Weed Identification Methodology by using Transfer Learning

Bushra Idrees¹, Amanullah Khan¹

¹Department of Computer Science, Air University Multan Campus.

Email: bushraidrees7777@gmail.com

ABSTRACT

From recent past years, Weed identification remained a hot topic for researchers. Majority of work focused on the detection of weed but we are trying to identify the weed via weed name. The unrivaled successes of deep learning make the researchers able to evaluate different weed species in the complex rangeland climate. Nowadays, with an increasing population, farming productivity needs to be increased a lot to meet the demand for accurate weed detection. Increased demand for an increase in the use of herbicides, resulting in environmental harm. In this research work, the picture of weed helps to detect and differentiate as per area, and its name. The main aim of this research is the identification of weed so that fewer herbicides can use. This research work will contribute to reducing the higher use of herbicides by helping clear identification of weed names through its features. We use transfer learning in machine learning. The deep Weeds dataset is used for the evaluation. For this, we use the deep learning model ResNet50 to get better results. The Deep Weeds dataset contains 17,509 images that are label and eight nationally recognized species of weed belonged to 8 across northern Australia locations. This paper declares a baseline for classification performance on the dataset of weed while utilizing the deep learning model ResNet-50 and it is a benchmark too. Deep learning model ResNet-50 attained an average accuracy classification of 96.16. The findings are high enough to make effective use of weed control methods in Pakistan for future field implementation. The results confirm that our System offers more effective Weed recognition than many other systems.

KEYWORDS: Weed Identification, Transfer Learning Deep Learning, ResNet50.

1. INTRODUCTION

In the era of precision agriculture, scientists, researchers and governmental organizations contribute considerably. It is possible if the agricultural land produces high-quality yield in the form of higher quality crops. As we know that, Pakistan has a higher agricultural sector, Pakistan is involved in the cultivation of various crops in those regions. It is the reason that agriculture plays a significant role in the economy of Pakistan. Quality food can be attained by evaluating the quality of crops and fresh methods to select for growth enhancement purposes such as vertical farming and many more. Higher quality crops can be produced by doing core research and introducing different techniques which could enhance crop growth and quality. In this regard,

many studies have been carried out on different types of crops for the improvement of their quality and their development in this respect.

In this Research, Our objective is to identify the weed in the crop using the deep learning ResNet-50 model. The major difference between humans and computers is intelligence that is the ability to make the right decisions at the right time based on past experiences. Computer scientists strive to bridge this gap to make computers intelligent. Computers are programmed to imitate intelligence by means of mathematical calculations upon previously learned experiences. This is known as Artificial Intelligence (AI). An important aspect of human intelligence is context awareness and perception. Humans have senses to take input in the form of image, sound, taste, smell, and touch.

Computers have mechanical sensors for

different types of inputs, for example, a camera for capturing visual input, a microphone for capturing sound, light sensors to capture the ambient light intensity and so on. After the input is received, the next step is perception. The human brain quickly infers the meaning of the input they receive. Humans majorly rely on their vision to discover the world around them and act accordingly. For humans, processing visual data is a simple task, on the other hand, when computers are given an image as input, it is a very difficult and complex task for computers to describe or understand the contents and/or focus parts of an image.

Computer Vision tackles the complexity of the image processing techniques which enables us to extract some meaningful information from the image/video which is provided. The main purpose of the computer vision domain is to make computers too smart to understand the visual data they gather from the world around.

In recent decades, due to the vast and easy availability of camera-fitted smartphones at affordable prices, capturing and sharing images has become a part of people's daily lives. Due to this change, there has been an influx of images available to the world through the internet. Humans have felt the need for computers to make sense of images as humans can by just looking at them. For example, when a human sees an image of a dog running on the beach, he can instantly tell that it is a dog that is running on the beach but for computers, it is not as simple as it sounds because the computer sees an image as an array of numbers. At a higher level Computer Vision tasks may be split into three steps:

- Image Acquisition
- Image Processing
- Image Analysis and Decision

The task of image caption generation using a computer can be decomposed into following steps:

- Image Features Extraction
- Training Model on Extracted Features
- Testing Model with Query Images

In the past, algorithms were written to manually detect edges, corners, points and other shapes in efforts to attempt to extrapolate features that could be used for classification. These algorithms include Harris [1], SURF [2], HOG [3], SIFT [4], BRISK [5] and FREAK [6]. However, these solutions were not scalable. They were affected by factors such as scaling,

translation, rotation and background clutter. For example, if a computer was shown straight images of cat and at test time a rotated image of a cat was presented, the computer might not be able to detect it correctly as a cat. Efforts were made to make the algorithms scale, rotation and translation invariant but the results were not good enough to compete human vision.

After the feature extraction phase, the next step is to analyze patterns in these features, so that a computer can "learn" to distinguish among images of different classes. The domain that deals with pattern recognition and modeling algorithms are known as Machine Learning (ML). In this domain, computers learn patterns from existing example data. With the development of fast computers and the availability of large labeled datasets in the 2000s, the recent trend after 2010 is to use ML algorithms and techniques to solve computer vision problems/tasks. Several machine learning methods exist including Decision Trees, K-Support Vector Machines (SVM) and K-Nearest Neighbors (KNN) but each has its own merits and demerits.

In the computer vision domain, image classification is known as the general work in the vision of a computer where computers try to classify an input image into a collection of a finite number of classes. For example, if a classifier model is trained on two classes that are, cats and dogs, given an input image, it should be able to classify it as a dog or cat based on the learned features. Object detection is the next step in computer vision after image classification where the object's positions are detected in the image. For example, if a classifier classifies an image as a cat then a detector's job is to detect the location of the cat in the image.

An artificial neural network with many cascaded layers of artificial neurons is known as Deep Neural Network (DNN) and Deep Learning is the branch of ML where the Deep Neural Networks are used to make computers mimic the working of human intelligence by learning manifold stages of features that correspond to diverse stages of abstraction. These multi-level features build up a conceptual hierarchy of the input data. Deep learning models learn to accomplish classification directly from input data such as images, sound or text. No hard-coded rules or feature extraction algorithms are required. Traditional neural networks have only 2 or 3 layers of neurons on the other hand, deep networks can have

hundreds of layers.

Starting in 2012, a specific type of network introduced which is a type of deep neural network known as Convolutional Neural Networks (CNNs) began dominating the image classification space – and while deep learning with CNNs has yielded amazing results, we continue to look for more computationally efficient, more accurate, and more descriptive models to apply to this task. As more novel Convolutional Neural Network (CNN) architectures are designed, the current industry standard is for them to be benchmarked through the ImageNet Challenge.

Data mining and Machine learning methods are widely used in different applications of the real world. Training data is an assumption of old methodologies for ML and test data is taken from the same domain, so the function of space input and the characteristics of the distribution of data are identical. In machine learning, real-world conditions, it is the argument that can't sustain. There are multiple cases where the training of data is too much expensive and very hard to obtain. It is then very essential to produce high-performance learners trained with much easier data obtained from multiple domains, it is then referred to as transformation learning [7].

2. PROBLEM STATEMENT

With the advancement of technology and research, it is observed that different kinds of weeds exist in the crops. Some crops didn't even progress due to weeding problems. It remained an alarming issue in the agriculture industry. So, with research, it is identified that there are over 30,000 species of weeds exist and we are going to do our research to identify which type of weed is rooted in the spoiled crop. Our research aim is to identify the type of weed through deep learning. This work will benefit the research community as well as in agriculture to detect and identify the weed to decrease herbicide use by spraying it in the field where weed exists.

3. LITERATURE REVIEW

With the availability of large labeled images data sets like Flickr8k, ImageNet and the development of fast, powerful dedicated processing hardware the interest in neural networks revived in the year 2009 onwards. Up till 2009, all the heavy lifting of training neural networks was being done by CPUs. Researchers were using multiple Processors or multi-core

processors to train neural nets but CPU are general-purpose hardware built to perform generic tasks of a computer. Nvidia, a company is known for manufacturing state-of-the-art GPUs for gaming started working on CUDA cores which are nowadays used in deep neural network training. It was determined by Andrew Ng determined that the speed of deep-learning systems could be increased by 100 times using Graphical Processing Units. This brings along several bundled benefits like ease of scalability and modularity. One can easily add GPU to an existing system, also it is fairly simple to add multiple GPUs and splitting the workload as done in GPUs consist of hundreds of processing cores that specially designed and optimized calculation in machine learning [8]. The article deals with the creation of a real-time algorithm for a specific weed Recognition system based on the Histogram Analysis of an image used to identify weeds [9]. Results analysis shows over 95 percent classification accuracy over 140 (broad and narrow) sample images with 70 samples from each weed group. In this article, author worked on the weed classification for identifying weed species for control [10]. Weeds are classified into two types based on the frequency of edges present in them. In this article, they take the case of corn crop where weed edge frequency is more than crop frequency. They introduce an algorithm that loads the image from the source and use three different methods to detect the type of the weed from the crop such as color segmentation, edge detection, filtering. The purpose to create such an algorithm is to classify and reuse the affected area of weed for more seeding. This particular area may be considered for further weed control operations, leading to more development [11]. Here the author represents the work that is based on a machine vision device for detecting weeds in vegetable Crops using outdoor pictures, avoiding problems with lighting and sharpness during the acquisition process. This application will be a weed removal mobile robot with a camera obscura (Latin for "dark room") module for monitoring lighting conditions. The aim of this paper is to develop a useful algorithm to classify weeds, using image filtering to extract color and area characteristics, then a method is implemented to mark and object in the scene, finally, an area-based classification is proposed, including sensitivity, specificity, positive and negative predicted values to assess algorithm efficiency [12]. The author worked on an image

dataset of 6 different species of fish, captured by a single camera from various pools near the Head Qadirabad, Chenab River in Punjab, Pakistan. The Fish-Pak dataset is very useful to compare different classifier variables like learning rate, momentum and their effect on overall results. Convolutional Neural Network (CNN) is one of the most commonly used image recognition architectures, based on visual characteristics. Six classes of data, i.e. The dataset included *Ctenopharyngodon idella* (Grass carp), *Cyprinus carpio* (Common carp), *Cirrhinus mrigala* (Mori), *Labeo rohita* (Rohu), *Hypophthalmichthys molitrix* (Silver carp), and *Catla* (Thala), with various photos. The Zoology Lab hosts Fish-Pak, under the joint affiliation of the Department of Computer Science and the Department of Zoology, Gujrat University, Gujrat, Pakistan. This paper Unmanned Air Vehicle (UAV) is used in various phases for data acquisition of wheat crops to capture the high quality of RGB images [13]. The proposed method facilitates weed, wheat, and barren land extraction in the field of wheat crops using background subtraction. The result shows that the method of context subtraction is good for grass, barren land and wheat detection. The author identified the weeds using image processing from the photographs of the site and classified them according to the classification of area and aspect ratio [14]. Their main purpose is to identify and eliminate the weed using image processing. They use two methods for weed detection such as Inter-Row weed detection and Inter-Plant weed detection. Author use three stages of paddy fields to detect weed. These three stages include when paddy is small when the crop size is greater than the weed size, when the paddy turns brown. The reason for creating such a system is to classify and reuse affected area of weed for further seeding. This particular area can be considered, resulting in further development, for more weed control operations. In this paper [15] the author worked on Device implementation of weed detection and removal technique for image processing. This includes easy detection of the edges related filters techniques such as Gaussian and Laplacian. Finally, it concludes with the results of the feature extraction which implement the ORB algorithm. An RGB image shall be taken as a sample to display the difference between weed and seed. This RGB image is further processed to detect weeds. They get an output after several steps where the weeds are isolated from the crop

taken in the sample image. In another article [16] this is based on the Image processing technique that detects and extracts weeds between the crops. The weeds are divided into three groups, based on shape and edge frequencies. The first is Less Edge Frequency Weeds, the second is medium Edge frequency Weeds and the third is High Edge Frequency Weeds. The weeds can be found in every area of the field. In this paper, the author considers the weeds between the crops and weeds between the rows. The author uses an image segmentation process that involves many pre-processing procedures. The clustering method is used there to partition the image into clusters. Information on weed prevalence in that specific area will be given to the Herbicide Sprayer unit. Although the usage of herbicides can be reduced by not spraying it to a large extent into the entire sector. However, the only downside to this method is that if there is a rise in the type of weed that is more than two weeks then the procedure would not apply. This report [17] is on the deep weeds. The author collects the dataset consists of 17,509 photos depicting eight distinct weed species of neighboring flora native to in situ Australia. Their intention is to use the collected data set to train weed control on the ground Robot; thus, to fit this requirement, they must customize the dataset. The prototype ground-based weed control robot combines high-resolution cameras and fast-acting solenoid sprayers to spray specified weed targets selectively. They use the most successful CNN models that are ResNet50 and InceptionV3 models. They have high dataset baseline performance, achieving an accuracy score of 95.1% and 95.7% respectively. ResNet-50 has good performance inference within the target robotic weed control applications, inferring 53.4 ms per image and 18.7 fps. This article is about vegetable identification [18]. Here the author used the method which is the combination of deep learning and image processing technology. Due to the deep learning, CenterNet trained model detect the vegetables and draw boundaries around the vegetables. This model got the precision of 95.6%. This paper is about weed identification in the fields of maize, potatoes and sunflowers [19]. As their top priority is to control of herbicides because it's not good for health. They used 93000 images gathered by RGB camera. Three different CNN models VGG16, ResNet-50 and Xception used to trained the images. The top-1 accuracy is

about 77% and 98%.

4. METHODOLOGY

The aim of this research is two-fold. The first is to identify the weed and the second is to bring the stability of prediction models. For this, we proposed the methodology which can be shown in Figure 4. 1. The famous CNN ResNet-50 model was picked for execution dependent on its solid presentation on profoundly factor datasets and their accessibility in TensorFlow and Keras backend. ResNet-50 was ILSVRC 2015 victor [20]. Its annual competition depicts that member models perceive 1,000 distinctive ImageNet object classes from more than five hundred thousand pictures.

In Large scale visual Recognition challenge champ, GBD-Net was not accepted in light of the fact that it needs within per-pixel marks, which are still not accessible for the Deep Weeds data set [21]. Moreover, rivalry champ of 2017, SE networks was not observed in light of a fact that the SE augmented models are not basically accessible in the TensorFlow and Keras backend [22]. The model then observed from now on, ResNet-50 is accessible in Keras along with weights that are pre-prepared in the TensorFlow backend [23].

ResNet model was proposed to solve the issue of diminishing gradient. The idea is to skip the connection and pass the residual to the next layer so that the model can continue to train. With ResNet models, CNN models can go

deeper and deeper. There are many variations for ResNet models and we chose Resnet50 here because ResNet-50 is a deep learning model which is pretrained for the classification of image of the ConvNet or CNN that is a neural network deep class. It is usually applied to the analysis of visual imagery. ResNet-50 is deep to 50 layers and then it is trained on the million images of ImageNet database 1000 categories. It is understood that there are multiple pretrained deep models that can be used such as VGG19, GoogleNet, AlexNet etc., but the ResNet-50 model is best and familiar for its less error rate and excellent performance generalization of the task recognitions. It is a useful tool. Furthermore, if you know Gradient Descent then you would encounter the Vanishing Gradient Issue. This issue can be tackled with the use of ResNet Model.

It can be observed from Table 1, that highly minimum execution of grouping ($80.11\% \pm 3$) is achieved with AlexNet model and accurately acquired with the ResNet-50 model with $95.23\% \pm 0.6$ by and large precision.

As far as the cost time is under consideration than Vgg16 models and ResNet contains the higher train time. The fundamental purpose behind this is ResNet-50 model is the biggest profound structure of learning along with other pre-prepared models (Vgg-16, AlexNet, ResNet-34 and ResNet-18). As ResNet-50 model allows the data move via network with residual connections, which is a

Sr.	Models	Average Training Time	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Overall (%)
1	AlexNet	133 s	80.37	76.64	84.11	81.78	77.57	80.09
2	Vgg-16	249 s	86.92	82.71	89.25	86.45	85.58	86.26
3	ResNet 18	131 s	87.38	85.51	90.65	89.25	84.58	87.47
4	ResNet 34	136 s	92.99	86.92	91.59	92.52	88.32	90.46
5	ResNet 50	248 s	94.86	95.33	96.26	94.39	95.33	95.23

gradient value does not fade via back propagation. The deepest structures of the model have the best order execution. During examination, we utilized the Vgg-16, AlexNet, Resnet-34, Resnet-50, and ResNet-18 deep learning pre-prepared models to group MR pictures in typical, cerebrovascular, neoplastic

provocative and degenerative classifications. All mentioned models are tried utilizing Harvard Medical SchoolMR mind pictures information base utilizing all accessible 1074 pictures. The outcomes depicts that ResNet-50 model acquired the most elevated exactness of $95.23\% \pm 0.6$ comparing with the five models. The

AlexNet has achieved the least presentation in arranging pictures into five classes. Along these lines we are utilizing ResNet-50 model.

The model ResNet-50, is accessible in Keras along with pre-prepared loads in a TensorFlow backend. So, the model is designed to calculate the 1000 diverse ImageNet Object classes. Unique ImageNet-prepared model was somewhat modified to characterize the 9 DeepWeeds species classes. It's calculated by supplanting their entirely connected 1,000 neuron layer with an entirely connected neuron layer that are nine.

Let,

b =no. of images per training batch

r =no. of pixel rows

c =columns

3 =no. of image color channels

then Network accepted

Input matrix= $b \times r \times c \times 3$

After eliminating the entirely connected 1,000 neurons ImageNet output layer, then the output from the network had the $b \times r \times c \times \alpha$ shape, here α gives the specific network, and α was the number of the features which are extracted for each $r \times c$ spatial location.

The spatial average pooling layer was used to change over the entirely convolutional $b \times r \times c \times \alpha$ yield to the $b \times r \times c$ shape that was deeply connected with the final $b \times 9$ weed classification layer. With 32 images/ batch and $224 \times 224 \times 3$ sized input images, the given global average pooling was perfectly identical to the 7×7 average pooling in model (ResNet-50). As the idea of DeepWeeds dataset and grouping task takes into consideration numerous species of weed to be available in each given picture. Figure1. Thusly, a reLU activation function was utilized for each weed-explicit neuron in the output layer. It permitted output of probabilities for every class to recognize a probability that a picture had a place with each class. A picture was

delegated as one of the weeds objectives if that the class' reLU-activated neuron likelihood was the best among all others and its probability was more prominent than $1/9 = 11.1\%$ (for example arbitrary speculation). The arbitrary estimate edge was executed to defeat the monstrous variety in the negative DeepWeeds class, which makes its objective likelihood be less firmly weighted towards explicit picture highlights than the 8 positive classes—whose pictures are higher predictable.

All labeled images which are 17,509 from the DeepWeeds were split into 60 percent training, 20 percent validation and 20 percent testing with $k=5$. Here, then Stratified random partitioning is performed to make sure classes even distributed within each subset, excluding negative class which is much larger than the even classes. Before the process of training, model was trained to the corresponding weights on ImageNet. The 9-neuron fully connected layer weight or also known as the output layers were initialized through the uniform random distribution. For training, the standard binary cross-entropy loss function was utilized.

To beat the exceptionally factor nature of the objective weed classification application, a progression of augmentations was performed on both the training and validation picture subsets to represent varieties in revolution, scale, shading, enlightenment and point of view. Picture augmentation was performed utilizing the OpenCV Library and its Python wrapper. All pictures were first resized to 256×256 pixels in size and arbitrarily enlarged for every epoch of training. Each picture was likewise arbitrarily turned in the scope of $[360, +360]$ degrees. At that point, each picture was arbitrarily scaled both vertically and on a level plane in the scope of $[0.5, 1]$. Each shading channel was arbitrarily moved inside the scope of ± 25 (for example roughly $\pm 10\%$ of the greatest accessible 8-bit shading encoding range $[0, 255]$). To represent brightening fluctuation, pixel force was arbitrarily moved inside the $[-25, +25]$ territory, moving all shading channels consistently. Likewise, pixel power was arbitrarily scaled inside the $[0.75, 1.25]$ territory. Irregular viewpoint transformations were applied to each picture to reproduce a huge variety of review distances and angles. At long last, the pictures were flipped evenly with a half likelihood and afterward trimmed to hold the 224×224 pixels needed for every input layer architecture. Without this broad augmentation, this CNN

Figure 1: Methodology overview

network definitely over-fitted the accessible pictures by retaining the training subsets.

4.1. Proposed Network Design

The fully working network design of research is given below. In this research, resNet50 model is used to determine the accuracy of weed identification.

4.2. Implementation

Now present a brief summary of the benchmark dataset and tools utilized for this research.

4.2.1. Datasets

In this research deep weeds dataset is used. It consists of 17,509 images across 9

categories. The DeepWeeds is a massive multi-class dataset consisting of 17,509 photos of 8 distinct weed species and Australian native plant life of various target (or negative) species. In order to train a CNN on such big data, it requires a few high-end GPUs hooked up to high-end systems and then the training should complete in about a week [8].

4.2.2. Tools Utilized

In light of its extensive application for both research and development on deep learning, Google Colab (Python & Tensorflow) is utilized all through this undertaking and every one of the outcomes is in this manner got from it.

Figure 2: Full working diagram of the proposed methodology

Due to lack of resources, we used a dataset with a fewer number of images and executed the experiments for a limited number of epochs. A machine with a core-i5 processor and 3 GPU's OF RTX 2080 Ti used for all experiments in this research.

5. EXPERIMENTAL AND RESULTS

This section presents our experimentation outcomes by applying deep learning model on Dataset of Deep Weeds. We executed the original ResNet50 model and utilized it as our base reference model to extract features across nine classes. We use transfer learning concept.

We use ResNet50 model to identify weeds. The model is trained with a minibatch size of 32 on three Nvidia 11GB GPU. The small

Figure 3: DeepWeeds dataset test pictures, in specifically:

(a) Lantana (b) Chinese Apple, (c) Parthenium (d) Parkinsonia, (e) Rubber vine, (f) Prickly acacia, (g) Siam weed, (h) Negatives and (i). Snake weed.

batch size is selected due to GPU memory constraints. We have total 17,509 images on which we have 60% training images, 20% testing images and 20% images are used for validation purposes. We follow the simple data augmentation with rotation, scale, color, illumination and brightness effect to produce augmented data

5.1. ResNet 50 Model Results

For training, Model takes almost 35 epoch's for better results and give us good accuracy. The summarized results are given below:

5.1.1. ResNet 50 Model initial results

The Table 2. shows the initial result of our experiments.

5.1.3. ResNet 50 model final epoch results

With the ResNet-50 CNN model, the DeepWeeds dataset was arranged to set up a benchmark standard of execution for future examination. Outcomes are Underneath,

arranged. We observe the precision (%), weighted top-accuracy (%), and also monitor the false positive rate (%) along with the 5 class folds.

Table 5 shows that overall top-1 classification accuracy of 96.16 percent was achieved by ResNet-50. This is an extraordinarily solid result of a pattern. The ResNet-50 model's slightly stronger presentation can be due to its too higher complexity, having weights of 23.5M trainable.

The accuracy of classification obviously varies via species to species. With 98.6 percent average top-1 accuracy, the most grounded presentation for the ResNet-50 model got in the negative class. This is a strong and impressive product of thinking about the immense diversity in this class. 98.2 percent at Parkinsonia, Siam weed at 97.5 percent, Lantana at 97.0 percent and Thorny acacia at 96.5 percent are weed species with solid execution. The plant species

Table 2. Calculated initial results of resNet50 after the experiment shows the results of three Epoch.

Sr.	Weeds	Precision	Recall	F1-Score	Support
1	Chinese Apple	0.00	0.00	0.00	225
2	Lantana	0.00	0.00	0.00	212
3	Parkinsonia	0.00	0.00	0.00	206
4	Parthenium	0.00	0.00	0.00	204
5	Prickly Acacia	0.00	0.00	0.00	212
6	Rubber Vine	0.00	0.00	0.00	202
7	Saim Weed	0.00	0.00	0.00	215
8	Snake weed	0.00	0.00	0.00	203
9	Negatives	0.52	1.00	0.68	1821
1	Chinese Apple	0.02	1.00	0.03	225
2	Lantana	0.00	0.00	0.00	212
3	Parkinsonia	0.00	0.00	0.00	209
4	Parthenium	0.02	1.00	0.03	203
5	Prickly Acacia	0.00	0.00	0.00	252
6	Rubber Vine	0.00	0.00	0.00	262
7	Saim Weed	0.00	0.00	0.00	225
8	Snake weed	0.00	0.00	0.00	273
9	Negatives	2.92	1.00	1.48	1893

1	Chinese Apple	0.02	1.00	0.03	235
2	Lantana	0.00	0.00	0.00	222
3	Parkinsonia	0.00	0.00	0.00	256
4	Parthenium	0.05	1.00	0.09	244
5	Prickly Acacia	0.00	0.00	0.00	242
6	Rubber Vine	0.03	1.00	0.05	252
7	Saim Weed	0.00	0.00	0.00	265
8	Snake weed	0.00	0.00	0.00	208
9	Negatives	3.01	1.00	1.50	1855

5.1.2. ResNet 50 Model 10 epoch results

Table 3. Calculated 10 epoch results of resNet50 Model

Sr.	Weeds	Precision	Recall	F1-Score	Support
1	Chinese Apple	0.05	1.00	0.09	225
2	Lantana	0.03	1.00	0.05	212
3	Parkinsonia	0.02	1.00	0.03	206
4	Parthenium	0.09	1.00	1.04	204
5	Prickly Acacia	0.90	1.00	0.94	212
6	Rubber Vine	1.90	1.00	1.31	202
7	Saim Weed	2.80	1.00	1.47	215
8	Snake weed	1.10	1.00	1.04	203
9	Negatives	6.42	1.00	1.73	1821

with the lowest arrangement accuracy are 95.8 percent Parthenium, 95.6 percent Rubber Vine, 95.5 percent Chinese apple and 90.8 percent Snake Weed. We estimate that the moderately low exhibition on all of the weed species may because of the weed containing less novel prominent features to train on.

Looking at Table 5 precision of the ResNet-50 model, it can very well to notice that confidence in the predictions which are positive of the model is very high (more noteworthy than 90 percent for all species). 98.2 percent at Parkinsonia, 97.5 percent at Siam Weed percent and Lantana at 97.0 percent are the three species with the most substantiated predictive positive Value. With a precision score of 98.6 percent, the Negative prediction class of high confidence. Snake weed, Chinee apple, Rubber vine, Parthenium and Prickly Acacia show the relative

lowest confidence in its positive prediction, from 90.8% to 96.5%. Because of similar picture features, this can be due to the high rate of confusion between these particular types of weeds. The model's average false positive rate is 0.53 percent, with the individual positive false rates for every species of weed being slightly low to less than 1 percent (as appeared in Table 5).

Results indicate that when deployed in the field, our model is considerable to result in minimum target rate of damage. For ResNet-50, negative class contains a moderately false high positive rate of 2.49 percent. This informs us that over 2% of targeted weeds are falsely classified they belong to class that is negative, since class includes non-target plant life. We assume that it is attributable to the wide variation in life of plant in the negative class all over 8 dataset's

rural locations. Then, it's useful to train a certain location model and then minimize the number of those targets that are misclassified when deploying such a model in the field.

For weed Identification, In this report they used ResNet-50 and Inspection-V3 models and their results are 95.7% and 95.1% respectively [18]. Table 4 also shows the results of report [17]. While in our research we are using

ResNet-50 model and our research results are 96.16%. Table 5 shows results. This paper has strong and surprising results due to use of 3 GPU's OF RTX 2080 Ti and by increasing the number of epoch's. This is illustrated in Figure 5. While the report used 2 GPU's and they did a smaller number of epoch's Figure 4 illustrates their results [17].

Table 4. The average accuracy test classification, false positive rate, precision, recall rate among all 5 validated folds for .y model [17]

Species	Top-1 accuracy (%)	Precision (%)	False positive rate (%)
	ResNet-50	ResNet-50	ResNet-50
Chinee Apple	88.5	91.0	0.61
Lantana	95.0	91.7	0.55
Parkinsonia	97.2	97.9	0.13
Parthenium	95.8	96.7	0.21
Prickly Acacia	95.5	93.0	0.46
Rubber Vine	92.5	99.1	0.05
Siam Weed	96.5	97.2	0.18
Snake Weed	88.8	90.9	0.55
Negatives	97.6	96.7	3.59
Weighted average	95.7	95.7	2.04

Table 5. Our research achieved best results for each species

Species	Top-1 accuracy (%)	Precision (%)	False positive rate (%)
	ResNet-50	ResNet-50	ResNet-50
Chinee Apple	95.5	95.5	0.55
Lantana	97.0	97.0	0.50
Parkinsonia	98.2	98.2	0.10
Parthenium	95.8	95.8	0.18
Prickly Acacia	96.5	96.5	0.36
Rubber Vine	95.6	95.6	0.08
Siam Weed	97.5	97.5	0.08
Snake Weed	90.8	90.8	0.50
Negatives	98.6	98.6	2.49
Weighted average	96.16	96.16	0.53

For weed detection, some papers did work on Histogram analysis [1], edge frequency [2], background subtraction method [5], some work on inter row and inter plant weed detection [6], while other use Laplacian filter with ORB algorithm [7] to detect the weed. But all of the above methods they used either work on real-time or only work on less number of images. While we are using CNN model to detect weed from the images and work amazingly on large number of datasets. This gives better results as compared to the other methods Moreover when the weed and crop color are the same

Figure 4: Resnet-50 model with top-1 accuracy 95.

Figure 5: resnet-50 model with top-1 accuracy 96.16%

convolution network (CNN) can be produced much better results.

6. CONCLUSION

The paper can be concluded with the facts that Image classification and detection are the core tasks in computer vision systems which lead us to all other important tasks. This research introduces an important factor through which we can generate more stable versions of our prediction models. In tests completed at benchmark subsets, the suggested approach beats various methodologies. The presented technique reports stable forecast model generation. Although we have a limited resource of GPU and RAM memory, our proposed method performs much better with respect to the hardware.

Limitations

One limitation that we noticed in our proposed method is that it takes times (directly proportional to the size of augmented data) more training time as compared to the original dataset, because dataset is increased due to data augmentation.

Future Work

The future directions are to reduce the model over fitting and to make it optimized for mobile devices and bring benefits of this technique to develop intelligent mobile apps as well as try this method on other data. Further research aimed to use Nvidia Jetson Nano for real time inference [25].

REFERENCES

- [1] R. Szeliski, Computer vision: algorithms and applications., England: Springer Science & Business Media, 2010.
- [2] T. Kanade, Three-dimensional machine vision, England: Springer Science & Business Media, 2012.
- [3] N. I. C. A. G. a. T. S. Sebe, Huang. Machine learning in computer vision, Springer Science & Business Media, , 2005.
- [4] C. G. a. M. S. Harris, "A combined corner and edge detector," In Alvey vision conference, vol. 15, no. 50, pp. 10-5244, 1988.

- [5] H. T. T. a. L. V. G. Bay, "Surf: Speeded up robust features," In European conference on computer vision Springer, vol. 1, no. 1, pp. 404-417, 2006.
- [6] N. a. B. T. Dalal, "Histograms of oriented gradients for human detection," IEEE computer society conference on computer vision and pattern recognition, vol. 1, no. 5, pp. 886-893, 2005.
- [7] Weiss., Karl., T. M., Khoshgoftaar. and D. Wang., " "A survey of transfer learning."," Journal of Big data , vol. 3, no. 1, pp. 1-10, 2016.
- [8] S. M. C. a. R. Y. S. Leutenegger, "BRISK: Binary robust invariant scalable keypoints," In 2011 International conference on computer vision, vol. 1, no. 1, pp. 2548-2555, 2011.
- [9] I. Ahmad, A. Naeem, Islam. and Muhammad., ""Real-time specific weed recognition system using histogram analysis"," In proceedings of world academy of science, engineering and technology.16. (2006), vol. 16, no. 1, pp. 1-10, 2006.
- [10] A. Paikerkari., V. Ghule., R. Meshram. and V. Raskar., ""Weed detection using image processing."," International Research Journal of Engineering and Technology (IRJET) , vol. 3, no. 3, pp. 1220-1222, 2016.
- [11] Molina-Villa., M. Alejandro. and L. E. Solaque-Guzmán., " "Machine vision system for weed detection using image filtering in vegetables crops."," Revista Facultad de Ingeniería Universidad de Antioquia , vol. 80, no. 1, pp. 124-130, 2016.
- [12] S. Z. H. Shah, H. T. Rauf., M. IkramUllah., M. S. Khalid., M. Farooq., M. Fatima. and S. A. C. Bukhari., ""Fish-Pak: Fish species dataset from Pakistan for visual features based classification."," Data in brief, vol. 27, no. 1, pp. 1-10, 2019.
- [13] S. Hameed and I. Amin., ""Detection of weed and wheat using image processing."," In 2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS), vol. 1, no. 1, pp. 1-5, 2018.
- [14] Avinash., M. A. E., V. Mr. and N. M. Gaunker., " "Weed Detection using Image Processing."," International Research Journal of Engineering and Technology (IRJET), vol. 3, no. 3, pp. 1220-1222., 2016.
- [15] R. Reddy, Anirudh. and G. Laasya., " "Design of Weed Detection System for Cotton field."," MAT journals, Journal of Image processing and Artificial intelligence, vol. 3, no. 3, pp. 1-10, 2017.
- [16] P. Sakthi and P. Yuvarani., ""Detection and Removal of Weed between Crops in Agricultural Field using Image Processing."," International Journal of Pure and Applied Mathematics , vol. 118, no. 8, pp. 201-206, 2018.
- [17] A. Olsen, D. A. Konovalov., B. Philippa., P. Ridd., J. C. Wood., J. Johns. and W. B. e. al., " "DeepWeeds: A multiclass weed species image dataset for deep learning."," Scientific reports , vol. 9, no. 1, pp. 1-12, 2019.
- [18] X. Jin, J. Che and Y. Chen, "" Weed Identification Using Deep Learning and Image Processing in Vegetable Plantation"," IEEE Access, vol. 1, no. 9, pp. 10940-10950, 2021.
- [19] G. Peteinatos, P. Reichel, J. Karouta, D. Andújar and R. Gerhards, ""Weed Identification in Maize, Sunflower, and Potatoes with the Aid of Convolutional Neural Networks"," Remote Sensing, vol. 24, no. 12, p. 4185, 2020.
- [20] Russakovsky., Olga., J. Deng., H. Su., J. Krause., S. Satheesh., S. Ma. and Z. H. e. al., ""Imagenet large scale visual recognition challenge."," International journal of computer vision, vol. 3, no. 115, pp. 1-12, 2019.
- [21] X. Zeng, W. Ouyang., J. Yan., H. Li., T. Xiao., K. Wang. and Y. L. e. al.,