Design of Deep Learning based fruit classification system platform

SooHo Jeong[1], **Hyun Yoe**[*]
1,* Department of Information and Communication Engineering, Sunchon National University, Suncheon-si, Jeollanam-do, Republic of Korea

Abstract - *Recent developments in artificial intelligence technology have shown good results in various fields. This paper aims to construct and introduce a new platform for classifying fruits applicable to small scale farmers based on Deep Learning. We used Google's open source library interface "Tesorflow" to classify a single fruit grade, and the accuracy of the constructed model was 0.73%. Using this system in the future, producers will be able to create higher value with the same fruit yield without relying on the imported fruit classification system, and consumers will be able to choose quality fruits that are price appropriate.*

Keywords: Deep Learning, Fruit Classification System, Tensorflow, Inception v3 model, Transfer Learning

1 Introduction

In Korea, it is essential to have a mechanism that automatically classification fruit quality due to labor cost increase due to lack of labor in rural production. In large scale farms, image processing based fruit classification system using computer vision system is spreading. However, this image processing based fruit classification system is very expensive due to dependence on imported products, and it is difficult to apply real farms because it does not fit the shape of fruits in Korea. [1][2]

Therefore, the system proposed in this study is a Deep Learning based fruit quality classification system that can be used in small farms. Recently, Deep Learning technology for the implementation of artificial intelligence is rapidly developing as the field of artificial intelligence advances. Currently, libraries using various algorithms are open sourced to utilize Deep Learning. [3] Among them, "Tensorflow", released by Google in 2015, is an open source library interface using a machine learning algorithm. Tensorflow is used in many areas such as voice recognition, computer vision, robotics, information retrieval, and natural language processing. [4]

The proposed system collects image data through Raspberry Pi camera and performs image learning by retraining the inception v3 model of Google's open source library, Tensorflow. Based on the completed Tesorflow model, it's graded accordingly.

2 Related Work

This chapter briefly describes the preliminary work that has been developed for this paper.

2.1 TensorFlow

TensorFlow is an open source software library for machine learning used in Google products. It was created by the Google Brain team for Google's research and product development and was released as an Apache 2.0 open source license on November 9, 2015. TensorFlow can run on multiple CPUs and GPUs in desktop environments such as Android and iOS as well as 64 bit Linux and MacOS. It is written in Python and C ++. Currently, TensorFlow is used in many fields such as voice recognition, computer vision, robotics, information retrieval, and natural language processing. [4]

2.2 CNN

CNN (Convolution Neural Network) is a type of Deep Learning that has achieved innovative results in various fields such as natural language processing and image recognition. CNN places a convolutional layer and a pooling layer in the hidden layer between the input and output layers. In these two layers, the process of lowering or sampling the resolution of the image is repeated. The convolution layer applies a weight filter to a part of the input image, and extracts features that may be helpful for classification, thereby creating a feature map. The pooling layer reduces the feature map by sub sampling the most important part of the feature map obtained from the convolutional layer. It reduces the size of the data while maintaining the characteristics, thereby preventing the change of the data due to the location change of the data and enhancing the performance of the neural network by reducing the data size. Based on these extracted features, classification is performed. [5]

2.3 Inception v3 Model

The Inception v3 model is a state-of-art image recognition model created by Google. Using the ImageNet

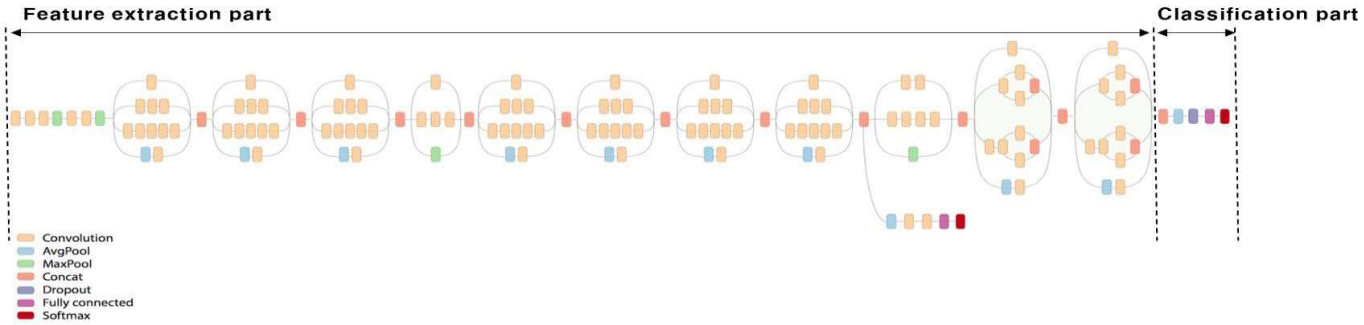

Figure 1. Composition of Inception v3 model

Large Visual Recognition Challenge 2012 data, the training model has a top-5-error rate of only 3.49%. For reference, when a human classified the ImageNet Large Visual Recognition Challenge 2012 data, the result was a top-5-error rate of 5.1%. [6]

The Inception v3 model consists of two parts as shown in the following Figure 1. [7]

- Feature extraction part with a convolutional neural network.

- Classification part with fully connected and softmax layers.

3 Main Subject

3.1 Hardware and System Configuration

The image acquisition hardware used for this study is Raspberry Pi Camera. The system uses Raspberry Pi 3 as the main controller to take an image of the fruit with the camera, and the first step runs a small neural network in TensorFlow to detect whether the image is a fruit. The image is then passed to the TensorFlow CNN neural network learning on a Linux server for further classification. The TensorFlow code is an Inception v3 model code provided by Google that modifies the convergence, pooling, and network design to match the number of classes and classes of pixels in the image, with minor modifications to the last layer.

3.2 Image Learning

We used the Inception v3 model, a new model of image recognition developed by Google, as a Tensorflow code for our experiments. [6]

Finding a suitable Neural Networks structure from the ground up for a new task each time, and learning it new is a lot of time and effort. Therefore, it is effective to bring the model structure and parameters of a proven model (eg, Inception model made by Google for image recognition) that has already worked well for a specific task, and retraining it to a new data set.

This process is called “Transfer Learning”. This paper categorizes crops through Transfer Learning. [7][8]

The following Figure 2 is the learning data for this paper.

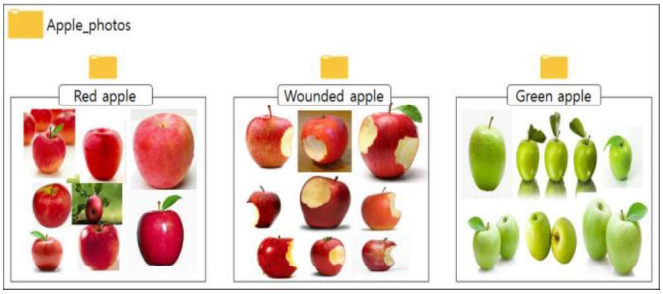

Figure 2. Image learning data (Apples)

When constructing the training data set, at least 30 image data must be prepared for each label, and there is no image size limitation.

3.3 System implementation

The following Figure 3 shows the overall system configuration of fruit classification image classification using Tensorflow.

First, to classify the image data, first create a folder with a different name in each folder and store the image data corresponding to the folder name. At this time, the jpg (jpeg) image does not need data preprocessing, so the file format is set to .jpg (jpeg).

Second, each image data collected is modified by code of classifying fruits by transfer learning of Google Inception v3 model, and image learning is carried out.

Finally, based on the completed Retraining_Inception v3 model, the test proceeds.

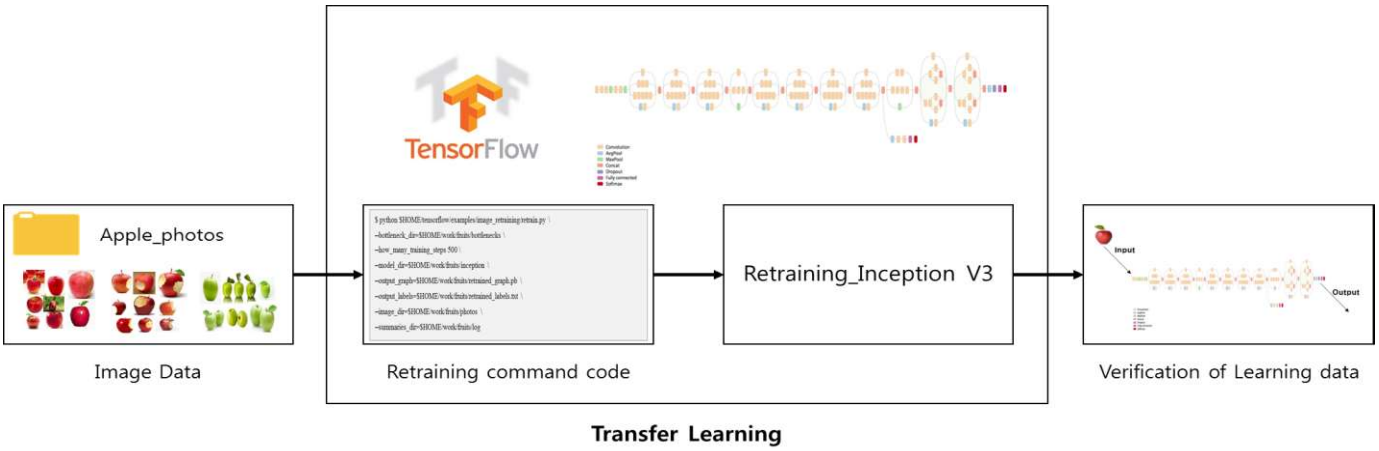

Figure 3. Tensorflow based fruit classification system whole structure and process

Figure 4 below shows the command code for Tensorflow for Inception v3 model retraining. Based on this code, it reads the fruits image in the folder and starts learning.

```
$ python $HOME/tensorflow/examples/image_retraining/retrain.py \
--bottleneck_dir=$HOME/work/fruits/bottlenecks \
--how_many_training_steps 500 \
--model_dir=$HOME/work/fruits/inception \
--output_graph=$HOME/work/fruits/retrained_graph.pb \
--output_labels=$HOME/work/fruits/retrained_labels.txt \
--image_dir=$HOME/work/fruits/photos \
--summaries_dir=$HOME/work/fruits/log
```

Figure 4. Tensorflow command code for retraining

The interpretation of the command code for retraining is as follows.

bottleneck_dir = Folder to convert and save pictures to learn

how_many_training_steps = Step setting

model_dir = Path to download the inception model

output_graph = Learned binary file (.pb) storage path for inference

output_labels = Label file storage path for inference

image_dir = Path where the original image is stored

summaries_dir = Path to store log file for tensorboard

4 Result

We carried out experiments based on the Inception v3 model, which was repaired. Experimental image data were collected with Red apple, wounded apple, and Green apple, and 30 images per label (image learning error when 30 image data counts or less) were collected. And we set the training step to 500.

```
red apple(score = 0.72952)
wounded apple(score = 0.26121)
green apple(score = 0.00927)
```

Figure 5. Experiment result

As a result of the experiment, it was observed that red apple was classified into 0.72959% probability as shown in the following Figure 5. However, there is a 0.26121% probability of a wounded apple of the same color except green apple, and the classification accuracy of this experiment is not so high. As a result of repeated experiments, it was confirmed that the more the number of training steps and the number of image data according to computer specifications, the higher the classification accuracy. In the future, it will be necessary to build up a basis for high classification accuracy by increasing the specifications of the experimental computer and collecting more image data.

5 Conclusions

This paper developed a Deep Learning based crop screening platform to classify fruit grades. We use Retraining_Inception v3, which is a transfer learning method of Inception v3, a Google image data training model. Image data was collected by referring to Google search image and ImageNet. [9]

An error due to image data occurred during the collection of image data through Retraining_Inception v3. This error was caused by insufficient number of image data per label due to the characteristics of the model, and it was found through experiment that at least 30 ~ 50 pieces of data per label should be smoothly operated by learning.

Using this system in the future, producers will be able to create higher value with the same fruit yield without relying on the imported fruit classification system, and consumers will be able to choose quality fruits that are price appropriate.

In addition, due to differences in learning speed and stability according to computer specifications due to the nature of Deep Learning, future research will need to show the accuracy of fruit classification according to the specifications of the experimental computer and the number of crop image data.

6 Acknowledgement

"This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-2013-1-00877) supervised by the IITP (Institute for Information & communications Technology Promotion)"

7 References

[1] Mehrdad Baigvand, Ahmad Banakar, Saeed Minaei, Jalal Khodaei, Nasser Behroozi-Khazaei ″Machine vision system for grading of dried figs″, Computers and Electronics in Agriculture, Vol. 119, Issue C, 158-165, Nov 2015

[2] SooHo Jeong, Hyun Yoe. ″Development of Machine Vision based fruit quality screening system″, Korean Institute of Communication Sciences, 100-101, Nov 2017

[3] Dohyun Kim, Jungheun Lee, Jihoon Yang. “Effective Structure of Tensorflow Multimodal Deep Learning for Video Classification”, Korea Information Science Society, 1842-1844, Dec 2016

[4] Tensorflow, https://www.tensorflow.org/

[5] Ciresan, Dan Claudiu, et al. "Convolutional neural network committees for handwritten character classification." Document Analysis and Recognition (ICDAR), 2011 International Conference on. IEEE, 18-21, Nov 2011.

[6] Image Recognition, https://www.tensorflow.org/versions/r0.8/tutorials/image_recognition/index.html#usage-with-python-api

[7] Codelabs “Image Classification Transfer Learning with Inception v3”, Google Developers, Jan 2017

[8] Image retraining, https://www.tensorflow.org/tutorials/image_retraining

[9] http://www.image-net.org/