International Journal of Engineering & Technology, 7 (3.3) (2018) 119-121

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper



Image recognition in the artificial agriculture officer

B Lokesh[1]*, Ravoori Charishma[2], Natuva Hiranmai[3]

[1] *Assistant Professor, Computer Science & Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad.*
[2] *B.E, Computer Science & Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad*
[3] *B.E, Computer Science & Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad*
**Corresponding author E-mail: Lokeshjoel@gmail.com*

Abstract

Farmers face a multitude of problems nowadays such as lower crop production, tumultuous weather patterns, and crop infections. All of these issues can be solved if they have access to the right information. The current methods of information retrieval, such as search engine lookup and talking to an Agriculture Officer, have multiple defects. A more suitable solution, that we are proposing, is an android application, available at all times, that can give succinct answers to any question a farmer may pose. The application will include an image recognition component that will be able to recognize a variety of crop diseases in the case that the farmer does not know what he is dealing with and is unable to describe it. Image recognition is the ability of a computer to recognize and distinguish between different objects, and is actually a much harder problem to solve than it seems. We are using Tensorflow, a tool that uses convolutional neural networks, to implement it

Keywords: *Android application; Artificial agriculture officer; Convolutional Neural Networks; Image Recognition; Tensor flow.*

1. Introduction

Image recognition and image classification is an easy task for anyone with a functioning brain. However, for a computer, it is a very difficult problem. As a human grows up, the brain subconsciously starts collecting visual information and making a large data store of objects and their relationships with the environment. If a child constantly comes across a variety of cars, eventually he will be able to distinguish between the different types, brands, and models. It would be naturally assumed, that if a child is able to learn to classify objects quickly, that it should be quite easy for a machine to as well. However, computers have a hard time differentiating images since it has to consciously take all factors into consideration such as lighting, size, colour, etc.
Image recognition generally has six main steps:

- Image Formatting: This step consists of actually acquiring the image and converting it into a digital format
- Conditioning: There are many parts of an image that are unnecessary for image recognition that were either added to the picture (as noise) during the process of digitization or they form part of the background. The conditioning step supresses or normalizes these uninteresting elements and highlights the more interesting parts.
- Labelling: Images generally have informative patterns that can be used to identify various objects. Patterns consist of adjacent pixels that share certain characteristics such that it can be deduced that they belong to the same structure. Edge detection techniques use the assumption that continuous adjacent pixels that have a great difference in intensity or colour mark the boundary between objects or objects and the background. Since, all discovered edges may not be important, the process of Thresholding filters out most of the insignificant edges. The edges considered important are then labelled.
- Grouping: Grouping turns the primitive edges found in the labelling step into lines by determining which of the edges belong to the same spatial event. In the first few steps the image is considered as a digital picture (consisting of pixel information). However, from grouping, the information regarding which spatial event that each pixel belongs to is stored in a logical data structure.
- Extracting: Extracting refers to feature extracting which includes creating a list of properties the each of the groups of pixels in the same spatial event. This list can consist of area, grey tone moments, inscribing circle, a group's centroid, etc. The properties of each group depend on whether it is considered an arc or a region. For a region, it is more useful to make note of the number of holes. In the case of an arc, its curvature would be important to know. Feature extraction also takes topological relationships among the objects into consideration. Do they touch? Where are they in relation to one another? Etc.
- Matching: The final step is to actually recognize the elements in the image. This is done by comparing the objects to a prior store of models and finding the best match (Template matching).

2. Image Recognition Techniques

2.1. Artificial Neural Network

Artificial Neural Networks are computer systems that were designed based on the working of the human brain. It uses artificial

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

neurons to model the tens of thousands of complex connections of biological neurons in the brain. In real life, a brain has billions of neurons each forming hundreds of paths as humans learn new things. In ANNs there are only a few layers of neurons that are much simpler each having few connections. Each connection (modelled after biological synapse) between the artificial neurons transmits signals from one to another. The recipient of the signal processes it and produces a response that is sent to the next neuron in the network. At the end layer an output is generated.
ANNs "learn" by way of examples. For example, in image recognition, they will look at hundreds of images of an object (without any prior knowledge about it) and learn to identify the same object in other images. The performance and accuracy of such systems depends on the structure of the network and the number of nodes.
Advantages:
- It is a non-parametric classifier
- It is a data driven self-adaptive technique
- It efficiently handles noisy inputs
- The computation rate is high

Disadvantages:
- It is semantically poor
- Training is time consuming
- Over-fitting
- It can be difficult to choose the network architecture

2.2. Decision Trees

Decision trees are classification support tools that use a tree like data structure to help make decisions. They use various features of a given data (training data set) to learn about the most probable outcome for a test data set. Decision trees consist of a collection of nodes and branches in a tree like structure. Each internal node represents a variable with different representations of it. Each branch represents each of the representations. The leaf nodes represent class labels. In image recognition, the goal is to build a tree that accurately classifies the given images.
Advantages
- Handles nonparametric training data
- Does not require extensive design or training.
- Provides hierarchical associations
- Provides a set of rules that are easy to interpret
- Simple
- Computationally efficient

Disadvantages
- Tree structure is prone to sampling
- Splitting is locally greedy
- Instability
- Complexity
- Cost
- Prone to over-fitting

2.3. Support Vector Machines

A support vector machine is a supervised learning model that analyzes data used for classification and regression analysis. Given a set of training examples, an SVM can build a model that exclusively categorizes different types of objects. It is considered to be a probabilistic binary linear classifier. Once the model has been thoroughly trained, new data sets can be given to the machine to be classified.
Support Vector Machines can also use unsupervised learning to build classification models. In unsupervised learning, the training data is absent. In other words, prior labels are not given to the objects. The machine "learns" by looking for similarities between the given data points and putting them in together in a way that makes the most sense.
A SVM separates the different categories of objects by creating a hyperplane or a set of hyperplanes in the n-dimensional space. A good separation is one that has the largest distance from each of the categories.
Advantages
- It offers flexibility over the choice in the form of the threshold
- Contains a nonlinear transformation
- It provides a good generalization capability
- The problem of over fitting is eliminated
- Reduction in computational complexity

Disadvantages
- Result transparency is low
- Training is time consuming
- Structure of algorithm is difficult to understand
- Determination of optimal parameters is not easy when there is nonlinearly separable training data.

2.4. Fuzzy Classification

Fuzzy classification is the categorization of objects into fuzzy sets based on a quantity called the truth value, which is defined by a membership function. The membership function is a function that indicates whether an object belongs to a certain set and by how much. If the result generated by the function is 1, the object completely belongs to the set. If the truth value generated is 0, it shows that the object does not have any characteristics that may associate it with a particular set. Any value in between shows a partial membership. A class is a set that is defined by a certain property, and all objects having that property are elements of that class. Classification is the process of grouping individuals having the same characteristics into a set. In fuzzy image recognition, stochastic associations are used to describe the qualities of a picture. These associations are used to determine which fuzzy set the image may belong to.
Advantages
- Can handle uncertainty
- Characteristics are described by discerning stochastic relationships.

Disadvantages
- Without prior knowledge the produced output is not efficient
- Precise solutions depend upon the direction of decisions

3. Tensorflow

Tensorflow is a tool for image recognition mainly designed using deep neural network models, specifically Convolutional Neural Networks. The latest model built by tensorflow implementing image recognition is the Inception V-3 model. It is trained to recognize various images from the ImageNet data set, which is an enormous image database, modelled after the WordNet, where each node represents an object defined by hundreds and thousands of images. Using Tensorflow as a tool has an advantage as it allows for transfer learning. Transfer learning refers to training an image recognition model that has already been trained on another data set or problem. This means that most of the processing that allows the model to distinguish objects has already been done. To make the model to recognize a new set of images, only the top-most layer must be re-trained. An alternative to the Inception V-3 model is the MobileNet, which gives more importance to optimization and being small than it does to accuracy.

Once the dataset has been created and imported into the Tensorflow directory, the next step is to create bottlenecks, which is the informal term used to refer to the layer of the network right before the last, which performs the actual classification. After all the bottlenecks have been created, the training of the final layer starts. The training works by supplying the cached value for each image into the bottleneck layer and the label for each image into a node called Ground Truth. These two inputs are necessary to calculate performance metrics of the training process

Table.1: Performance Metrics of Model Training

Performance Metric	Definition
Training Accuracy	Percentage of images from current training batch labelled with correct class
Validation Accuracy	Precision of labelling on randomly selected images from another set
Cross Entropy	A function that indicates the progress of learning

To know the real measure of how well the model is performing is to see how well it distinguishes images not in its training data set. This is the validation accuracy. If the validation accuracy stays low while the training accuracy increases, this means that the network is over-fitting. The main purpose of the training is to make the cross entropy low.

The Tensorflow training process consists of 4000 training steps, where each step takes 10 random images from the training data set, finds their bottlenecks and sends them to the final layer to get the model's predictions. These predictions are then checked for validity by comparing them to the genuine labels. The weights in the network are updated accordingly. As the process continues, the accuracy increases. At the end of the script, a final test of accuracy is run using a small sample of the training data set that was kept separately. This evaluation is a marker of how well the trained model will perform the required classification

3.1. Convolutional Neural Networks

A Convolutional Neural Network (CNN) is essentially a network that uses multiple copies of the same neuron. This allows for computationally large networks with a fairly low number of parameters (which are the values describing the behaviour of the neurons). This is similar to calling the same function many times in computer programming.
CNNs require much less pre-processing than other image classification algorithms, which means that the network learns to process its own data before classification whereas other traditional algorithms need to be hard-engineered.

A CNN consists of many hidden layers that can be made up of convolutional layers, pooling layers, fully connected layers, and normalization layers**.**

3.1.1. Convolutional Layer

The convolution layer performs a convolution operation on the input, and passes the result on the next layer. The convolution operation is a function that resembles the response of a biological neuron when it receives visual stimuli. Each of these convolutional neurons only processes data in its own receptive field. This type of operation offers a huge advantage over a fully connected feed-forward neural network in image classification, as a it requires fewer neurons and parameters.

3.1.2. Pooling Layer

Global or Local Pooling refers to the combination of outputs of one layer into a single neuron, being used as the input to the next.

- Max Pooling uses the highest value from each of a cluster of neurons.
- Average Pooling uses the average value of each cluster of neurons.

3.1.3. Fully Connected Layer

Fully connected layers are those where every neuron in one layer is connected to every neuron the next.

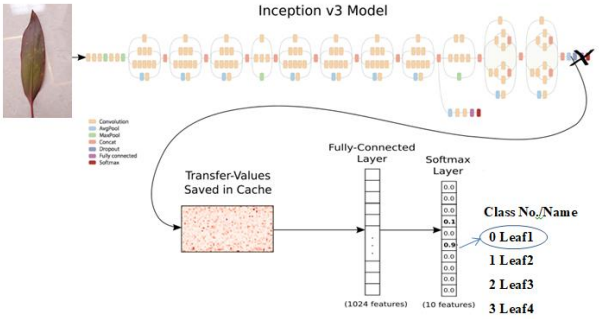


Fig 1: Training Model Using Tensorflow

4. Conclusion

This paper attempts to provide an understanding of Image Recognition, the different types of techniques used to achieve it, and the tool Tensorflow, and its performance metrics. It also goes into depth about the specific deep neural network being used, the convolutional neural network. Tensorflow is being used to implement the classification of images in the android application, the Artificial Agriculture Officer. The application uses the provided deep neural network to train the model to recognize an image set consisting of a variety of labelled plant diseases.

References

[1] Sharwari Gaikwad, Rohan Asodekar, Sunny Gadia, and Vahida Z. Attar, "AGRI-QAS question-answering system for agriculture domain," in Advances in Computing, Communications and Informatics, 2015 International Conference.

[2] Pooja Kamavisdar1 , Sonam Saluja2 , Sonu Agrawal3" A Survey on Image Classification Approaches and Techniques" International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 1, January 2013

[3] Sonal Athavale,Neelabh Sao," Classification on moving object trajectories", International Journal of Advanced Technology & Engineering Research ISSN No: 2250-3536 Vol 2, Issue 2, May 2012

[4] FatihErtam, GalipAydin, "Data classification with deep learning using Tensorflow" International Conference on Computer Science and Engineering (UBMK), October 2017

[5] JaspreetKaur and Vishal Gupta, "Effective Question Answering Techniques and their Evaluation Metrics", International Journal of Computer Applications (0975 – 8887) Vol 65– No.12, March 2013

[6] .JianxinWu,"Efficient HIK SVM Learning for Image Classification", IEEE Transactions on Image Processing, Vol. 21, No. 10, October 2012

[7] XiaohongYu, and Hong Liu Huangshan, "Image Semantic Classification Using SVM in Image Retrieval" P. R. China, 26-28, December 2009

[8] Y. Koren, "Collaborative filtering with temporal dynamics," in Proc.KDD, pp. 447–456.Paris, France, 2009