Improving Manual Concrete Building Facade Inspection with Machine Learning

IE3100R Systems Design Project Group 2 (AY 2018/2019) Department of Industrial Systems Engineering & Management
Group Members: Lim Yan Hwee Jonathan, Yang Jia Long, Yeo Hwee Cheng Clara, Yuan Jiahui, Zheng Yurong
SDP Professor: Liu Yang | Industrial Supervisor: Cindy Chiow

Company Overview

Airbus Aerial is a commercial drone startup under Airbus which leverages on existing aerospace technology to provide imagery services across applications such as insurance, agriculture, building inspections etc.

Project Overview

Project Motivation

Problems with Current Manual Building Facade Inspection process

Time-Consuming | Costly | Unsafe | Subjective

As such, drones are increasingly deployed to speed up inspection processes.
Machine Learning is used to process image data, to standardise defect identification process.

Key Objectives

To evaluate the usage of CNN machine learning models in facade defects inspection

Skillsets Applied

Statistical data analysis
Machine Learning using Python programming

Methodology

Data Collection → Base CNN Architecture → Hyperparameter Optimisation → Accuracy Analysis → Further Improvement (MobileNet CNN) → Recommendations & Conclusion

Data Collection

Scraping

To build database

Slicing

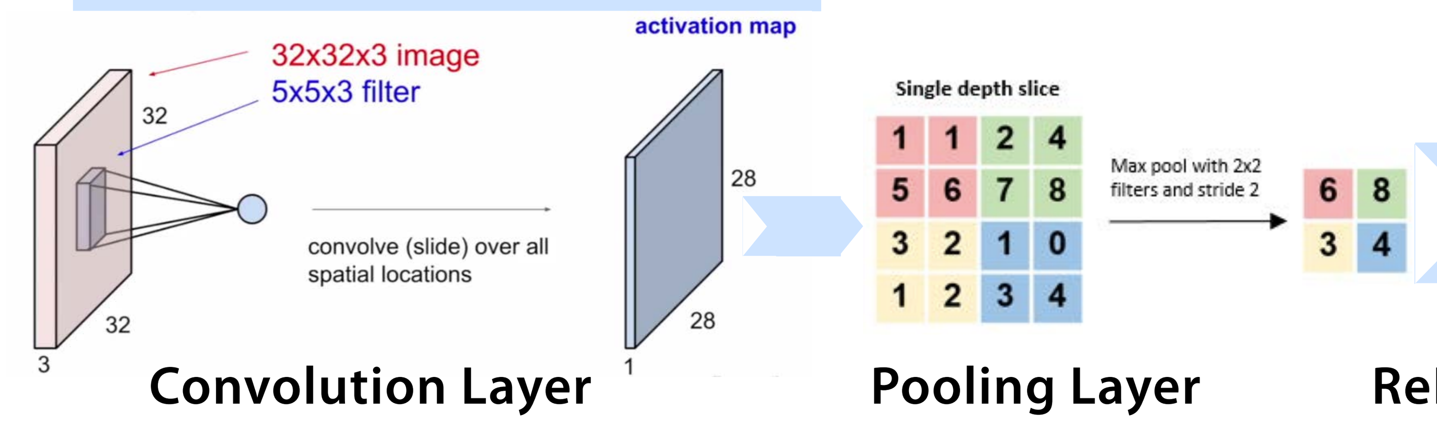
To increase images in database

Rotation

To increase robustness of training data set

(Before) (After)

Explanation of CNN Layers



Activation Function → Fully Connected Layer → Softmax Function

Convolution Layer

Filter is slid over the input image to calculate dot product output across which enables the model to learn specific features

Pooling Layer

Max pooling reduces spatial size and thus the number of parameters and computation

ReLU Function

Performs non-linear transformations

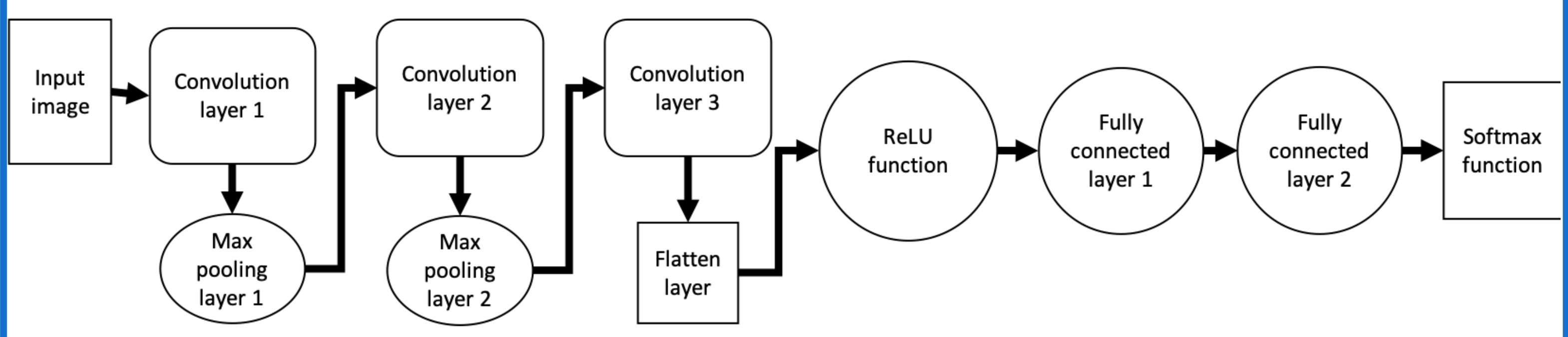
FC Layer

Sum weighting of features in previous layers

Softmax Function

Outputs a probability distribution of the input image being in each class

Base CNN Architecture

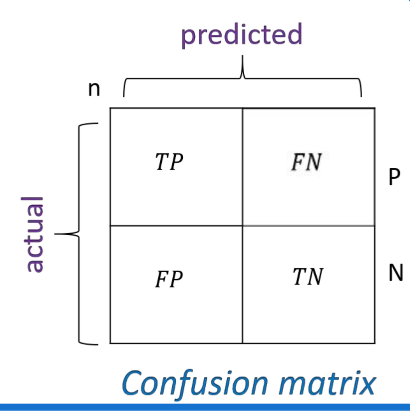


Input layer is customized to take in images of dimensions 224*224*3, which represents width, height and depth

Hyperparameter Optimization

Varying 3 different hyperparameters for 5 chosen intervals to find out the combination that leads to the highest accuracy amidst <u>125</u> configurations

Epoch (EP)	50, 75, 100, 125, 150
Batch Size (BS)	16, 32, 64, 128, 256
Learning Rate (LR)	0.00001, 0.00005, 0.0001, 0.0005, 0.001



Confusion matrix

Performance measurement for machine learning classification algorithm

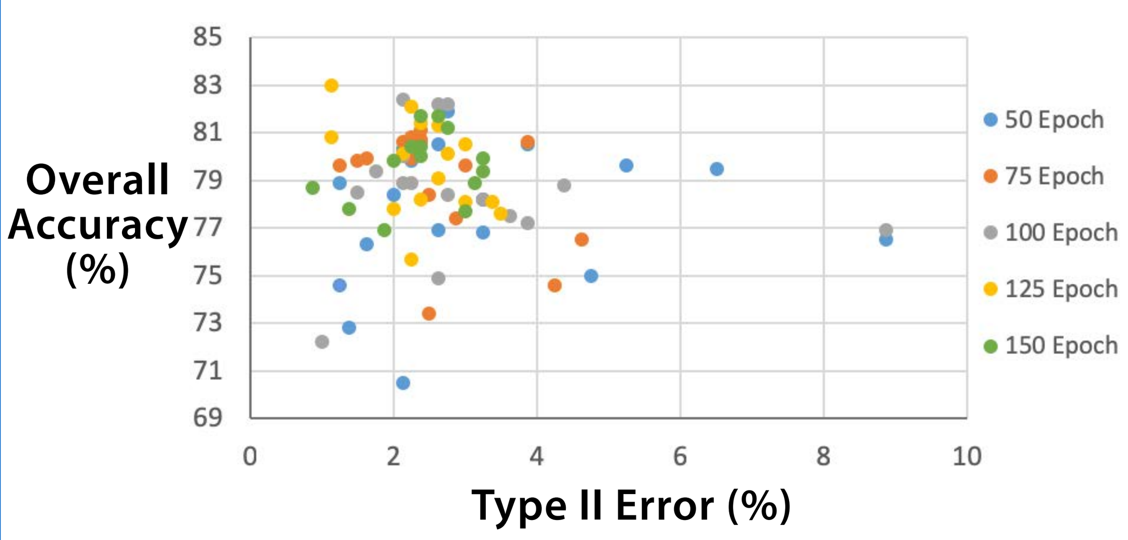
$$\text{Accuracy} = \frac{TP + TN}{TN + FP + FN + TP}$$

Type II Error

False Negatives(FN): the images for which the algorithm predicted no defects but the images have defects

When defects are not detected, this poses a safety risk

Accuracy Analysis



Type II error is another metric that should be considered when deciding the best combination

The best accuracy 83% belongs to the parameter combination of **<u>125 EP, 32 BS and 0.0001 LR</u>**.

A better combination can be further finetuned to achieve better accuracy by using smaller intervals for each hyperparameter.

Benchmarking

1

Chaiyasarn et al.	Our Base CNN Archi.
Inspect heritage buildings	Inspect building facades
2 classes	5 classes
Test accuracy 67.5%	Test accuracy 83%

2

Cha et al.	Our Base CNN Archi.
Trained with 40,000 images	Trained with 3,000 images
2 classes	5 classes
Test accuracy 97%	Test accuracy 83%

Further Improvements

MobileNetV2

Using **<u>depth wise separable convolutions</u>** which replace traditional convolutions - reduce computation and parameters.

1 <u>Model with Transfer Learning</u>

Transfer Learning: Where pre-trained weights from Imagenet is used. Hence model does not have to learn from scratch

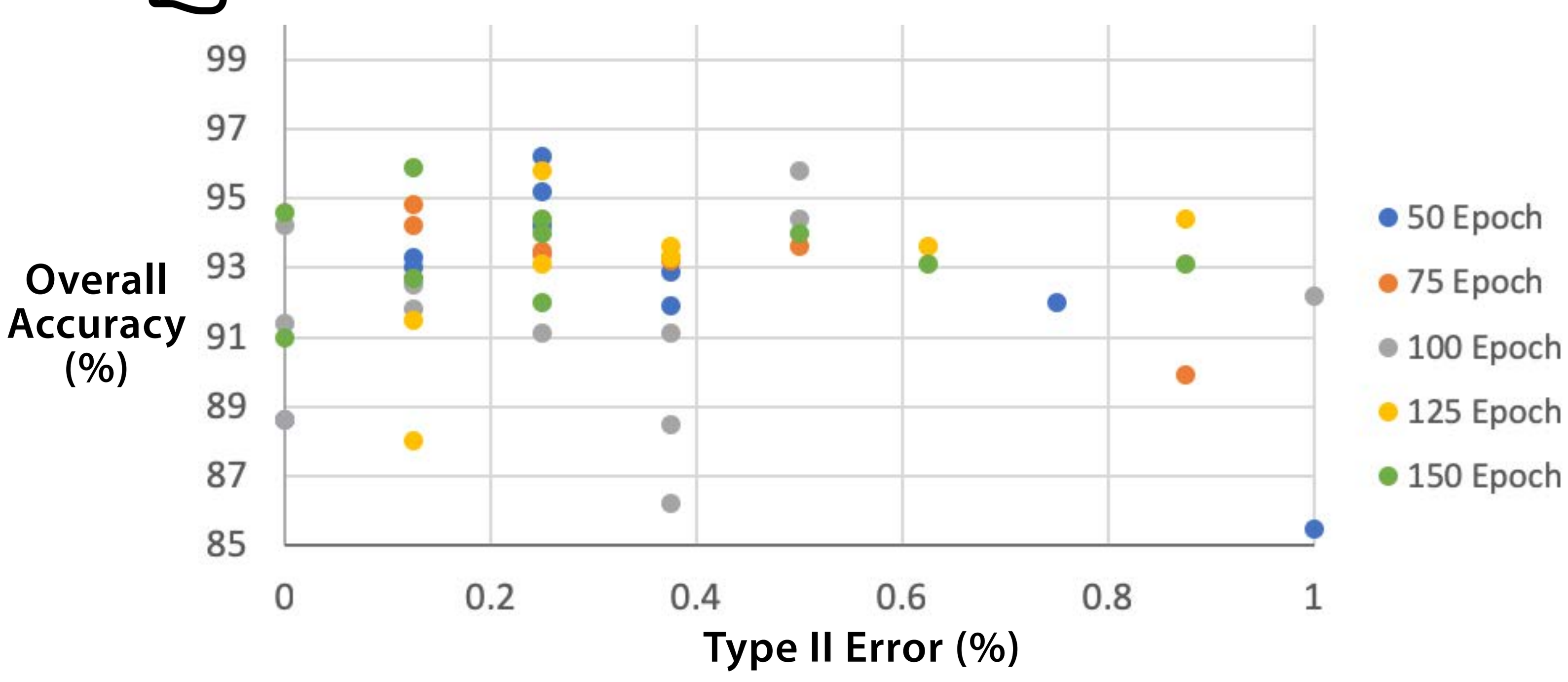
Accuracy:
38.7% to 63%
Type II error:
11.8% to 65.8%

Earlier layers are untrainable, which are crucial for defect detection

2 <u>Model w/o Transfer Learning</u>

Best performing accuracy of 94.6% with 0% Type II error

Hyperparameter combination of 150 EP, 16 BS and 0.0001 LR



Limitations

- Limited capability of CPU
  - Unable to run 128, 256 BS on MobileNetV2
- Lack of representative data for all classes
  - Limited to 1000 images per class
- Long training time required for MobileNetV2

Future Direction & Conclusion

- Use actual image data captured by Airbus drones
- Identify multiple defects in a single image
- Overall accuracy and Type II errors should be considered as performance indicators for hyperparameter optimization
- MobileNetV2 is preferred with its higher overall accuracy