Aim

The objective of this project is to classify a Fruits image dataset using Convolutional Neural Network into the following classes: (0) Apples, (1) Bananas, (2) Mixed Fruits, (3) Oranges.

Brief description of experimentation with layers in our CNN models

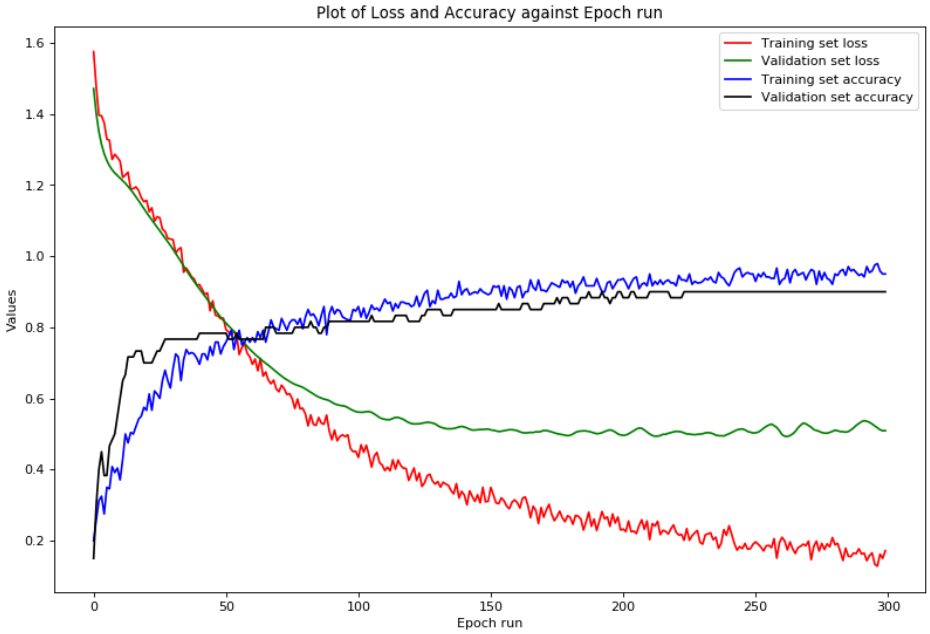
Performed the following to evaluate the impact on model accuracy:

1. Varying degrees of feature extraction: used different stackings of Convolutional (Conv2D) layers with varying number of filters and filter sizes
2. Varying degrees of dimensionality reduction on feature maps: used different stackings of Max Pooling layers with varying window sizes
3. Varying number of Dropout layers and fractions of neurons: added different stackings of Dropout layers with different dropout rates
4. Varying number of epochs

Specifications and Evaluation of our CNN models

	CNN Model 1	**CNN Model 2**	**CNN Model 3**	**CNN Model 4**
Model Specifications				
Conv2D Layers	Layer 1: 32 filters, size = 3x3	Layer 1: 32 filters, size = 3x3	Layer 1: 16 filters, size = 5x5 Layer 2: 32 filters, size = 3x3	Layer 1: 32 filters, size = 3x3 Layer 2: 64 filters, size = 4x4 Layer 3: 64 filters, size = 3x3 Layer 4: 128 filters, size = 5x5
MaxPooling Layers	3x3 window	6x6 window	Two 2x2 windows	Two 2x2 windows; padding = same
Dropout Layers	Layer 1: 0.5 Layer 2: 0.25	0.5	0.5	Layer 1: 0.25 Layer 2: 0.5
Flatten Layers	1	1	1	1
Dense Layers	64 units with ReLu activation function; 4 units with softmax activation function	4 units with softmax activation function	256 units with ReLu activation function; 4 units with softmax activation function	32 units with ReLu activation function; 4 units with softmax activation function
Epochs	100	300	100	100
Model Evaluation				
Training Loss	0.0627	0.1767	0.0241	0.0623
Training Accuracy	0.9917	0.9417	1.0000	0.9667
Validation Loss	0.6363	0.4992	0.7232	0.8913
Validation Accuracy	0.8500	0.9000	0.8833	0.9000
Training Time	8.3695s	20.2121s	9.8789s	20.8939s
Prediction Time	0.0375s	0.0409s	0.0403s	0.0459s

Plot of Loss and Accuracy of our best model (CNN Model 2)



Model with highest accuracy on the test data

The most accurate models are CNN Model 2 and Model 4 as they both have a prediction accuracy of 90%. However, CNN Model 2 is marginally faster to train than Model 4.

Reflection of lessons learnt in building our network

On Disproportionate Dataset:

Volume of input data/images									
Datasets	**Apples**		**Bananas**		**Mixed Fruits**		**Oranges**		**Total**
Train	75	31.3%	73	30.4%	20	8.3%	72	30.0%	240
Test	19	31.7%	18	30.0%	5	8.3%	18	30.0%	60

Across all 4 CNN models, images of mixed fruits were inaccurately predicted most of the time, while apples, oranges and bananas had more accurate predictions. This is due to the small proportion of mixed fruit images in the training dataset. Since the model was trained with only 20 images of mixed fruits (8.3% of the dataset), it is not able to predict images of mixed fruits as accurately compared to the other classes.

On Model Architecture and Accuracy:

CNN Model 2 consistently produces a higher accuracy than the other CNN models. It has a simple structure which utilises 1 Conv2D layer of 32 filters of size 3 pixels by 3 pixels, a pooling layer with pooling window of size 6x6, 1 dropout layer, 1 flatten layer and lastly, 1 dense layer with a softmax activation function. This suggests that a complex model architecture does not necessarily imply a greater accuracy. CNN Model 2 and 4 have similar accuracies despite Model 4 having a more complex architecture.

On Size of Input Image:

The CNN models were ran using two different input image sizes – 34px by 34px and 128px by 128px – to observe the impact of the size on accuracy and duration. The models using the larger image size had substantially longer training duration, but there were no significant improvements in accuracy. On the other hand, a smaller input image size of 34px by 34px is better as not only was training duration shorter, but accuracy is also comparable at 90%. Hence, a larger input image size does not necessarily lead to greater accuracy.