual Learning for Image Recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, June 27–30, 2016: 770–778.

Janiesch C., Zschech P., Heinrich K. (2021): Machine learning and deep learning. Electronic Markets, 31: 685–695.

Kiefer J., Wolfowitz J. (1952): Stochastic estimation of the maximum of a regression function. The Annals of Mathematical Statistics, 23: 462–466.

Krizhevsky A., Sutskever I., Hinton G.E. (2012): ImageNet classification with deep convolutional neural networks. In: Proceedings of the 26th Conference on Neural Information Processing Systems (NIPS 2012), Lake Tahoe, Dec 3–8, 2012: 1097–1105.

Le V.N.T., Apopei B., Alameh K. (2019): Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods. Information Processing in Agriculture, 6: 116–131.

Le V.N.T., Ahderom S., Alameh K. (2020): Performances of the LBP-based algorithm over CNN models for detecting crops and weeds with similar morphologies. Sensors, 20: 2193.

https://doi.org/10.17221/77/2022-RAE

Lowry C.J., Smith R.G. (2018): Weed control through crop plant manipulations. In: Jabran K., Chauhan B.S. (eds): Non-Chemical Weed Control. London, Academic Press: 73–96.

Ma Y., Feng Q., Yang M., Li M. (2016): Detection of wine grape leaves based on HOG. Computer Engineering and Applications, 52: 158–161.

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

Olsen A., Konovalov D.A., Philippa B., Ridd P., Wood J.C., Johns J., Banks W., Girgenti B., Kenny O., Whinney J., Calvert B., Azghadi M.R., White R.D. (2019): DeepWeeds: A multiclass weed species image dataset for deep learning. Scientific reports, 9: 2058.

Pan S.J., Yang Q. (2010): A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering, 22: 1345–1359.

Pizer S.M., Johnston R.E., Ericksen J.P., Yankaskas B.C., Muller K.E. (1990): Contrast-limited adaptive histogram equalization: Speed and effectiveness. In: Proceedings of the 1st Conference on Visualization in Biomedical Computing, Atlanta, May 22–25, 1990: 337–345.

Razfar N., True J., Bassiouny R., Venkatesh V., Kashef R. (2022): Weed detection in soybean crops using custom lightweight deep learning models. Journal of Agriculture and Food Research, 8: 100308.

Sabzi S., Abbaspour-Gilandeh Y., García-Mateos G. (2018): A fast and accurate expert system for weed identification in potato crops using metaheuristic algorithms. Computers in Industry, 98: 80–89.

Simonyan K., Zisserman A. (2015): Very deep convolutional networks for large-scale image recognition. In: Proceedings of the 3rd International Conference on Learning Representations (ICLR), San Diego, May 7–9, 2015: 1–14.

Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. (2015): Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, June 7–12, 2015: 1–9.

Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z. (2016): Rethinking the inception architecture for computer vision. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, June 27–30, 2016: 2818–2826.

Tang J., Wang D., Zhang Z., He L., Xin J., Xu Y. (2017): Weed identification based on K-means feature learning combined with convolutional neural network. Computers and Electronics in Agriculture,135: 63–70.

Vapnik V. (1999): The Nature of Statistical Learning Theory, Statistics for Engineering and Information Science. 2nd Ed. Berlin, Springer Science & Business Media.

Wang P., Tang Y., Luo F., Wang L., Li C., Niu Q., Li H. (2022): Weed25: A deep learning dataset for weed identification. Frontiers in Plant Science, 13: 1053329.

Wu Z., Chen Y., Zhao B., Kang X., Ding Y. (2021): Review of weed detection methods based on computer vision. Sensors, 21: 3647–3670.

Yang J., Wang Y., Chen Y., Yu J. (2022): Detection of weeds growing in alfalfa using convolutional neural networks. Agronomy (MDPI), 12: 1459.

You J., Liu W., Lee, J. (2020): A DNN-based semantic segmentation for detecting weed and crop. Computers and Electronics in Agriculture, 178: 105750.

Zimdahl R. (2018a): Weed Classification, in Fundamentals of Weed Science, 5th Ed. In: Zimdahl (ed.). Cambridge, Academic Press: 47–60.

Zimdahl R. (2018b): Weed Classification, in Fundamentals of Weed Science, 5th Ed. In: Zimdahl (ed.). Cambridge, Academic Press: 17–46.

Zou K., Chen X., Wang Y., Zhang C., Zhang F. (2021): A modified U-Net with a specific data argumentation method for semantic segmentation of weed images in the field. Computers and Electronics in Agriculture, 187: 106242.

Received: September 30, 2022
Accepted: February 21, 2023
Published online: July 1, 2023