Convolutional Neural Networks and Transfer Learning in Grass Feature Recognition

Ayan Patel, Kiley Roberson, Nick Papadakis, and Caleb Watts

Abstract— **The ability to identify unwanted weeds and grass in an agricultural field is useful and profitable for farmers. It allows farmers to localize pesticides and remove only the grass that is detrimental, preserving the healthy plants. This paper proposes an approach to classifying types of grass using image classification by a transfer learning neural network. Existing solutions for plant classification have been explored successfully, but many are specialized towards recognizing leaves and weeds instead of grass. Recent developments in image classification based on deep neural networks give a promising direction for recognizing grass types.**

Our goal was to classify images of uprooted grass based on their taxonomy. Our approach utilized transfer learning based on the ResNet50 implementation from ILSVRC 2015, with modified outputs in order to classify grasses based on our chosen classes. We obtained a data set of 200 distinct images of grass with corresponding labels. Image augmentation was performed to add diversity to the images. Due to the limited number of images, we decided to approach a simpler problem than taxonomy as a proof of concept. We categorize images by the plant characteristics, giving us two main classes, flowering and grassy plants. The resulting model was overfitted, achieving an accuracy of 0.89 in training, but only 0.50 in validation.

I. Introduction

The presence of unwanted or invasive weeds and grasses in agriculture, grasslands, and pastures can have a significant negative impact on the quality of land [3]. In our natural grasslands, invasive species of grass can crowd out native plant species, causing unnatural fire cycles and depleting water sources [11]. Several studies have shown that in high drought areas, grasslands often serve as a better carbon sink than neighboring forests, as they recover quicker from drought and fire damage [4]. However, dead or invasive grasses can put these carbon sink grasslands in danger, damaging them and making them more susceptible to fires [11].

In agriculture the presence of unwanted grasses and weeds can lead to depleted and damaged crops, as well as an over spraying of expensive and environmentally damaging weed poisons [3]. Current methods of controlling unwanted grass and weeds from farm land involves blanket spraying, a method of weed control that causes contamination to nearby waterways and negatively effects wildlife and other beneficial plant life [8].

Manually locating areas of grasslands with invasive grass species, or doing selective spraying simply on crops affected by weeds, is time consuming and laborious. Manual grass classification and determining if a grass is an invasive species requires a skill that only a limited number of people have, making it difficult to scale. For these reasons, the employment of automatic and computer driven classification systems would promote the health of grasslands, agriculture, and our environment.

The data that we received of local grasses from Cal Poly's Swanton Pacific Ranch was not large enough for us to be able to preform classification of the grasses by their species, so we instead focused on classifying the images by features of the grasses. Specifically, we focused on determining if the plant is a flowering plant or a grassy non-flowering plant. We decided to use a neural network approach based on transfer learning. Image classifiers have recently achieved a high level of accuracy for a variety of applications, and transfer learning allows for trained solutions to be used for new applications. We used a network trained for the ILSVRC competition, ResNet, in order to classify images of grass strains. Our initial data set features grasses that have been extracted from the ground and put against a white background.

II. Summary of Past Approaches

[5] investigated using artificial networks with a scheme for texture extraction in order to identify strains of broadleaf and grass weeds. The solution used Gabor wavelets in order to extract texture information. The approach also featured a convolutional step in addition to the Gabor wavelets. The ANN featured in the network was a three-layer feedforward backpropagation network, with four input nodes and two output nodes which corresponded to broadleaf and grass classes. This approach predates more modern image classification methods, such as AlexNet and ResNet. It is also a "shallower" network than our approach will be, with their approach only having four layers. Our initial approach will instead utilize transfer learning with ResNet, a deep network.

In 2006, [6] sought to develop a method to map invasive nonindigineous plants in North American rangelands. They cite that traditionally, invasive plant mapping is done on the ground by hand, which is very costly. They sought to create a more efficient approach by collecting hyperspectral imagery data using aerial photography. This data was collected in two different sites in Madison County, Montana. They then used this data to train a random forest classifier. The researchers focused solely on two species of plant: leafy surge (*Euphorbia esula L.*) and spotted knapweed (*Centaurea maculosa Lam.*). They were able to achieve a 84% average accuracy rate for the classification of spotted knapweed and 86% accuracy rate for the classification of leafy spurge. This study is related to ours as it also sought to enable more efficient identification of invasive plant species using a machine learning approach. The data set and classifier implementation is very different from ours however. Their data was hyperspectral imaging captured from a plane, while our is pictures of uprooted grass species taken with a traditional camera. For classification, the researchers used a random forest classifier, while we utilize a deep neural network.

In [2], a deep convolutional neural network was created to classify plant species. The system was trained on 22 plant species and had an accuracy of 86.2 percent. A total of 10,413 images were used to train and test on. The main purpose of this system was to identify weeds and keep them under control through the use of herbicides. Some plants were grouped into families instead of species, as similar species were hard to distinguish but easily identifiable into a family. Pre-processing was done to make all the background colors white so that they would not interfere with the classification. Their neural network consisted of convolutional layers, batch normalization, max pooling, and relu activation function, with these layers repeated several times. The last few layers consisted of dropout and fully connected layers with the last one using softmax activation. In total, the network had around 1.2 M learning parameters. The output of the network was a vector with a space for each of the 22 species. This approach is similar to what we are trying to accomplish. Our network will have slightly different layers and we will have less species labels to classify images. In addition, because we have less data than this system to train on, we may end up taking a pre-trained network and building on top of that to get our ideal network for this project.

A similar weed classification study was done by Smith et al. in [8]. Convolution neural networks and transfer learning were employed to classify weeds found in grassland images. In the study the performance of both a conventionally trained CNN with a ResNet architecture, and a transfer learning model using the convolutional layers of an ImageNet pre-trained network were compared on increasingly smaller datasets. The models developed in this study were a 26 layer ResNet network employing dropout, and a transfer model using the convolutional layers of a MobileNet network fed into two dense layers trained on the subject data sets. Performance analysis found that conventionally trained CNN's achieve a higher accuracy given a large enough data set, however, as the data set decreases to as little as 25 samples per class, the conventionally trained ResNet does no better than random guessing, while the transfer learning model still maintains accuracy as high as 83.5%. Due to the limited amount of data collected for our particular problem of classifying grass types, we have elected to use a transfer learning approach. However, rather than using a MobileNet network we will employ a pre-trained ResNet model.

In [9], researchers utilized a fine-tuning transfer learning approach using a pre-trained Alexnet implementation for detection of thoraco-abdominal lymph nodes and classification of interstitial lung disease. They used the fine-tuning approach of training all layers except the last at a rate 10 time smaller than the default rate. They set the last fully-connected layer to be initialized randomly and be trained from scratch. The rationale behind this approach is that it allows the model to accommodate the new object categories for the new application that the neural network is being used for. Similar to our problem, the size of the data sets that the researchers had to work with were much smaller than the ImageNet data set than their starting network was trained on. Whereas ImageNet has millions of images, the data sets used for each of these problems were on the order of hundreds. Despite this small data set, the researchers were able to achieve state-of-the-art performance for mediastinal LN detection using this transfer learning approach.

III. Implementation

A. Data Set

The data set was given to us by Cal Poly's Swanton Pacific Ranch. We were given approximately 300 images but had to remove duplicates and images that did not fit in our two classes. Due to this, we had a total of 87 images for the flowering class and 116 images for the grassy class. We randomly removed 20% of the images from each class to use for validation. Each image was taken after the grass had been removed from the ground and was photographed against a white background as shown in figure 1.

(a) Flowering (b) Grassy

Fig. 1: Images of grass

B. Image Augmentation

Our original data set consisted of images with a resolution of 3024x4032. After categorizing these images into 2 groups, we down-sampled the images by a factor of 10 to get images with a resolution of 300x400. This is because such large images would take a very long time to train a model on, and even with this level of downsampling features were still preserved so that it was easy for the human eye to discern between the two different plant classifications.

Due to the small size of our data set, we employed considerable data augmentation in our approach. Our approach used flipping, rotation, zooming, horizontal and vertical shifts, and color normalization by subtracting the mean. More extensive techniques for generating data such as GANs were also considered. However, research into similar classification problems on constrained data sets showed they give comparable performance, with traditional augmentation sometimes outperforming GANs [7].

C. Architecture

Due to the limited amount of data we possess for our classes, a transfer learning method was employed to take advantage of already learned features from a model trained on thousands of images [10]. Our neural network builds off of a ResNet50 model that has been trained on ImageNet. Our model was implemented using Python and Keras, employing Keras's provided ResNet50 model with its pre-trained ImageNet weights.

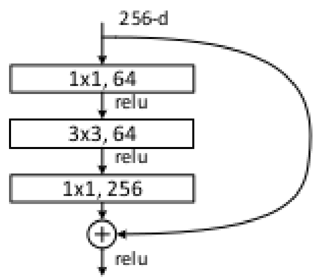

Fig. 2: ResNet50 Block consisting of 3 layers

The inputs to the model are 300x400x3 vectors representing the grass images with color. The design of our implementation uses ResNet50 models with the final output layer removed, followed by several fully connected layers that allow the model to learn for the domain of our grass recognition problem, similar to works done by [1]. Three fully connected layers with ReLU activation follow the transfer model, with a final output layer employing a sigmoid function, as our classifier is binary. Because we have such a limited amount of data, we implemented dropout between the fully connected layers with a percentage of 0.5 to try and prevent over-fitting of the data.

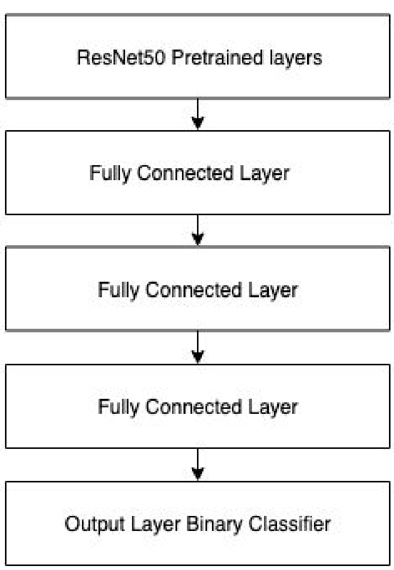

Fig. 3: Architecture

We chose to use stochastic gradient descent as our optimizer, with a learning rate of 0.0001. Similar to [9], we initialize our final layers to be random and allow them to train from scratch. However, unlike [9], we completely froze the weights for our starting ResNet50 model's layers in order to speed up training while hopefully achieving a similar effect.

IV. Results

After running the neural network model on our data set for 100 epochs, we obtained final accuracies of 0.89 on the training set and 0.50 on the validation set. Generally, the accuracy of the model on the training set improved over time with mild oscillation. In contrast, the accuracy on the validation set oscillated wildly and showed no clear sign of improvement over time. The discrepancy in the accuracies can be attributed to the small size of our data set, which led to a pronounced trend of overfitting to our training set of data. The model was successfully training on the training set shown in figure 4 by the loss decreasing and the accuracy increasing. However, since the model was overfitted, we do not see any improvement for the validation set.

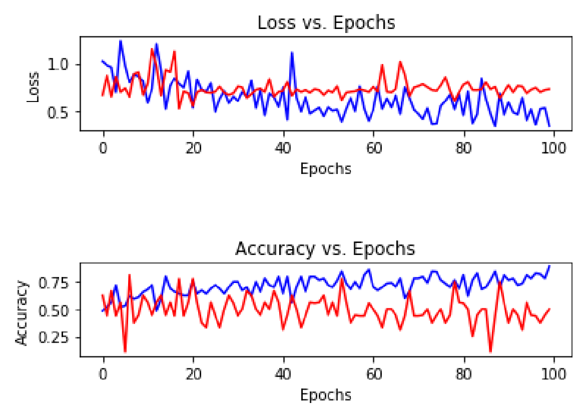

Fig. 4: Loss and Accuracy Graphs

V. Discussion

The results of our neural network on the validation set did not achieve the accuracies that we had hoped for. This is due to overfitting because of the limited number of images that we have in our dataset. In order to improve the results of our neural network, we would need a larger dataset with more images representing each class. The design of our neural network took into consideration the fact that we had limited images for each class, using transfer learning to leverage the learned weights of networks that had been trained on larger datasets, similar to the implementation done by [8]. However we did not see the same positive results as [8] did, this is because while they too were dealing with a limited number of images, our dataset suffered from significantly less diverse images and contained a smaller number for each of our classes.

During training we saw a decrease in the loss function with a corresponding increase in accuracy over time in the training set, showing that learning was occurring. However, validation results were often no better than random guessing, which can be attributed to overfitting produced by having limited data.

The neural network seemed to work correctly and we believe that simply training on a larger data set will drastically improve our results. In addition, the data set must include pictures of grass against a plain background. Most of the images we had were a few weeks old, so most of the grass contained dead leaves or flowers. Additional images would need to be of grass with a similar state in order to preserve our current classes, since fresher grass might have visual differences with dead grass within the same class. Alternatively, the dataset that we used for this paper could be entirely replaced by a dataset containing live grasses, where visual differences between flowering and grassy grasses would be more prominent.

VI. Conclusion

We created a neural network based on the ResNet50 model to be able to determine whether an image of grass was flowering or grassy. Due to the lack of sufficient data, even after image augmentation, are network was over-fitted to the training set. This led to a low accuracy in our validation set. However, we believe if we are able to obtain more data, this model can be trained to be much more accurate to be able to determine whether a grass image is flowering or grassy. In addition, we can extend the classification groups and identify whether a species is invasive or not, but this would require a much more robust dataset.

References

[1] Cagatay Catal Hamdi Yalin Yalic Huseyin Temucin Bedir Tekinerdogan Aydin Kaya, Ali Seydi Keceli. Analysis of transfer learning for deep neural network based plant classification models. *Science Direct*.

[2] Mads Dyrmann, Henrik Karstoft, and Henrik Skov Midtiby. Plant species classification using deep convolutional neural network. *Biosystems Engineering*, 151:72 – 80, 2016.

[3] Australian Government. Impact of weeds.

[4] K. Kerlin. Grasslands more reliable carbon sinks than trees. 07 2018.

[5] B. L. Steward L. Tang, L. Tian. Classification of broadleaf and grass weeds using gabor wavelets and an artificial neural network. *American Society of Agricultural Engineers*, 46(4):1247–1254, 2003.

[6] Rick L Lawrence, Shana D Wood, and Roger L Sheley. Mapping invasive plants using hyperspectral imagery and breiman cutler classifications (randomforest). *Remote Sensing of Environment*, 100(3):356–362, 2006.

[7] Jason Wang Luis Perez. The effectiveness of data augmentation in image classification using deep learning. 12 2017.

[8] Mark F. Hansen Wenhao Zhang Melvyn L. Smith Lyndon N. Smith, Arlo Bryne. Weed classification in grasslands using convolutional neural networks. 09 2019.

[9] Hoo-Chang Shin, Holger R Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M Summers. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE transactions on medical imaging*, 35(5):1285–1298, 2016.

[10] Qiang Yang Sinno Jialin Pan. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*.

[11] United States Geological Survey. Invasive grasses, vegetation, and weeds.