Transfer Learning Based Model for Classification of Cocoa Pods

Juliana R C P de Oliveira[1] Roseli Ap. Francelin Romero[2]

Institute of Mathematical and Computer Sciences,
University of São Paulo, Av Trabalhador São Carlense - São Carlos, Brazil
[1]jurcpdo@gmail.com, [2]rafrance@icmc.usp.br

July 5, 2018

Outline

Motivation

Cocoa plantations provide a unique challenge to automation.

- The trees cannot grow in regular shapes.
- The cocoa pods grow directly from the trunk of the tree.
- The plantation cannot be made in rows to provide space for systematic harvesting.

Proposal

Automated flying drones coupled with small harvesters.

- Capable of navigating unpredictable and thickly forested terrain.
- Effectively identifying cocoa pods that are ready to be harvested.
- Equipped with tools for removing the cocoa pod from the tree, to be collected by automated harvesters.

The drones will require software sophisticated enough to identify and harvest only ripe cocoa pods within an acceptable margin of error.

Identification

A deep learning model trained to identify ripe cocoa pods.

- Built upon a pre-existing model trained to identify the presence of a cocoa pod within a still image.
- A new feature is added to identify whether the image contains a cocoa pod that is ready to be harvested.

Issues:

- Requires large amounts of data to form a sophisticated model.
- Training a sufficiently sophisticated model necessitates powerful hardware and great lengths of time.

Transfer Learning

Due to lack of publicly available resources, Transfer Learning is used.

- Transfer Learning uses pre-existing weights trained in a related subject.
- Although it requires a big amount of data for the pre-training, the specific dataset can be composed of a few hundred objects.
- Architectural limitation, models are not interchangeable between pre-training and training.

Inception-Resnet-v2

A model that combines residual connections and the Inception modules.

- 185 convolutional blocks across its 50 layers and one completely dense block.
- Multiple convolutional filters processed with the same input, allowing the model to extract both general and local features.
- Residual Networks ensure an increase of performance after each layer.

Architecture

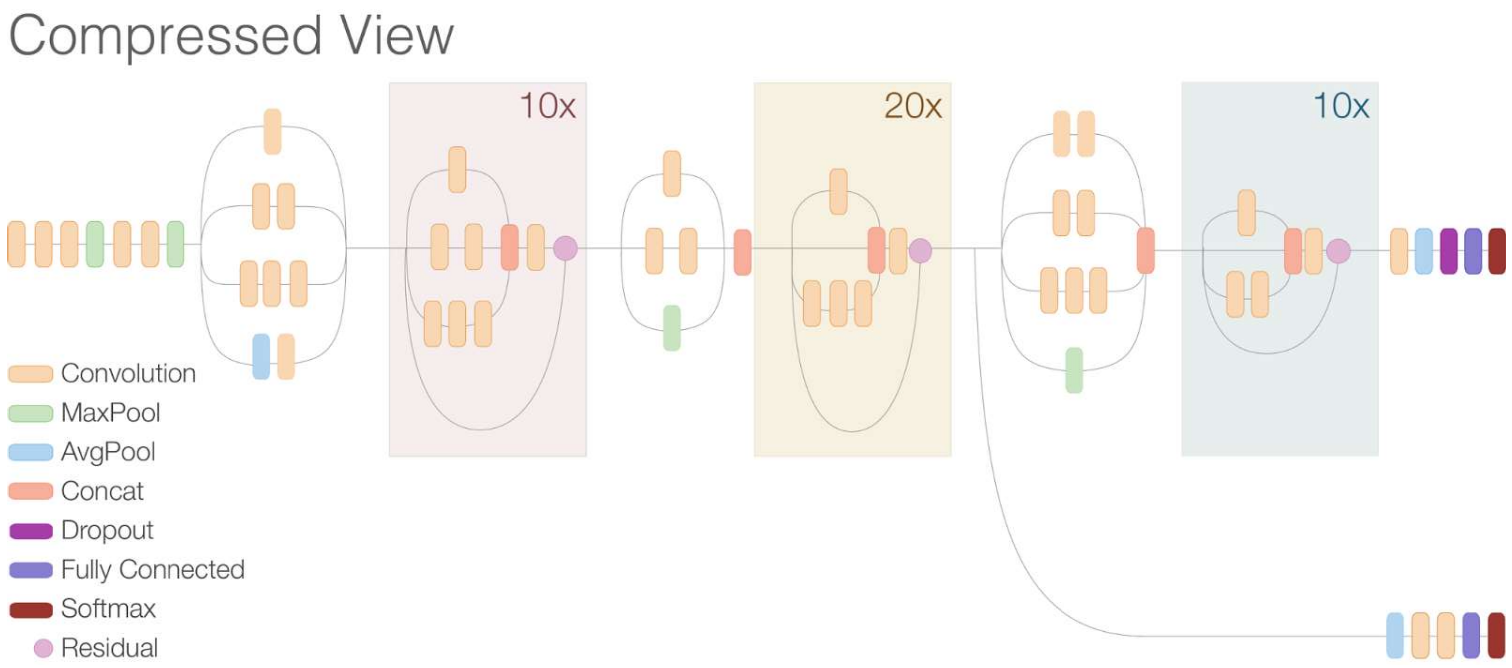

Figure: Inception-Resnet-v2 model architecture. Source: A. Alemi, Improving inception and image clas- sification in tensorflow, 2016. [Online]. Available: https://research.googleblog.com/2016/08/improving-inception-and-image.html

Database

Requirements:

- Images of cocoa pods at various stages of development.
- From as many different angles and settings as possible.

A total of 1243 images of un/ripe pods was collected.

- ImageNet (online, open source image database).
- Frames extracted from several video clips, which provided substantial variation.

Figure: Sample of cocoa images from the database.

Methodology

Three types of Transfer Learning were applied.

- **Total training**: Training of all layers using the weights of the checkpoint as the starting point.
- **Partial training**: Freezing the weights of most hidden layers, training the last convolutional reduction and upper layers (*dropout*, and dense output layer).
- **Minimum training**: Freezing all hidden layers, only training the upper layers (*dropout*, and dense output layer).

The checkpoint provided with Inception-Resnet-v2 was used as the pre-training.

- The checkpoint contains weights of the pre-training using over 1.2 million images and more than 1000 categories.

Results

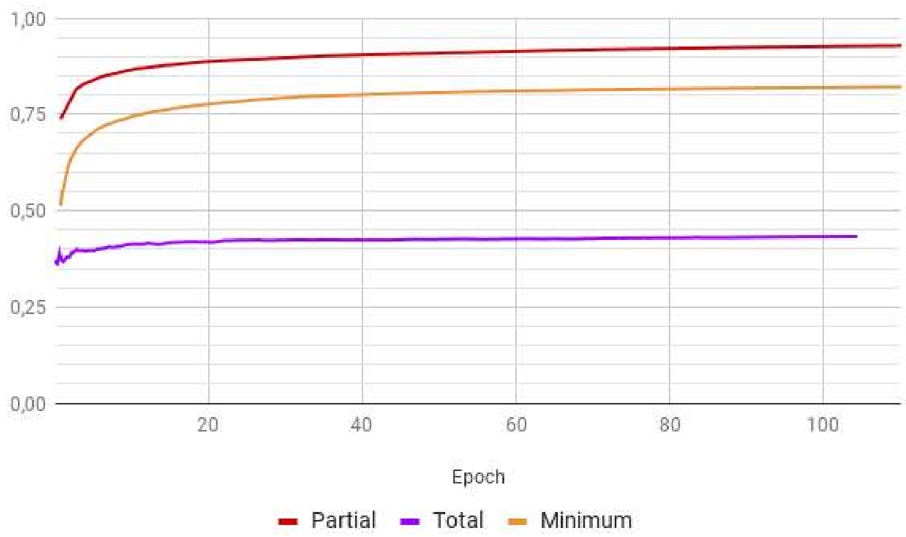

Figure: Accuracy per epoch.

Table: Accuracy comparison by each type of training.

	Total	Partial	Minimum
Epochs	105	975	434
Training	0.4328	0,9743	0,8363
Evaluation	0.3846	0,8956	0,8297

Conclusion

An Inception-Resnet-v2 model was successfully trained to identify ripe cocoa pods with an accuracy of 90%.

Improvements can be made by expanding the dataset:

- Wider variety of angles and settings.
- More examples of differing stages of development, especially between ripe/unripe.
- More powerful hardware would drastically reduce training times.