IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 17, 2020, accepted January 20, 2021, date of publication February 2, 2021, date of current version February 10, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3056577

Identification of Tobacco Crop Based on Machine Learning for a Precision Agricultural Sprayer

MUHAMMAD TUFAIL[1,2], (Member, IEEE), JAVAID IQBAL[3], (Member, IEEE), MOHSIN ISLAM TIWANA[3], (Member, IEEE), MUHAMMAD SHAHAB ALAM[1], ZUBAIR AHMAD KHAN[2], AND MUHAMMAD TAHIR KHAN[1,2]

[1]Advanced Robotics and Automation Laboratory, National Center of Robotics and Automation, Peshawar 25000, Pakistan
[2]Department of Mechatronics Engineering, University of Engineering and Technology, Peshawar 25120, Pakistan
[3]Department of Mechatronics, National University of Science and Technology, Islamabad 24090, Pakistan

Corresponding author: Muhammad Tufail (tufail@uetpeshawar.edu.pk)

This work was supported by the Higher Education Commission of Pakistan through the National Center of Robotics and Automation, Pakistan.

ABSTRACT Agrochemicals, which are very efficacious in protecting crops, also cause environmental pollution and pose serious threats to farmers' health upon exposure. In order to cut down the environmental and human health risks associated with agrochemical application, there is a need to develop intelligent application equipment that could detect and recognize crops/weeds, and spray precise doses of agrochemical at the right place and right time. This paper presents a machine-learning based crop/weed detection system for a tractor-mounted boom sprayer that could perform site-specific spraying on tobacco crop in fields. An SVM classifier with a carefully chosen feature combination (texture, shape, and color) for tobacco plant has been proposed and 96% classification accuracy has been achieved. The algorithm has been trained and tested on a real dataset collected in local fields with diverse changes in scale, orientation, background clutter, outdoor lighting conditions, and variation between tobacco and weeds. Performance comparison of the proposed algorithm has been made with a deep learning based classifier (customized for real-time inference). Both algorithms have been deployed on a tractor-mounted boom sprayer in tobacco fields and it has been concluded that the SVM classifier performs well in terms of accuracy (96%) and real-time inference (6 FPS) on an embedded device (Raspberry Pi 4). In comparison, the customized deep learning-based classifier has an accuracy of 100% but performs much slower (0.22 FPS) on the Raspberry Pi 4.

INDEX TERMS Crop and weed detection, machine-learning, precision agriculture.

I. INTRODUCTION

Tobacco, considered as a cash crop in many countries (e.g., China, Brazil, India, and Pakistan), is a highly agrochemicals dependent crop. According to Food and Agriculture Organization (FAO) of the United Nations, approximately 6 million metric tons of tobacco was produced on approximately 5 million hectares of land worldwide (in over 125 countries) in 2018. Huge amounts of agrochemicals are applied at various times on tobacco over its three-month growing season to get maximum yield. It has drastic environmental impact when compared to other agricultural cash crops and causes health and socioeconomic problems for populations specially in low-income and middle-income countries. Even in the United States, approximately 27 million pounds of different agrochemicals (which included different insecticides, fungicides, and suckercides) were applied on tobacco from 1994 to 1998 [1].

Tobacco plants are mostly affected by pests and diseases, which necessitates the application of large quantities of pesticides on it. The tobacco industry considers pesticides significantly essential for the economical production of tobacco [3], [4]. Therefore, the heavy reliance on pesticides for controlling pests in tobacco fields exacerbates the risk of farmers' exposure to pesticides. Tobacco farmers in Pakistan fail to protect themselves from getting exposed to pesticides primarily for two reasons: firstly they are unaware of the risks associated with pesticide exposure due to lack of information, and secondly they still use the conventional rudimentary pest/weed control techniques due to lack of availability

The associate editor coordinating the review of this manuscript and approving it for publication was Hao Ji.

FIGURE 1. Intra-Row Spacing Between Tobacco Plants.

of advanced application equipment. The environmental and human health risks associated with agrochemical application are mainly a result of imprecise spraying i.e. over-application or uniform broadcast spraying indiscriminately throughout the entire field. Therefore, there is a need to develop intelligent application equipment that could detect plants/weeds and precisely spray at the right place and right time.

Tobacco plants are commonly grown in straight rows, and the intra-row spacing between the plants ranges from 0.5 to 1 meter, as shown in Figure 1. Performing broadcast spraying indiscriminately on the entire tobacco field, especially at the early growth stages, results in unnecessary off-the-target application on bare soil between any two consecutive plants, which therefore accounts for environmental pollution and pesticide leaching into the ground.

Hence, for conducting selective/site-specific pesticide spray only on tobacco plants, the application equipment must have the ability to detect and classify tobacco plants from weeds, know where they are located inside the rows, and then accordingly spray only on the targets i.e. tobacco plants.

A major problem encountered in the implementation of vision-based site-specific spraying is the accurate detection and classification of crop plants and weeds. To address the challenge of precise spraying, considerable efforts have been made by the research community for developing scientific and technological solutions for spot/site-specific application of agrochemicals, see for example [2]–[7]. All these solutions have focused on the use of techniques that automatically senses plants and weeds in real-time, and sprays agrochemicals in a targeted style, that too only when necessary. A state of the art plant-weed sensing and monitoring solutions can be found in [8]. Several surveys on machine vision and imaging processing techniques i.e. pre-processing, segmentation, feature extraction, and classification for plant and weed detection are available in [9]–[13].

The work presented here is part of a larger project aimed to design, fabricate, and test a precision agricultural sprayer for crops and orchards. The project has been carried out at the Advanced Robotics and Automation Laboratory, one of the 11 affiliated labs with Pakistan's National Center of Robotics and Automation. This paper focuses on the development of a machine learning-based tobacco crop/weed detection and classification system for site-specific spraying.

The main objectives of this research work can be summarized as follows:

1) Development of a vision-based learning model for the detection and classification of crops and weeds.
2) Selecting the best machine learning algorithm for accurately predicting/detecting tobacco crop and weeds. Performance of the selected algorithm must be tested on real datasets collected from farm fields under natural conditions.
3) Integration of the learning-based vision model in the tractor-mounted spraying system for autonomously performing spot/site-specific spraying.

This paper is divided into seven sections. Section 2 presents a critical assessment of related work available in the literature. A description of the steps involved in vision-based classification problems is presented in Section 3. Experimental hardware of the spraying system is briefly discussed in section 4. Results and discussion of traditional machine-learning algorithms applied to the tobacco and weeds classification problem, the development framework, and real-time performance comparison on different target machines are presented in Section 5. Section 6 concludes the paper.

II. STATE OF THE ART

Significant attention has been paid by the scientific community to the development of precise and targeted spraying solutions for reducing pesticide/herbicide inputs so as to cut-down their side effects [14]–[22]. Discerning the crop field scenario i.e. understanding the crop/plant requirements or finding where weeds are located in field is an essential key component of intelligent precision spraying system. Computer vision, a technology that enables a machine to automatically analyze and understand the visual world, has been immensely improved over the past few years and is therefore now widely being used in many precision agriculture problems. Numerous vision-based crop and weed detection techniques, in the aforementioned context, have been proposed [22]–[31]. A brief summary on the previous work done on vision-based crop/weed detection and classification is presented in Table 1.

Despite the significant progress in the field of computer vision, there still exist limited robust vision-based crop-weed detection techniques that perform well in complex scenarios such as fields densely populated with weeds, crop plants that highly resemble with weeds, varying illumination levels, different soil textures, and varying shape/color features at different growth stages.

Most of the studies conducted thus far on crop/weed detection are not experimentally validated on tough and challenging real-world datasets like the one presented here in this paper. This research study proposes and evaluates the application of SVM classifier for the detection tobacco plants and weeds in real field scenarios.

TABLE 1. Summary of Vision-based crop/weeds detection and classification.

Reference	Crop	Method/Approach	Sensor	Features	Accuracy	Limitation
[32]	Wheat, maize, sunflower	Region-based segmentation method using a blob colouring analysis	RGB color camera	Normalized green channel, spatial similarity	87%-100%	Not tested in real-time
[33]	Cabbage and carrots	Fuzzy logic approach	CCD camera	Morphological and color features	88% for cabbage and 72% for carrots	Unable to separate the plants that grow together
[34]	Cotton	Plant recognition algorithm, based on a minimum distance function (MDF)	CCD camera coupled with a 20 mm focal length lens	Size	82.75%	Requires a controlled illumination chamber
[35]	Wheat	Discriminant analysis classifier	Optical sensor (phototransistors)	Color indices	70%	Low accuracy
[36]	Corn	Bayesian classifier	Digital camera	Size and area	Not reported	Not tested outside in field environments
[37]	Lettuce	Multi-waveband Bayesian classification model	Hyperspectral camera	RVI and NDVI indices	90.3%	Sensitive to seasonal variation in optical properties of lettuce
[38]	Cotton	Support vector machines (SVM)	Digital camera	Shape features and Hu moments	94%	Not tested in field
[39]	Maize	Fast Image Processing (FIP) and Robust Crop Row Detection (RCRD)	Digital video camera	Shape and size Detects	95% of weeds and 80% of crop	Performs well only under uncontrolled lighting in real-time
[40]	Tobacco	Fuzzy comprehensive evaluation (FCE)	Color camera	Color, size, shape and surface	94% for trained and 72% non-trained tobacco leaves	Tested only for grading tobacco leaves
[41]	Tobacco	K-Nearest Neighbor (K-NN) classifier	Digital camera	Textural features	70-80%	Tested only for grading tobacco leaves
[42]	Spinach and chard crops	Thresholding classifier	RGB color camera	Color and size	90%	Cannot estimate the exact plant positions
[43]	Sugar beets	Random forest classifier and Markov Random Field	4-channel camera (RGB+NIR)	Appearance and geometric features	96%	Performance decreases when the appearance of the plants changes
[44]	Sugar beets	CNN classifier	RGB+NIR camera	No hand crafted features	97%	Only for early growth stages, since the approach is not able to deal with overlapping plants
[45]	Maize	Bayesian classification approach	Color camera	Mahalanobis distance	84-91%	Computationally expensive and works only in initial growth stages
[46]	Romaine Lettuce	Crop signaling technique	RGB camera	Does not rely on size, shape or spatial pattern	98.93%	A compound is to be applied to the crop plants for generating optical signals
[47]	Corn	Decision Trees	CCD Camera	Combination of spectral, shape and textural features	95%	Applicability to actual field conditions not tested
[48]	Carrot	Random Forest classifier	Multispectral down-looking monocular camera	Shape, contour, and statistical features	93.8%	Intra class overlapping plants cannot be split into different plant regions
[49]	Tobacco	Convolutional Neural Network (CNN)	DSLR camera	Haar-like features	96.25%	Tobacco grading
[50]	Rice	Support vector machines and random forest classifier	Digital color cameras	Texture, color, and shape features	91.36%	Not tested in actual fields
[51]	Sugar beets	SVM and ANN	RGB camera	Shape features	92.67% by ANN and 93.33% SVM	Only shape features were used, other features such as color and texture were not tested

FIGURE 2. Sample images from ARAL tobacco dataset.

III. METHODOLOGY OF MACHINE LEARNING-BASED IMAGE CLASSIFICATION

The following sub-sections describe the methodology to develop the proposed classification solution for tobacco and weeds.

A. COLLECTION OF DATASET

Acquiring a pertinent dataset is the most tedious yet crucial and important task in computer vision applications.

The process of data collection becomes even more challenging and expensive in the agricultural domain because dedicated fields are needed so that a large number of RGB, or in some cases near-infrared (>720mm), or red light (620-680mm) spectral images of crops/weeds can be gathered at different growth stages and under various lightening conditions.

In the work presented here, a dataset was created that consisted of images of weeds and tobacco plants captured in Pakistan in the tobacco fields at Swabi, Khyber Pakhtunkhwa (34°09'07.3"N 72°21'36.2"E). The images were taken at early growth stages at various day timings and different lighting conditions using digital color cameras. The images in the dataset were categorized into: (1) tobacco, and (2) weeds. Each category had 97 images that were used to train different machine learning models for the detection and classification tasks. For the deep learning algorithm, the dataset comprised of 2200 labeled images of tobacco and weeds. Figure 2 shows sample images from the aforementioned dataset.

B. IMAGE PRE-PROCESSING

In image pre-processing, several image enhancement techniques, including normalization, conversion to binary or grayscale, alignment and resizing, noise removal, and contrast/sharpness enhancement, etc., are used for improving raw images.

In the given application, the most important pre-processing step was to segment the green objects (tobacco and weeds) in images from the background (soil). This was performed by converting the RGB images into the CIE L*a*b color system which works like human perception. L*a*b is a uniform three-dimensional real number space with the three dimensions (channels) being: L* (luminosity from black to white); a* (hue along green to red); and b* (saturation along blue to yellow).

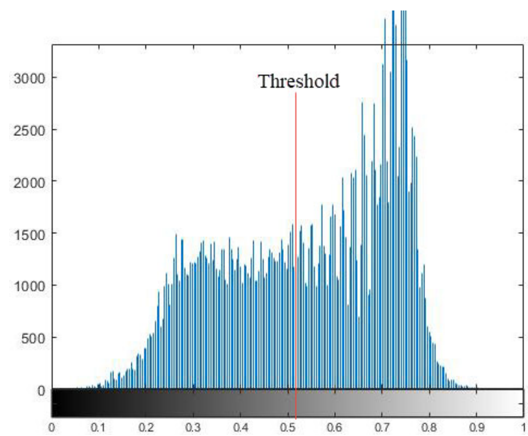

FIGURE 3. Image histogram (a* channel).

For the given dataset, segmentation is performed using the Otsu's method for which the threshold is computed using the a* channel due to its bimodal nature (Figure 3). This was not possible in the RGB space. The resulting segmented image is shown in Figure 4 with soil being automatically removed.

FIGURE 4. Segmented image.

C. FEATURE SELECTION AND EXTRACTION

The feature selection and extraction step is the most important step that involves the extraction of some selected features, i.e. vital information, from the pre-processed images. The most common features chosen in image classification problems are (a) geometric features, such as shapes, contours, lines, edges, points, and (b) color features, such as intensity and color histograms. When choosing features for the tobacco/weeds classification problem, the best candidates to consider were color, texture, or shape features. The following subsections present justification and description of the selected features.

1) TEXTURE FEATURES

Texture in an image is a measure of spatial distribution of intensities in a neighborhood.

FIGURE 5. Example image for texture segmentation.

It was observed that tobacco images have a repeating pattern of local variations in intensities. This can be demonstrated by applying Gabor filters to an example image (Figure 5). The Gabor filter which is a group of Gabor wavelets automatically determines the boundaries between tobacco and non-tobacco objects (weeds) based on their texture characteristics. The extracted Gabor texture features are input to a k-means clustering algorithm. This classifies textured region of tobacco from other texture classes (weeds) as shown in Figure 6. It is evident from Figure 6-(a) that the tobacco plant has prominent texture features as compared to the surrounding objects.

Other statistical methods to characterize texture include Gray Level Co-occurrence Matrix (GLCM) or gray-level spatial dependence matrix, contrast, entropy, and homogeneity etc. These methods give a set of measurements (feature vector) of texture in a region.

In order to compute the Haralick texture features for an image, the GLCM, **P**, is first computed which is a two-dimensional array that stores the number of pairs of pixels with similar gray levels.

$$\mathbf{P}_d[i, j] = n_{ij} \tag{1}$$

where n_{ij} is the number of occurrences of pairs of gray levels (i, j) at offset $\mathbf{d} = (d_x, d_y)$ in the image. The elements of $\mathbf{P}_d[i, j]$ are normalized to lie between 0 and 1 and, therefore, can be interpreted as probabilities. A variety of numeric features [52] are extracted from the GLCM to quantify the image texture characteristics.

For the given dataset, the offset pair $\mathbf{d} = (0, 1)$ was used which means the horizontal proximity of the pixels (i.e., pixel next to the pixel of interest in the same row) was considered to calculate second order statistics based on a symmetrical co-occurrence matrix. The following metrics were computed and used in the classification: angular second moment (Energy), contrast, correlation, variance, inverse difference moment (homogeneity), sum average, sum variance, entropy, different variance, difference entropy, information measure of correction I and II, and maximal correlation coefficient.

2) SHAPE FEATURES

Geometric moments are statistical features. For a given image function $f(x, y)$, geometric moments of order $(p + q)$ are calculated as,

$$m_{pg} = \sum_{x=1}^{N}\sum_{y=1}^{N} f(x, y)(x)^p(y)^q \tag{2}$$

For a region of interest, m_{00} is the area (calculated from a binary image). From this, the center of gravity is calculated as $\bar{x} = m_{10}/m_{00}$ and $\bar{y} = m_{01}/m_{00}$.

The central moments which are translation-invariant are then computed as,

$$\mu_{pq} = \sum_{x=1}^{N}\sum_{y=1}^{N} f(x, y)(x - \bar{x})^p(y - \bar{y})^q \tag{3}$$

Scale-invariance is achieved by,

$$v_{pg} = \frac{\mu_{pq}}{\mu_{00}^{(1+(p+q)/2)}}, \quad p + q \geq 2 \tag{4}$$

Rotation-invariance is achieved by,

$$\phi_1 = v_{20} + v_{02} \tag{5}$$
$$\phi_2 = (v_{20} - v_{02})^2 + v_{11}^2 \tag{6}$$

Invariance to general linear transformations is achieved by,

$$I_1 = \mu_{20}\mu_{02} - \mu_{11}^2 \tag{7}$$

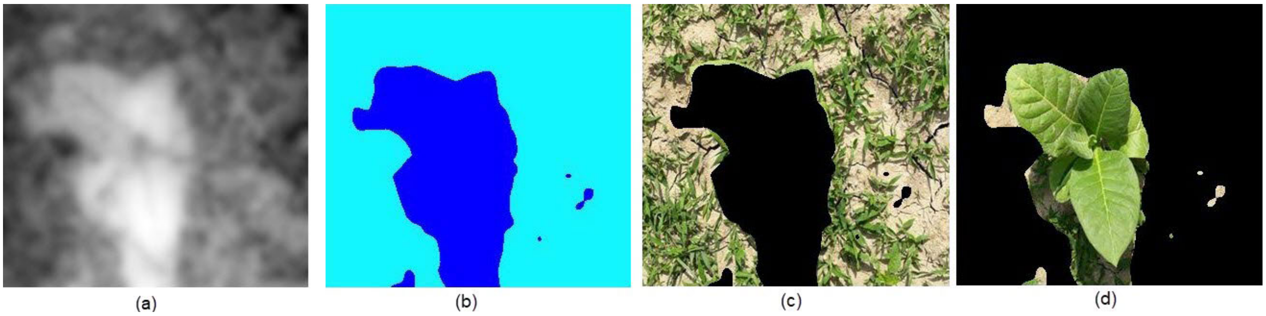

FIGURE 6. Texture-based segmentation using Gabor filters (Orientation between [0, 135] degrees in steps of 45 degrees).

$$I_2 = (\mu_{30}\mu_{03} - \mu_{21}\mu_{12})^2 - 4(\mu_{30}\mu_{12} - \mu_{21}^2)(\mu_{21}\mu_{03} - \mu_{12}^2) \quad (8)$$

$$\psi_1 = \frac{I_1}{\mu_{00}^4} \quad (9)$$

3) COLOR FEATURES

For color features extraction, color moments are used which provide distinctive features to differentiate objects based on their color. The basic idea of color moments is based on probability distribution of image intensities which can be characterized by moments such as mean, variance, and skewness. These three are the central moments of intensity distribution and can be easily found for all color spaces such as RGB, HSV, and L*a*b. In this work, central moments of input images (in L*a*b color space) are found giving nine moments in total. These are defined as follows [53]:

$$E_i = \sum_{j=1}^{N} \frac{1}{N} p_{ij} \quad (10)$$

$$\sigma_i = \sqrt{\left(\frac{1}{N}\sum_{j=1}^{N}(p_{ij} - E_i)^2\right)} \quad (11)$$

$$s_i = \sqrt[3]{\left(\frac{1}{N}\sum_{j=1}^{N}(p_{ij} - E_i)^3\right)} \quad (12)$$

where i is the color channel, p_{ij} is the j-th image pixel of i-th color channel and $N = m \times n$ (m, n being the number of columns and rows in the image, respectively).

D. CLASSIFICATION

The dataset, comprising of images, was labeled and split into two sub-sets, namely, training and testing datasets. The training dataset was used for training the classification model, and the testing dataset was used for verifying the accuracy of the classifier.

The training data is formulated as:

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n). \quad (13)$$

where x represents a feature vector $[x_1, x_2, \ldots, x_m]^T$ and y represent the class (desired output) to which the input feature vector belongs.

The selection of an apt classifier that is best suited to the nature of data and application/problem being targeted is a real challenge. For the data formulated in the previous step, the classification algorithm (also called regression) aims at learning a function f which approximates the training data in the best possible manner i.e.,

$$y_i = f(x_i) \quad \forall i \in 1, \ldots, n. \quad (14)$$

The algorithm finds regularities in patterns of the labeled training data and can be classified using various approaches such as those based on the type of learning (supervised vs. unsupervised), assumption on the distribution of data (parametric vs. non-parametric), and modeling technique (statistical vs. neural network approaches). The output of this step in the given context is to distinguish plants from weeds as well as classify weeds further based on their species (single-leaf or whole plant) and their leaf size (i.e. broad-leaf or narrow-leaf).

Lastly, in order to evaluate the accuracy and generalization ability of the algorithm, all possible sources of errors in classification are identified. Labeled test data is used for evaluating the learned function. Once the errors in classification results are minimized to a certain threshold, the algorithm is deployed and exposed to unseen data.

In deep learning based object detection, a Convolutional Neural Network (CNN) is used which is a specialized type of deep Artificial Neural Networks (ANN) and is widely used as a state-of-the-art in computer vision. The architecture of a CNN consists of an array of layers of neurons such as convolutional layers, max-pooling or average-pooling layers, and fully-connected layers. Following here, the steps involved in the deep learning approach have been explained.

Initially, an image as a 3D object (with length, width, and height being the dimensions) is input to the neurons in the first layer. Subsequent layers are divided into two general categories (Figure 7):

1) Feature extraction layers: Inspired by the working of the biological visual cortex, only sub-regions or small patches of the input image (also called receptive fields)

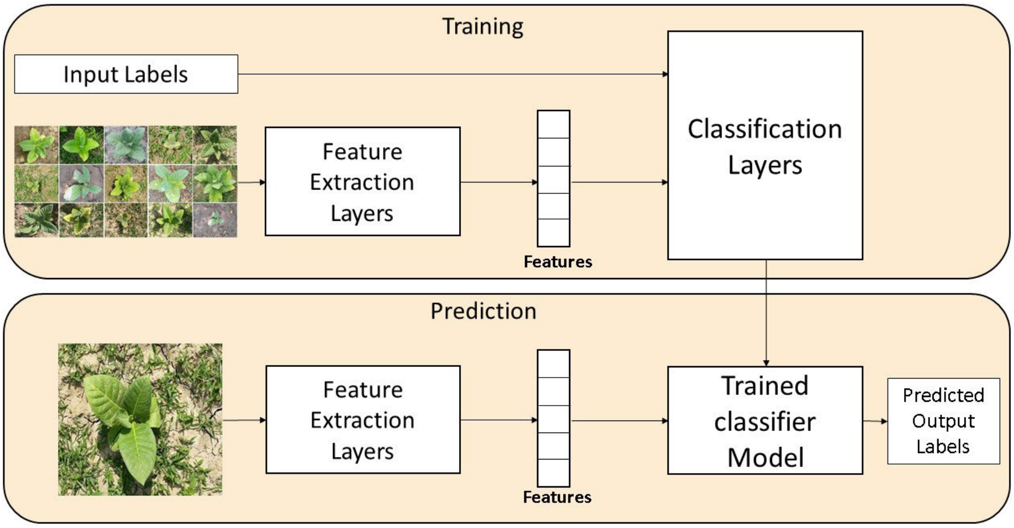

FIGURE 7. Deep learning framework for object detection.

connect to the neurons in the convolutional layer which applies sliding convolutional filters (learned automatically) to the input to generate feature maps. The neurons in the hidden layers perform further convolution operations and sub-sampling (pooling) to extract (learn) feature hierarchies which are the intermediate representations of the input data and are helpful to discriminate amongst object classes.
The feature extraction network is typically a pre-trained CNN such as AlexNet, VGG16, Residual Network (ResNet) 18, MobileNet v2, GoogLeNet and others. With minor modifications, these architectures can be used for new predictive modeling applications. For the current problem, ResNet18 was selected for feature extraction due to its low requirement for lower computation resource and training time.

2) Classification layers: The extracted features (local information) are then input to the units in one or more fully-connected layers which multiplies them by a weight matrix and adds to them a bias vector to discover underlying patterns specific to the dataset. These feature vectors then serve as inputs to the output layers (softmax and classification) that computes the output classification probabilities.

A log loss between the predicted class labels and the ground-truth class labels is used as an evaluation metric for the classification model. The deep learning algorithm iteratively improves the classification accuracy (thereby reducing loss) using an optimization algorithm such as Stochastic Gradient Descent with Momentum (SGDM) to reach at a trained model that has an optimized set of parameters (i.e., weights and biases).

IV. FLUID FLOW CONTROL OF THE PROTOTYPE SPRAYER

The fluid flow control system of the sprayer is represented via a block diagram shown in Figure 8, and the prototype model of the boom sprayer developed is shown in Figure 9. The flow control system comprised of a controller and essential hardware (sensor and actuators etc.) for controlling the application rate according to the field requirements. A fixed displacement diaphragm pump (PRO-PUMPS DP-80, 12V DC) driven by an electric DC motor was used to deliver the agrochemical at 5.5 liters/min and at a desired pressure of 60 psi to the four flat fan nozzles mounted on the boom. Each nozzle was controlled individually via ON/OFF solenoid valves which thereby allowed for more accurate application of agrochemicals. Application rate through the nozzles was controlled by changing duty cycles of the PWM solenoid valves (UNI-D UW-15, 12V DC) according to the feedback (reference) signal obtained from the vision system. A 100% duty cycle i.e. a fully ON signal resulted in maximum flow rate whereas a zero percent duty cycle i.e. an OFF signal resulted in no flow.

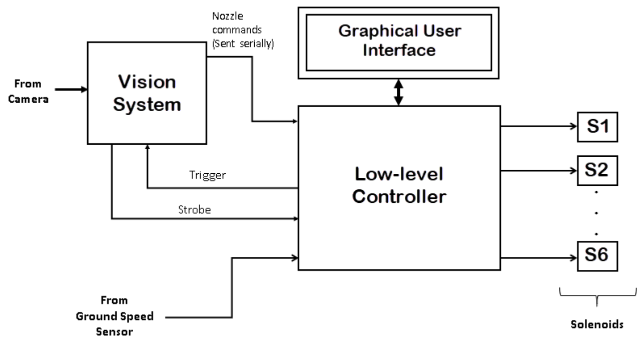

FIGURE 8. Block diagram of precision agricultural sprayer.

Information about flow rate and pressure in the system was provided to the control unit by two flow meters (YS-201 Hall Effect Flow Meter) and a pressure sensor (WIKAI Type A-10, 15 Bar). An electronic proportional control valve (BURKET 1094, 24V DC, 4 – 20mA) was used to maintain the desired pressure in the system. When the pressure in the system

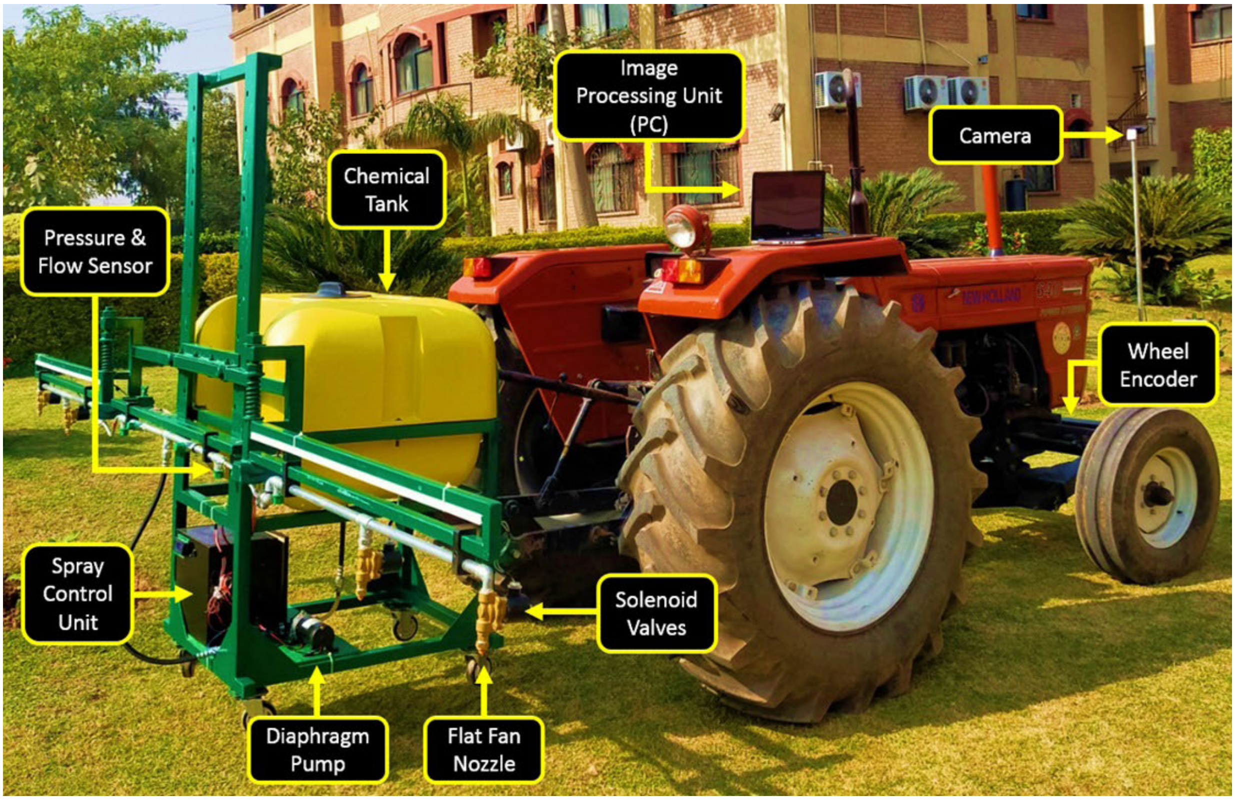

FIGURE 9. Precision agricultural sprayer.

exceeded the set pressure due to closure of any outlet/nozzle, the EPV mounted on the by-pass return line regulated the excess flow back to the tank. An additional line controlled by a solenoid valve was used for agitation purposes. The pump's flow output and pressure were set greater than the boom's flow and pressure requirement in order to compensate for the additional flow required for agitation and line pressure losses. Information about the ground speed of the tractor was obtained via a rotary incremental encoder (OMRON E6H-CWZ6C, 1000 P/R) mounted on the front wheel of the tractor.

FIGURE 10. Vision system (Raspberry Pi 4, Intel Movidius Neural Compute Stick 2, and PiCam).

V. RESULTS AND DISCUSSION

A. DESCRIPTION OF THE DEVELOPMENT FRAMEWORK

For real-time inference the commonly used portable devices include the NVIDIA Jetson Nano, Jetson TX1, and Raspberry Pi 4. For the implementation of this work, a Python virtual environment in Python 3 was setup on a Raspberry Pi 4 (Model B with 4 GB RAM) which had the Raspbian Stretch Linux distribution running on it. The vision system also included a Raspberry Pi camera and a Vision Processing Unit (VPU) i.e., Intel Movidius Neural Compute Stick (NCS) as shown in Figure 10. The Raspberry Pi camera module version 2.0 has an 8-megapixel Sony IMX219 sensor, a physical sensor of size 3.6×2.76 mm, a resolution of 3280×2464, and a frame rate of 60 FPS (at 1080p). The camera was mounted at a height of 1.8 meters facing downward. The popular open source neural network API Keras was used which runs on top of the default Google TensorFlow backend engine. Different reasons to choose Keras include its modularity, speed, and ease of use.

For comparison, image features were learned using a pre-trained deep neural network i.e., convolutional neural network (CNN) on a laptop with MATLAB Deep Learning and Statistics and Machine Learning Toolboxes 2020b. To extract image features, the pre-trained deep network ResNet18 architecture [54] was used which is pre-trained on ImageNet. Like other networks, ResNet18 extracts features in a hierarchical fashion where the network's layers extract low-level features followed by the extraction of high-level features by deep

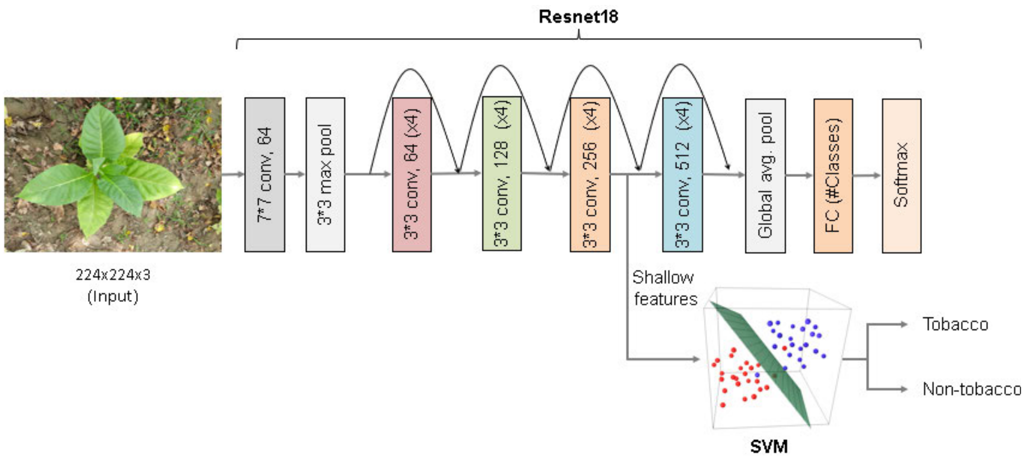

FIGURE 11. SVM with shallow features from ResNet18.

TABLE 2. Accuracy during training (cross validation).

True Class		Predicted Class: Tobacco	Predicted Class: Weeds
	Tobacco	340	24
	Weeds	6	354
		98.4%	93.6%

TABLE 3. Accuracy during prediction.

True Class		Predicted Class: Tobacco	Predicted Class: Weeds
	Tobacco	47	1
	Weeds	2	44
		95.9%	97.8%

hidden layers. For the task at hand, ResNet18 was modified by replacing its last (output) layer with an SVM classifier. Originally the last layers of a CNN network outputs probabilities of each label for a given input image. These probabilities are calculated by an activation function based on outputs of the previous (hidden) layers, weights, and a bias. For an SVM classifier, these probabilities can be treated as features and are used for training.

The generalization ability of a CNN model depends on the number of feature maps learned by the hidden layers. In the proposed algorithm, the number of hidden layers was decreased and its effect on the performance (both its accuracy and training time) of the network as a feature extractor was studied. As a result, the 13th intermediate layer of the network was used for activation (Figure 11) which computes 256 feature maps, each of size 14×14, and gives the same accuracy as that of the original network. Reducing size of the network helped improve the processing time during training and validation.

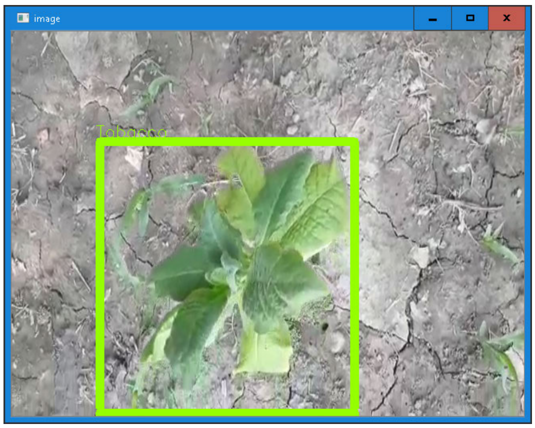

FIGURE 12. Real-time object detection.

For further comparison, the customized deep learning-based object detection was compared with a similar deep learning approach i.e., a Single Shot Detector (SSD). For training, a dataset of 364 tobacco images was used where each image contained at least one instance of a labelled tobacco plant. The dataset was split into test (70%) and validation (30%) sets. As explained in the previous section, the SSD architecture has two sub-networks: (1) feature extraction network, and (2) a classification network. For feature extraction, a pretrained MobileNet-v2 network was fine-tuned. The subsequent detection sub-network composed of a few convolutional layers and a fully-connected layer that modified the MobileNet-v2 architecture into a single-class SSD network for tobacco detection. Like before, the network input size used was $224 \times 224 \times 3$. For data augmentation, the images in the training set was randomly flipped and scaled.

To deploy the deep learning application on the Inte's Movidius Vision Processing Unit (VPU), the Intel's Open Visual Inference and Neural network Optimization (OpenVINO) kit has been used. The Intel Distribution of OpenVINO toolkit 2021.1 for Windows 10 provides CNN-based deep learning inference engines for all the Intel

TABLE 4. Comparison of frame rates (in hertz) for the algorithms presented.

Feature Extraction	Classification	Device	Real-time Speed (in FPS)	Accuracy
Haralick + Hu + color moments	SVM	Embedded PC (Raspberry Pi 4)	6.25	96%
Haralick + Hu + color moments	SVM	Laptop PC (Intel Core i5-7200 2.5GHz CPU, 8GB RAM with Integrated HD Graphics 620)	12.5	96%
CNN (Modified ResNet18)	SVM	Embedded PC (Raspberry Pi 4 + Intel Neural Computing Stick 2)	0.55	100%
CNN (Modified ResNet18)	SVM	Laptop PC (Intel Core i5-7200 2.5GHz CPU, 8GB RAM with Integrated HD Graphics 620)	4.95	100%
CNN (MobileNet v2)	CNN (SSD)	Embedded PC (Raspberry Pi 4 + Intel Neural Computing Stick 2)	5	81%
CNN (MovileNet v2)	CNN (SSD)	Laptop PC (Intel Core i5-7200 2.5GHz CPU, 8GB RAM with Integrated HD Graphics 620)	11	81%

hardware categories i.e., CPU, GPU, and VPU. Other components of the toolkit installed included the Intel Deep Learning Deployment Toolkit, the Model Optimizer (that converts Caffe/TensorFlow/MXNet/ONNX models to the Intermediate Representation format) and Open Model Zoo (a repository of pre-trained models).

B. RESULTS AND DISCUSSION

Accuracy of the proposed approach in this study is shown in Tables 2 and 3. Figure 12 shows the detection result of tobacco plant in real-time.

Table 4 summarizes result of the real-time performance comparison on different target machines.

VI. CONCLUSION

The research work presented in this paper deals with the detection of crop (tobacco) and weeds by utilizing RGB cameras mounted on a boom sprayer in a real field. A classification and detection framework based on SVM was built that could detect crops and weeds, and extract features in real-time. Accuracy and real-time performance analysis was conducted for the traditional SVM and the deep learning algorithm. Moreover, the dataset presented herein was taken on real tobacco fields having multifaceted natural scenarios i.e. high weed densities, overlapped and occluded weeds by tobacco plants, varying illumination levels due to sunny/clouded weather conditions and day timings, different soil humidifies, varying orientations of plant leaves, and soil deposits on leaves of tobacco plants and weeds. Thus, the said dataset helped in providing a truly tough and comprehensive assessment of the machine learning and deep learning-based detection and classification models.

It was observed that the fusion of color moments in L*a*b space, Haralick texture features, and Hu shape moments gives both best accuracy and real-time performance for tobacco and weeds classification. On the basis of these features, first training and then real-time detection was performed. Accuracy of each machine learning algorithm was computed, with the customized deep learning-based algorithm giving 100% while the SVM algorithm giving 96%. However, for real-time detection on an embedded platform (Raspberry Pi 4), the SVM classifier achieved frame rate between 5 to 6 compared to 0.22 FPS for the customized deep learning algorithm. For the traditional deep learning approach (SSD with MobileNet v2), though the FPS achieved was considerably higher, the accuracy achieved was low (81%). This makes the developed proposed solution (SVM with hand-engineered Haralick, Hu, and color feature) suitable to be embedded in an agricultural precision sprayer. Future work involves incorporating and testing fuzzy logic based control of the variable-rate precision sprayer which will have the ability to differentiate between plants and weeds in real-time, perform site-specific/targeted spray on either crops or weeds, and also change flow-rate according to the plant requirements (canopy size, infestation level, etc.).

REFERENCES

[1] P. A. McDaniel, G. Solomon, and R. E. Malone, "The tobacco industry and pesticide regulations: Case studies from tobacco industry archives," *Environ. Health Perspect.*, vol. 113, no. 12, pp. 1659–1665, Dec. 2005.

[2] Q. U. Zaman, T. J. Esau, A. W. Schumann, D. C. Percival, Y. K. Chang, S. M. Read, and A. A. Farooque, "Development of prototype automated variable rate sprayer for real-time spot-application of agrochemicals in wild blueberry fields," *Comput. Electron. Agricult.*, vol. 76, no. 2, pp. 175–182, May 2011.

[3] A. Ruckelshausen, P. Biber, M. Dorna, H. Gremmes, R. Klose, A. Linz, and F. Rahe, "BoniRob—An autonomous field robot platform for individual plant phenotyping," *Precis. Agricult.*, vol. 9, no. 841, p. 1, 2009.

[4] B. Chostner, "See & spray: The next generation of weed control," *Resource Mag.*, vol. 24, no. 4, pp. 4–5, 2017.

[5] R. Berenstein, O. B. Shahar, A. Shapiro, and Y. Edan, "Grape clusters and foliage detection algorithms for autonomous selective vineyard sprayer," *Intell. Service Robot.*, vol. 3, no. 4, pp. 233–243, Oct. 2010.

[6] A. Stentz, C. Dima, C. Wellington, H. Herman, and D. Stager, "A system for semi-autonomous tractor operations," *Auto. Robots*, vol. 13, no. 1, pp. 87–104, 2002.

[7] M.-A. Michaud, C. Watts, and D. Percival, "Precision pesticide delivery based on aerial spectral imaging," in *Proc. CSBE/SCGAB*. Edmonton, AB, Canada: American Society of Agricultural and Biological Engineers, Jul. 2006, p. 1.

[8] C. Fernández-Quintanilla, J. M. Peña, D. Andújar, J. Dorado, A. Ribeiro, and F. López-Granados, "Is the current state of the art of weed monitoring suitable for site-specific weed management in arable crops?" *Weed Res.*, vol. 58, no. 4, pp. 259–272, Aug. 2018.

[9] A. Wang, W. Zhang, and X. Wei, "A review on weed detection using ground-based machine vision and image processing techniques," *Comput. Electron. Agricult.*, vol. 158, pp. 226–240, Mar. 2019.

[10] D. C. Slaughter, D. K. Giles, and D. Downey, "Autonomous robotic weed control systems: A review," *Comput. Electron. Agricult.*, vol. 61, no. 1, pp. 63–78, Apr. 2008.

[11] N. Li, X. Zhang, C. Zhang, L. Ge, Y. He, and X. Wu, "Review of machine-vision-based plant detection technologies for robotic weeding," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2019, pp. 2370–2377.

[12] K. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, p. 2674, Aug. 2018.

[13] J. Behmann, A.-K. Mahlein, T. Rumpf, C. Römer, and L. Plümer, "A review of advanced machine learning methods for the detection of biotic stress in precision crop protection," *Precis. Agricult.*, vol. 16, no. 3, pp. 239–260, Jun. 2015.

[14] H. Y. Jeon and L. F. Tian, "Direct application end effector for a precise weed control robot," *Biosyst. Eng.*, vol. 104, no. 4, pp. 458–464, Dec. 2009.

[15] S. Kaur, S. Pandey, and S. Goel, "Plants disease identification and classification through leaf images: A survey," *Arch. Comput. Methods Eng.*, vol. 26, no. 2, pp. 507–530, Apr. 2019.

[16] M. D. Cicco, C. Potena, G. Grisetti, and A. Pretto, "Automatic model based dataset generation for fast and accurate crop and weeds detection," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 5188–5195.

[17] D. Hall, F. Dayoub, J. Kulk, and C. McCool, "Towards unsupervised weed scouting for agricultural robotics," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 5223–5230.

[18] D. Hall, F. Dayoub, T. Perez, and C. McCool, "A transplantable system for weed classification by agricultural robotics," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 5174–5179.

[19] T. U. Rehman, Q. U. Zaman, Y. K. Chang, A. W. Schumann, and K. W. Corscadden, "Development and field evaluation of a machine vision based in-season weed detection system for wild blueberry," *Comput. Electron. Agricult.*, vol. 162, pp. 1–13, Jul. 2019.

[20] C. Sullca, C. Molina, C. Rodríguez, and T. Fernández, "Diseases detection in blueberry leaves using computer vision and machine learning techniques," *Int. J. Mach. Learn. Comput.*, vol. 9, no. 5, pp. 656–661, Oct. 2019.

[21] X. P. Burgos-Artizzu, A. Ribeiro, A. Tellaeche, G. Pajares, and C. Fernández-Quintanilla, "Improving weed pressure assessment using digital images from an experience-based reasoning approach," *Comput. Electron. Agricult.*, vol. 65, no. 2, pp. 176–185, Mar. 2009.

[22] J. M. Peña, J. Torres-Sánchez, A. I. de Castro, M. Kelly, and F. López-Granados, "Weed mapping in early-season maize fields using object-based analysis of unmanned aerial vehicle (UAV) images," *PLoS ONE*, vol. 8, no. 10, Oct. 2013, Art. no. e77151.

[23] Z. Fan, J. Lu, M. Gong, H. Xie, and E. D. Goodman, "Automatic tobacco plant detection in UAV images via deep neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 876–887, Mar. 2018.

[24] H. Ali, M. I. Lali, M. Z. Nawaz, M. Sharif, and B. A. Saleem, "Symptom based automated detection of citrus diseases using color histogram and textural descriptors," *Comput. Electron. Agricult.*, vol. 138, pp. 92–104, Jun. 2017.

[25] Z. Iqbal, M. A. Khan, M. Sharif, J. H. Shah, M. H. U. Rehman, and K. Javed, "An automated detection and classification of citrus plant diseases using image processing techniques: A review," *Comput. Electron. Agricult.*, vol. 153, pp. 12–32, Oct. 2018.

[26] H. T. Søgaard and I. Lund, "Application accuracy of a machine vision-controlled robotic micro-dosing system," *Biosyst. Eng.*, vol. 96, no. 3, pp. 315–322, Mar. 2007.

[27] A. I. de Castro, M. Jurado-Expósito, J. M. Peña-Barragán, and F. López-Granados, "Airborne multi-spectral imagery for mapping cruciferous weeds in cereal and legume crops," *Precis. Agricult.*, vol. 13, no. 3, pp. 302–321, Jun. 2012.

[28] H. S. Midtiby, S. K. Mathiassen, K. J. Andersson, and R. N. Jørgensen, "Performance evaluation of a crop/weed discriminating microsprayer," *Comput. Electron. Agricult.*, vol. 77, no. 1, pp. 35–40, Jun. 2011.

[29] F. Castaldi, F. Pelosi, S. Pascucci, and R. Casa, "Assessing the potential of images from unmanned aerial vehicles (UAV) to support herbicide patch spraying in maize," *Precis. Agricult.*, vol. 18, no. 1, pp. 76–94, Feb. 2017.

[30] K.-H. Dammer, "Real-time variable-rate herbicide application for weed control in carrots," *Weed Res.*, vol. 56, no. 3, pp. 237–246, Jun. 2016.

[31] T. Rumpf, C. Römer, M. Weis, M. Sökefeld, R. Gerhards, and L. Plümer, "Sequential support vector machine classification for small-grain weed species discrimination with special regard to Cirsium arvense and Galium aparine," *Comput. Electron. Agricult.*, vol. 80, pp. 89–96, Jan. 2012.

[32] C. Gée, J. Bossu, G. Jones, and F. Truchetet, "Crop/weed discrimination in perspective agronomic images," *Comput. Electron. Agricult.*, vol. 60, no. 1, pp. 49–59, Jan. 2008.

[33] J. Hemming and T. Rath, "PA—Precision agriculture: Computer-vision-based weed identification under field conditions using controlled lighting," *J. Agricult. Eng. Res.*, vol. 78, no. 3, pp. 233–243, 2001.

[34] R. D. Lamm, D. C. Slaughter, and D. K. Giles, "Precision weed control system for cotton," *Trans. ASAE*, vol. 45, no. 1, p. 231, 2002.

[35] N. Wang, N. Zhang, J. Wei, Q. Stoll, and D. E. Peterson, "A real-time, embedded, weed-detection system for use in wheat fields," *Biosyst. Eng.*, vol. 98, no. 3, pp. 276–285, Nov. 2007.

[36] A. Tellaeche, X. P. Burgos-Artizzu, G. Pajares, and A. Ribeiro, "A vision-based method for weeds identification through the Bayesian decision theory," *Pattern Recognit.*, vol. 41, no. 2, pp. 521–530, Feb. 2008.

[37] D. C. Slaughter, D. K. Giles, S. A. Fennimore, and R. F. Smith, "Multi-spectral machine vision identification of lettuce and weed seedlings for automated weed control," *Weed Technol.*, vol. 22, no. 2, pp. 378–384, Apr. 2008.

[38] X. Li and Z. Chen, "Weed identification based on shape features and ant colony optimization algorithm," in *Proc. Int. Conf. Comput. Appl. Syst. Model. (ICCASM)*, vol. 1, Oct. 2010, Art. no. V1-384.

[39] X. P. Burgos-Artizzu, A. Ribeiro, M. Guijarro, and G. Pajares, "Real-time image processing for crop/weed discrimination in maize fields," *Comput. Electron. Agricult.*, vol. 75, no. 2, pp. 337–346, Feb. 2011.

[40] F. Zhang and X. Zhang, "Classification and quality evaluation of tobacco leaves based on image processing and fuzzy comprehensive evaluation," *Sensors*, vol. 11, no. 3, pp. 2369–2384, Feb. 2011.

[41] D. S. Guru, P. B. Mallikarjuna, S. Manjunath, and M. M. Shenoi, "Machine vision based classification of tobacco leaves for automatic harvesting," *Intell. Autom. Soft Comput.*, vol. 18, no. 5, pp. 581–590, Jan. 2012.

[42] M. A. Molina-Villa and L. A. Solaque-Guzmán, "Machine vision system for weed detection using image filtering in vegetables crops," *Revista Facultad de Ingeniería Universidad de Antioquia*, vol. 80, pp. 124–130, Sep. 2016.

[43] P. Lottes, M. Hoeferlin, S. Sander, M. Muter, P. Schulze, and L. C. Stachniss, "An effective classification system for separating sugar beets and weeds for precision farming applications," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 5157–5163.

[44] A. Milioto, P. Lottes, and C. Stachniss, "Real-time blob-wise sugar beets vs weeds classification for monitoring fields using convolutional neural networks," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 4, p. 41, Aug. 2017.

[45] I. D. García-Santillán and G. Pajares, "On-line crop/weed discrimination through the mahalanobis distance from images in maize fields," *Biosyst. Eng.*, vol. 166, pp. 28–43, Feb. 2018.

[46] R. Raja, T. T. Nguyen, D. C. Slaughter, and S. A. Fennimore, ''Real-time weed-crop classification and localisation technique for robotic weed control in lettuce,'' *Biosyst. Eng.*, vol. 192, pp. 257–274, Apr. 2020.

[47] F. Lin, D. Zhang, Y. Huang, X. Wang, and X. Chen, ''Detection of corn and weed species by the combination of spectral, shape and textural features,'' *Sustainability*, vol. 9, no. 8, p. 1335, Aug. 2017.

[48] S. Haug, A. Michaels, P. Biber, and J. Ostermann, ''Plant classification system for crop/weed discrimination without segmentation,'' in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Mar. 2014, pp. 1142–1149.

[49] C. S. Marzan and C. R. Ruiz, Jr., ''Automated tobacco grading using image processing techniques and a convolutional neural network,'' *Int. J. Mach. Learn. Comput.*, vol. 9, no. 6, pp. 807–813, Dec. 2019.

[50] R. Kamath, M. Balachandra, and S. Prabhu, ''Paddy crop and weed discrimination: A multiple classifier system approach,'' *Int. J. Agronomy*, vol. 2020, pp. 1–14, Jun. 2020.

[51] A. Bakhshipour and A. Jafari, ''Evaluation of support vector machine and artificial neural networks in weed detection using shape features,'' *Comput. Electron. Agricult.*, vol. 145, pp. 153–160, Feb. 2018.

[52] C. Solomon and T. Breckon, *Fundamentals of Digital Image Processing: A Practical Approach With Examples in MATLAB*. Hoboken, NJ, USA: Wiley, 2011.

[53] P. Maheshwary and N. Srivastava, ''Prototype system for retrieval of remote sensing images based on color moment and gray level co-occurrence matrix,'' *J. Comput. Sci. Issues*, vol. 3, pp. 20–23, Aug. 2009.

[54] K. He, X. Zhang, S. Ren, and J. Sun, ''Deep residual learning for image recognition,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

MUHAMMAD TUFAIL (Member, IEEE) received the B.Sc. degree in computer system engineering and the M.Sc. degree in mechatronics engineering from the Asian Institute of Technology, Thailand, and the Ph.D. degree in mechanical engineering with specialization in manufacturing and mechatronics from the University of British Columbia (UBC), Canada. He then worked as a Postdoctoral Research Fellow with the Industrial Automation Laboratory, UBC, in 2016. Since 2017, he has been working as an Assistant Professor with the Department of Mechatronics Engineering, University of Engineering and Technology (UET), Peshawar. He is also the Co-Director of the Advanced Robotics and Automation Laboratory, UET, and a Co-PI with the National Center of Robotics and Automation, Pakistan. His research interests include robot control, computer vision, and embedded systems.

JAVAID IQBAL (Member, IEEE) received the bachelor's and master's degrees in mechanical engineering from the College of Electrical and Mechanical Engineering (CE & ME), National University of Science and Technology (NUST), Islamabad, and the Ph.D. degree from the University of New South Wales, Australia. He is currently performing his duties as the Dean of the Faculty of Engineering, CE & ME, NUST. His research interests include industrial simulation, machine learning, mechanical systems, and mobile robots among others.

MOHSIN ISLAM TIWANA (Member, IEEE) received the B.E. degree in mechatronics engineering from the National University of Sciences and Technology (NUST), Pakistan, in 2007, and the Ph.D. degree in biomedical engineering from the University of New South Wales (UNSW), Sydney, Australia. He received the Rectors Gold Medal for the B.E. degree. For his pioneering work on artificial limb design, he was given the Presidential Award. He received the Global Innovation Award for the Ph.D. degree, in 2012, and the prestigious Technology Transformer of the Year Award by CRDF Global, USA, in 2013.

MUHAMMAD SHAHAB ALAM received the B.S. and M.S. degrees in mechatronics engineering from Air University, Islamabad, Pakistan, in 2011 and 2015, respectively. From 2015 to 2019, he was a Lecturer with the Department of Mechanical Engineering, Sarhad University of Science and Information Technology, Peshawar, Pakistan. Since 2019, he has been a Research Associate with the Advanced Robotics and Automation Laboratory (ARAL), National Center of Robotics and Automation, Pakistan. His research interests include agricultural robotics and automation and machine vision.

ZUBAIR AHMAD KHAN received the bachelor's degree in mechatronics engineering from the National University of Science and Technology, Peshawar, Pakistan, and the master's and Ph.D. degrees in mechatronics engineering from the University of Engineering and Technology, Peshawar, in 2012 and 2019, respectively. He is currently a Lecturer with the Department of Mechatronics Engineering, University of Engineering and Technology. His research interests include manufacturing automation, robotics, and intelligent control systems.

MUHAMMAD TAHIR KHAN received master's degree in mechatronics from the University of New South Wales, Sydney, Australia, and the Ph.D. degree from the University of British Columbia, Vancouver, BC, Canada, in 1999 and 2010, respectively. He was a Postdoctoral Fellow with the Industrial Automation Laboratory, University of British Columbia, for two years until January 2012. He is currently a Professor with the Department of Mechatronics Engineering, University of Engineering and Technology, Peshawar, Pakistan. His research interests include robotics and intelligent control systems.