Machine Learning Engineer Nanodegree

Capstone Proposal

Matthew Crummey
7-20-2019

Domain Background

The continued integration of automation into everyday life will require further refined computer vision. Computer vision is the process in which computers are trained to interpret and understand the visual world [1]. Some of the industries that use computer vision are: Manufacturing, Health Care, Insurance and Retail [1, 2]. One of the applications of computer vision within retail is the recognition of food, which is used in cashier-less stores like Amazon GO [2]. By being able to accurate identify fruit from images taken from video cameras, a store can know two main things: 1. When a customer has "purchased" it and 2. When the shelf is running low on a fruit. My person motivation for the project is that since learning about Computer vision I have always been fascinated with how computers can be programmed to identify objects in images.

Problem Statement

The main objective of this project is to use machine learning to be able to identify fruits within images, thus this is a computer vision problem. The input is an RGB picture of a piece of fruit and the output is the prediction vector of what type of fruit it is. The problem is to train an algorithm that when given an image of a fruit, can accurately identify what type of fruit it is.

Dataset

The dataset used is the fruit-360 dataset from Kaggle.com (https://www.kaggle.com/moltean/fruits) [3]. This data set has approximately 70000 100px by 100px RDB images (about 6Kb each) of fruit that are already split into a training (53177 images) and testing (17845 images) set. I will split the training set into a training set and a validation set to allow the model to take advantage of validation during training. Due to the fact that the images are relatively small in size, all images within their respective categories will be used, which creates the breakdown seen below in Table 1.

Table 1: Number of Images used in each category

	Images Used
Training	45200
Validation	7977
Testing	17845

Both the training and testing have subfolders with the name of each fruit as the name of the subfolder. Within the dataset different pieces of fruit are given different numbers (Apple Golden 1 vs Apple Golden 2); for this project I will strip out the numbers so all Apple Golden will be train, validated and tested the same. Originally, without the above trimming, there are 103 targets; after the trimming there are 87 targets. The highest count of any one target in either the training set or target set is 5% of the total images. I do not believe this upsets the distribution of the images such that it would cause the model to become overfitted. A full table with the count of each target in the training and testing sets can be seen in Appendix 1 Table 2.

Solution Statement

The solution to this problem is to accurately determine which fruit is displayed in the image. This problem will be best solved by using a convolutional neural network. A convolutional neural network (CNN) is ideal for image classification problems because by using convolutional layers, groups of pixels are considered at once, which can allow feature detection to occur[4]. CNNs use 3D representation of 2D images where each layer is one of the colour channels [4].

Benchmark Model

The benchmark model for this problem will be a one-layer CNN, in that only one convolution layer will be used, followed by one pooling layer, one dropout layer and one densely connected layer. The benchmark model will be used to test:

1. the accuracy of the prediction of the testing set
    a. Including with what confidence each prediction is made
2. the accuracy of predicting internet sourced images

Evaluation Metrics

The first evaluation metric for this project will be the Accuracy Score of the testing set, after training and validation have been done. A deeper dive will be done into the numbers that are used to compute the Accuracy Score as well, to see how confident the model is in its prediction. The second evaluation metric will be the Accuracy Score of internet sourced images. This will be done because the images used for testing and training have a white background and I want to see how well the model does with images with backgrounds. As stated above, it is not believed that the slight imbalance in terms of representation of certain fruits (ie. Tomatoes and Grapes) will cause an issue in either the training, testing, or the accurate computation of the accuracy.

Project Design

The first part of any CNN design is to process the images into numerical representation. For 100px by 100px RGB images this means creating a 100 by 100 by 3 3D matrix, with each layer representing one colour, and each cell of each layer will be normalized to have a number between 0 and 1 based on how strongly that pixel was coloured in the original image. The targets (names of the fruits) is created by taking the name of the folder was in: ie. Test/Apple Braeburn/3_100.jpg is a Test image of a Braeburn Apple) As mentioned above, the images have already been split into a Training and Test set, and I will further split the Training set into a validation and training set.

Once this preprocessing has been done, the actual model can be created. The benchmark model has the following layers:

```
_________________________________________________________________
Layer (type)             Output Shape          Param #
=================================================================
conv2d_2 (Conv2D)        (None, 100, 100, 16)     448
```

```
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 33, 33, 16)       0
_________________________________________________________________
dropout_4 (Dropout)        (None, 33, 33, 16)       0
_________________________________________________________________
flatten_2 (Flatten)        (None, 17424)          0
_________________________________________________________________
dense_4 (Dense)           (None, 87)            1515975
=================================================================
Total params: 1,516,423

Trainable params: 1,516,423

Non-trainable params: 0
```

After the model's creation it can be complied and fitted using the training and validation images created during the pre-process. Once the fitting has been done, the test accuracy can be computed using the testing images.

The final model will be more complex than the benchmark model but will follow roughly the same format with several convolution layer-pooling layer pairs, followed by dropout, flattening and densely connected layers.

References

[1] SAS, "Computer Vision," 2019. [Online]. Available: https://www.sas.com/en_us/insights/analytics/computer-vision.html. [Accessed 20 07 2019].

[2] CB Insights, "Beyond Amazon Go: The Technologies And Players Shaping Cashier-Less Retail," 09 10 2018. [Online]. Available: https://www.cbinsights.com/research/cashierless-retail-technologies-companies-trends/#why. [Accessed 20 01 19].

[3] M. Oltean and H. Muresan, "Fruit recognition from images using deep learning," *Acta Universitatis Sapientiae, Informatica,* vol. 10, no. 1, pp. 24-42, 2018.

[4] skymind, "A Beginner's Guide to Convolutional Neural Networks (CNNs)," [Online]. Available: https://skymind.ai/wiki/convolutional-network. [Accessed 20 07 2019].

Appendix 1: Target Counts

Table 2: Count of each target in both the full training and testing set

Target	Training	Training Percent	Testing	Testing Percent
Apple Braeburn	492	0.93%	164	0.92%
Apple Crimson Snow	444	0.83%	148	0.83%
Apple Golden	1465	2.75%	489	2.74%
Apple Granny Smith	492	0.93%	164	0.92%
Apple Pink Lady	456	0.86%	152	0.85%
Apple Red	1413	2.66%	472	2.64%
Apple Red Delicious	490	0.92%	166	0.93%
Apple Red Yellow	1164	2.19%	383	2.15%
Apricot	492	0.93%	164	0.92%
Avocado	427	0.80%	143	0.80%
Avocado ripe	491	0.92%	166	0.93%
Banana	490	0.92%	166	0.93%
Banana Lady Finger	450	0.85%	152	0.85%
Banana Red	490	0.92%	166	0.93%
Cactus fruit	490	0.92%	166	0.93%
Cantaloupe	984	1.85%	328	1.84%
Carambula	490	0.92%	166	0.93%
Cherry	1230	2.31%	410	2.30%
Cherry Rainier	738	1.39%	246	1.38%
Cherry Wax Black	492	0.93%	164	0.92%
Cherry Wax Red	492	0.93%	164	0.92%
Cherry Wax Yellow	492	0.93%	164	0.92%
Chestnut	450	0.85%	153	0.86%
Clementine	490	0.92%	166	0.93%
Cocos	490	0.92%	166	0.93%
Dates	490	0.92%	166	0.93%
Granadilla	490	0.92%	166	0.93%
Grape Blue	984	1.85%	328	1.84%
Grape Pink	492	0.93%	164	0.92%
Grape White	1943	3.65%	654	3.66%
Grapefruit Pink	490	0.92%	166	0.93%
Grapefruit White	492	0.93%	164	0.92%
Guava	490	0.92%	166	0.93%
Hazelnut	464	0.87%	157	0.88%
Huckleberry	490	0.92%	166	0.93%
Kaki	490	0.92%	166	0.93%
Kiwi	466	0.88%	156	0.87%

Target	Training	Training Percent	Testing	Testing Percent
Kohlrabi	471	0.89%	157	0.88%
Kumquats	490	0.92%	166	0.93%
Lemon	492	0.93%	164	0.92%
Lemon Meyer	490	0.92%	166	0.93%
Limes	490	0.92%	166	0.93%
Lychee	490	0.92%	166	0.93%
Mandarine	490	0.92%	166	0.93%
Mango	490	0.92%	166	0.93%
Mangostan	300	0.56%	102	0.57%
Maracuja	490	0.92%	166	0.93%
Melon Piel de Sapo	738	1.39%	246	1.38%
Mulberry	492	0.93%	164	0.92%
Nectarine	492	0.93%	164	0.92%
Orange	479	0.90%	160	0.90%
Papaya	492	0.93%	164	0.92%
Passion Fruit	490	0.92%	166	0.93%
Peach	1230	2.31%	410	2.30%
Peach Flat	492	0.93%	164	0.92%
Pear	492	0.93%	164	0.92%
Pear Abate	490	0.92%	166	0.93%
Pear Kaiser	300	0.56%	102	0.57%
Pear Monster	490	0.92%	166	0.93%
Pear Red	666	1.25%	222	1.24%
Pear Williams	490	0.92%	166	0.93%
Pepino	490	0.92%	166	0.93%
Pepper Green	444	0.83%	148	0.83%
Pepper Red	666	1.25%	222	1.24%
Pepper Yellow	666	1.25%	222	1.24%
Physalis	492	0.93%	164	0.92%
Physalis with Husk	492	0.93%	164	0.92%
Pineapple	490	0.92%	166	0.93%
Pineapple Mini	493	0.93%	163	0.91%
Pitahaya Red	490	0.92%	166	0.93%
Plum	1767	3.32%	597	3.35%
Pomegranate	492	0.93%	164	0.92%
Pomelo Sweetie	450	0.85%	153	0.86%
Quince	490	0.92%	166	0.93%
Rambutan	492	0.93%	164	0.92%
Raspberry	490	0.92%	166	0.93%
Redcurrant	492	0.93%	164	0.92%

Target	Training	Training Percent	Testing	Testing Percent
Salak	490	0.92%	162	0.91%
Strawberry	492	0.93%	164	0.92%
Strawberry Wedge	738	1.39%	246	1.38%
Tamarillo	490	0.92%	166	0.93%
Tangelo	490	0.92%	166	0.93%
Tomato	2627	4.94%	877	4.91%
Tomato Cherry Red	492	0.93%	164	0.92%
Tomato Maroon	367	0.69%	127	0.71%
Tomato Yellow	459	0.86%	153	0.86%
Walnut	735	1.38%	249	1.40%