

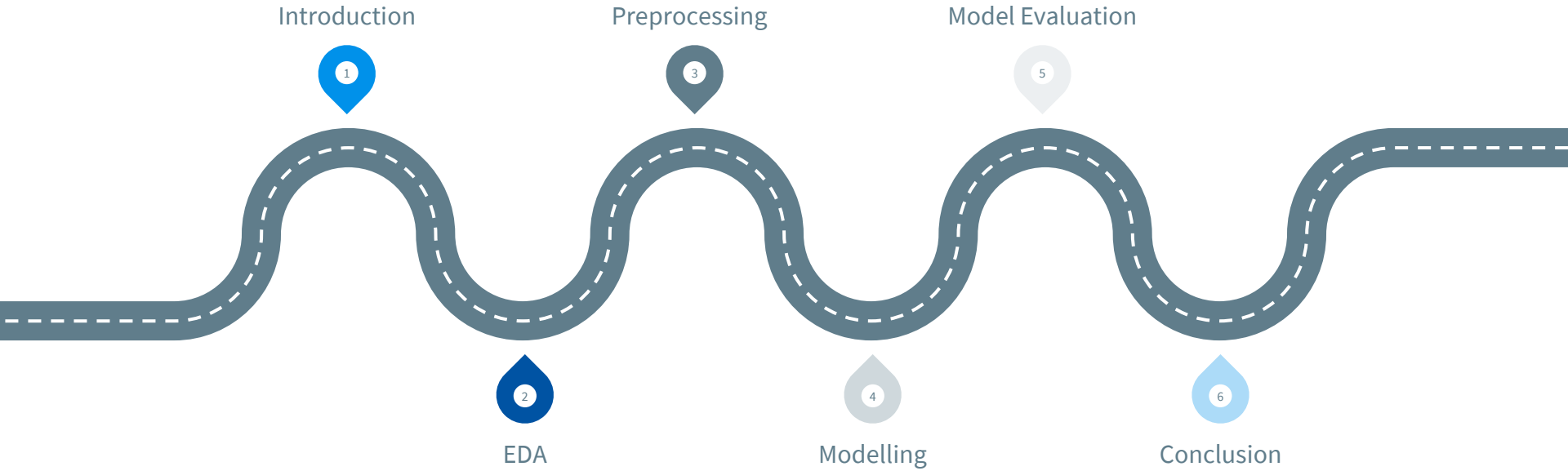
A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines. Some nodes are highlighted with blue circles, and others with blue dots. The lines are thin and grey, creating a mesh-like structure.

# Distracted Driver Detection

**Rifqi Alkhatib**  
General Assembly DSI 20

A decorative network diagram in the bottom-right corner, similar to the one in the top-left, featuring a complex web of interconnected nodes and lines. Some nodes are highlighted with blue circles, and others with blue dots. The lines are thin and grey, creating a mesh-like structure.

# Agenda



A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines. The nodes are represented by small circles, some of which are larger and have concentric circles, suggesting a hierarchical or multi-layered structure. The lines are thin and gray, connecting the nodes in a non-linear fashion.

# Introduction

A decorative network diagram in the bottom-right corner, similar to the one in the top-left. It shows a cluster of nodes connected by lines, with some nodes being larger and having concentric circles. The overall style is minimalist and technical.

## Background

- ◎ High road fatality rates in Singapore
- ◎ Distracted driving a major cause of road accidents
- ◎ >80% of drivers admit to using phone while driving
- ◎ >90% think that it is unsafe



## Problem Statement

- ◎ Explore use of dashboard cameras to improve these statistics
- ◎ Create an image classification model that can detect the distraction state of the driver



# EDA

## Data Source – Kaggle Competition

- ◎ Images of drivers in various states of distraction
- ◎ 10 Classes
- ◎ Images curated from controlled experiment
- ◎ Train and test data split on the drivers\*
  - > 22,000 Train images
  - > 79,000 Test images
- ◎ Evaluated on multi-class log loss

# Classes

c0: Safe Driving



c1: Texting - Right



c2: Talking on the Phone - Right



c3: Texting - Left



c4: Talking on the Phone - Left



c5: Operating the Radio



c6: Drinking



c7: Reaching Behind



c8: Hair and Makeup



c9: Talking to Passenger





## Insights from Data

- ◎ Drivers' body position the most distinguishing feature
- ◎ Features consistent across classes
  - Type of car
  - Lighting
  - Quality of image
- ◎ Features inconsistent within same class
  - Driver features
  - Camera angle

A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines. The nodes are represented by small circles, some of which are larger and have concentric circles, suggesting different levels of connectivity or importance. The lines are thin and gray, creating a mesh-like structure.

# Preprocessing

## Distinct Driver Issue

- ◎ Train and test data split on the drivers\*
- ◎ Model must be able to generalize to new drivers
- ◎ Many similar images
- ◎ Might end up overfitting to drivers' features

Class: c0, File: img\_208.jpg

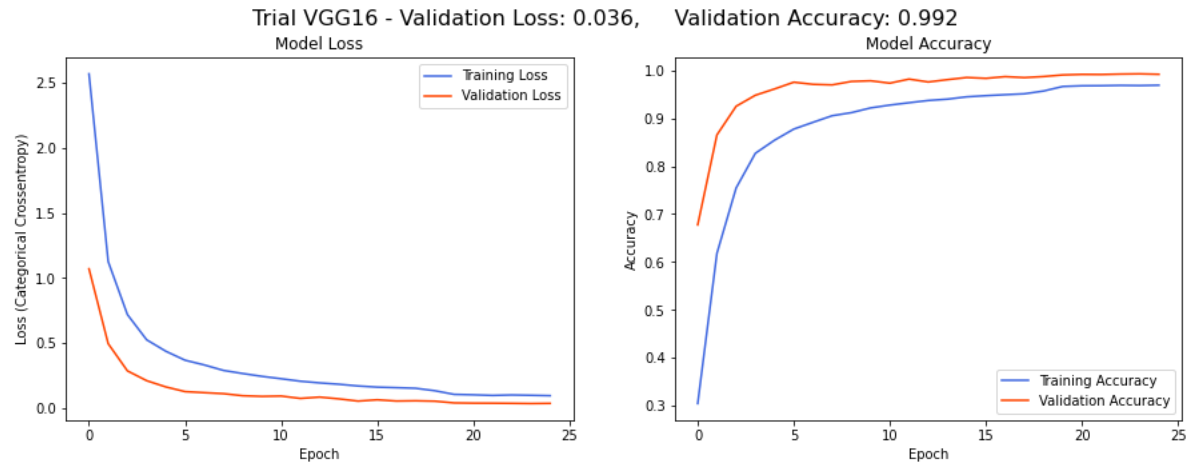


Class: c0, File: img\_231.jpg



## Trial Model

### Transfer learning from pre-trained CNN model: VGG16



### Kaggle score – 1.78549

## Split by Driver

- ◎ Random train-validation split leads to poor generalization
- ◎ Split by unique driver instead
  - 26 unique drivers in train data
  - 19 – 7 split
- ◎ Model should train on non-driver related features

A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines. The nodes are represented by small circles, some of which are solid grey and others are hollow with a grey outline. The lines connecting them are thin and grey, creating a dense, organic structure that tapers off towards the right.

# Modelling

## Methodology

- ◎ Transfer Learning
  - Build classifier on pre-trained CNN Models
    - ◎ VGG16
    - ◎ EfficientNetB4
- ◎ Fine-tuning
  - 2 rounds
- ◎ Optimised on categorical cross-entropy
- ◎ 25 Initial epochs + 25 Fine-tuning epochs per round
- ◎ Callbacks
  - Checkpoint, EarlyStopping, ReduceLROnPlateau

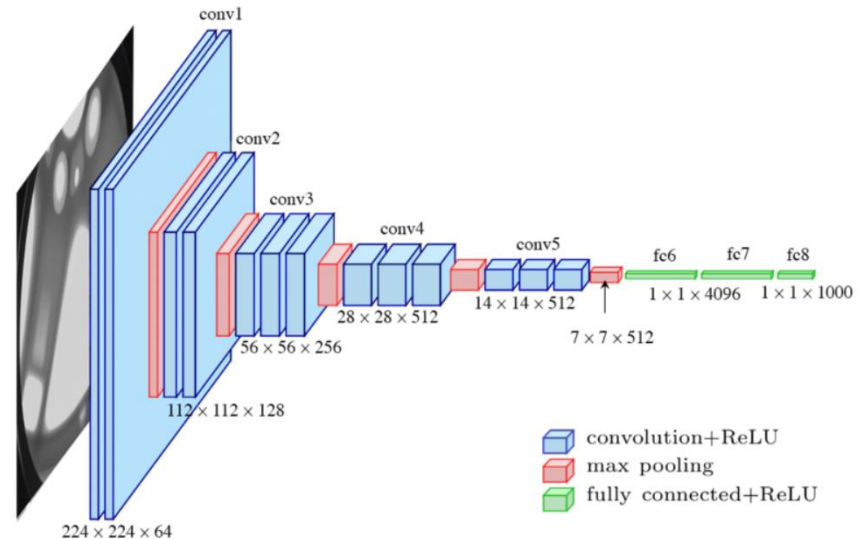
## Baseline Model

- ◎ Predict all classes to have same probability
- ◎ Baseline score: **2.303**



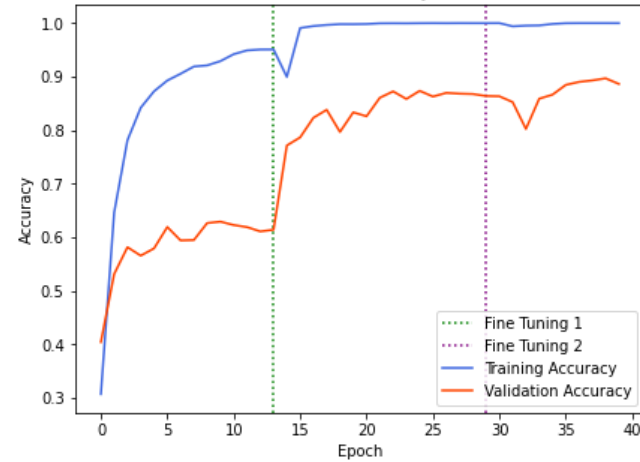
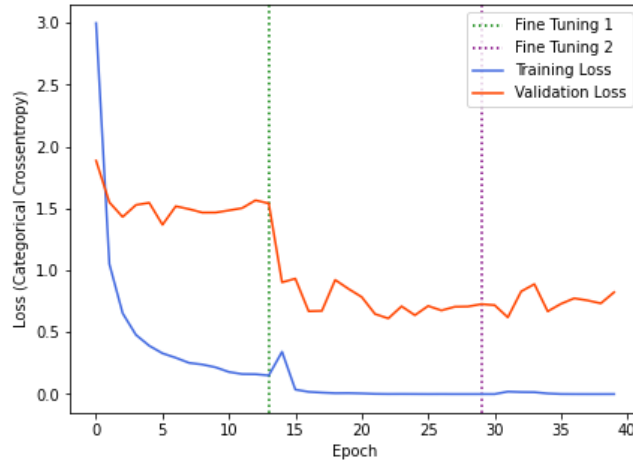
# VGG16

- Trained on ImageNet
  - 90.1% Top-5 Acc
- 5 Convolutional blocks
- 3 Fully connected layers
- Simple architecture
- Large size



# VGG16

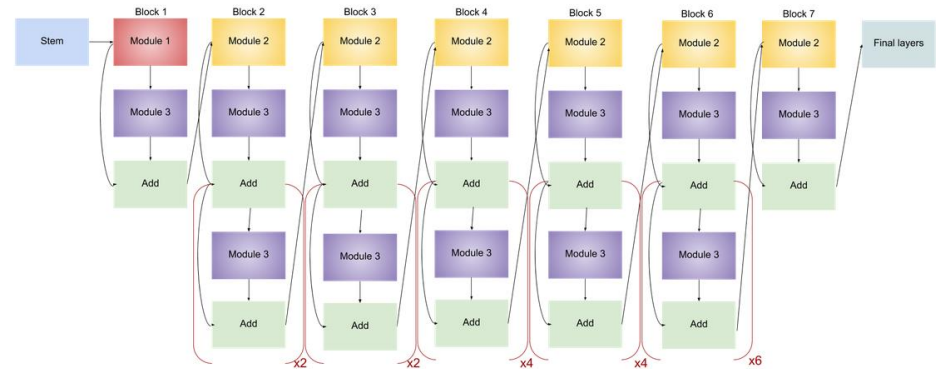
Fine-Tuned VGG16 - Validation Loss: 0.620, Validation Accuracy: 0.853



🕒 Kaggle Score: **0.68395**

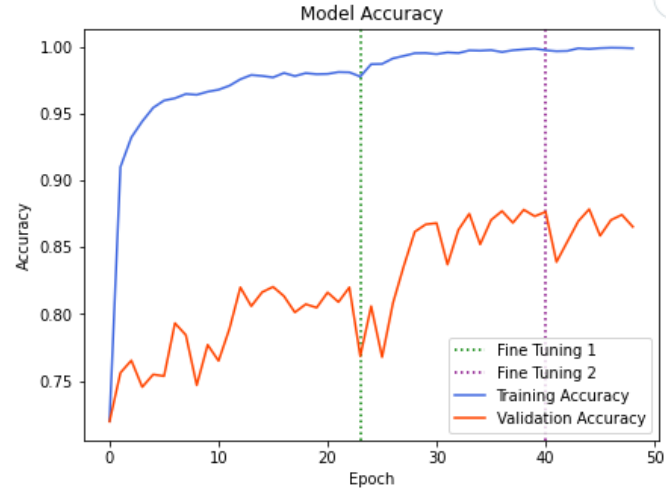
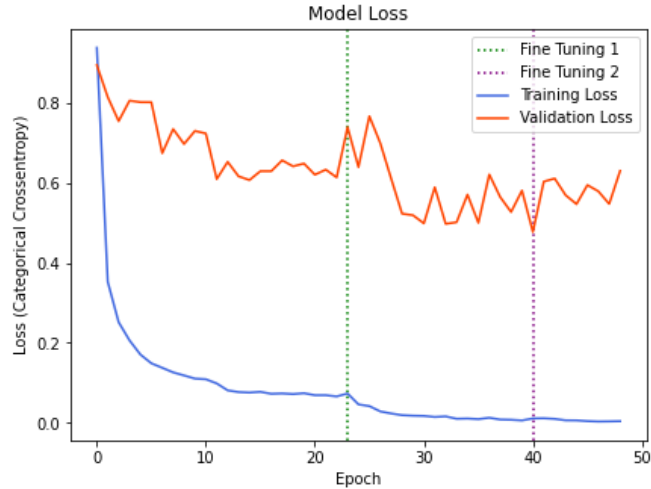
# EfficientNetB4

- ⦿ Trained on ImageNet
- ⦿ Focused on efficiency of parameters
- ⦿ Repeated modules linked in blocks
- ⦿ 8 different architectures
- ⦿ B4 provides good balance



# EfficientNetB4

Fine-Tuned EfficientNetB4 - Validation Loss: 0.478, Validation Accuracy: 0.877



🕒 Kaggle Score: **0.59852**

## Selection of Best Model

Model	Data	Fine-Tuning	Test Score (Kaggle)	Val Loss	Val Accuracy	Initial Epochs	Fine-Tuning Epochs	Total Epochs
VGG16	Randomly Shuffled	No	1.78549	0.049	0.987	25	-	25
VGG16	Split by Driver	Yes	0.68395	0.62	0.853	13	17 + 8	38
EfficientNetB4	Split by Driver	Yes	0.59852	0.478	0.877	24	16 + 9	49

- ⦿ EfficientNetB4 performs best
- ⦿ Splitting by driver reduced overfitting

## Final Kaggle Submission

- ◎ Ensemble method
- ◎ Cross-validate best model over 4 folds
  - Data is not randomly shuffled
  - Expose the model to maximum amount of information
- ◎ Fine-tuned EfficientNetB4
- ◎ Final Score: **0.30096**
  - Improvement from single fold (0.598)

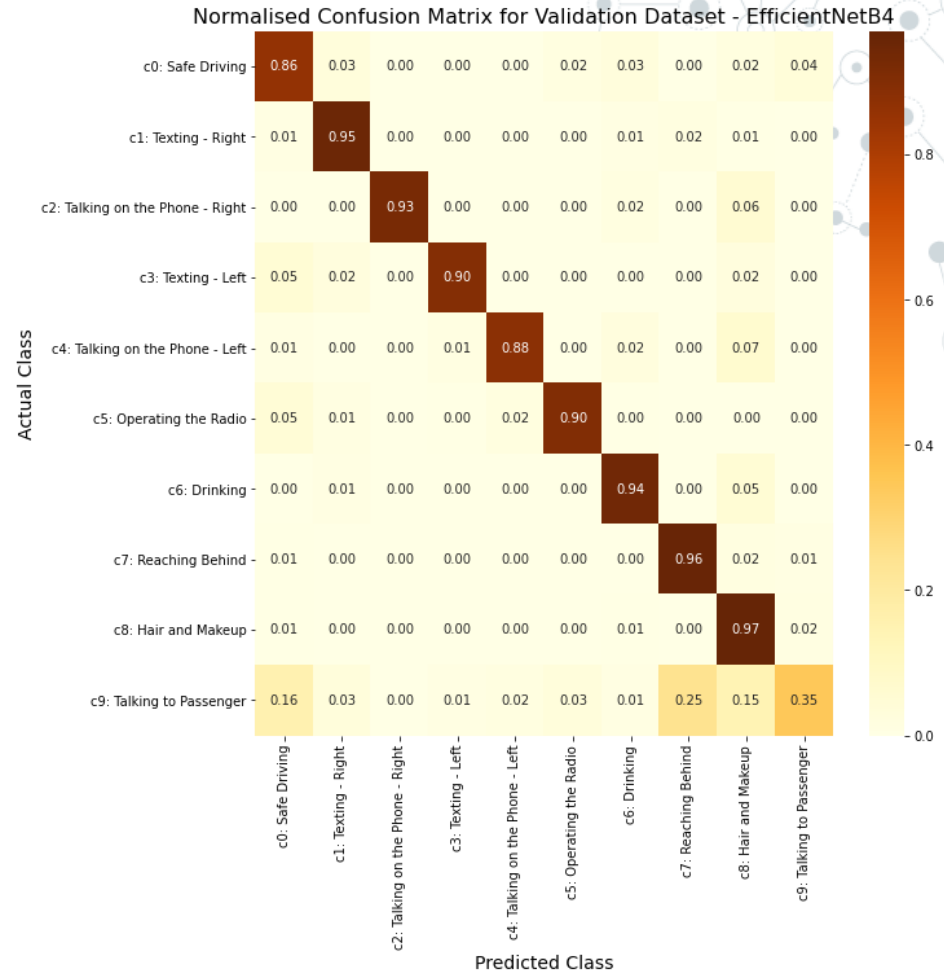


# Model Evaluation



# Confusion Matrix

- ⦿ Generally performs well
- ⦿ Class c9: Talking to Passenger





## Misclassifications

- ⊙ Sources of ambiguity for each class
- ⊙ Particularly noticeable for c9
- ⊙ Class labels not always consistent

Actual: c0, Predicted: c9



Actual: c1, Predicted: c7



Actual: c2, Predicted: c8



Actual: c3, Predicted: c1



Actual: c4, Predicted: c8



Actual: c5, Predicted: c0



Actual: c6, Predicted: c8



Actual: c7, Predicted: c8



Actual: c8, Predicted: c0



Actual: c9, Predicted: c7



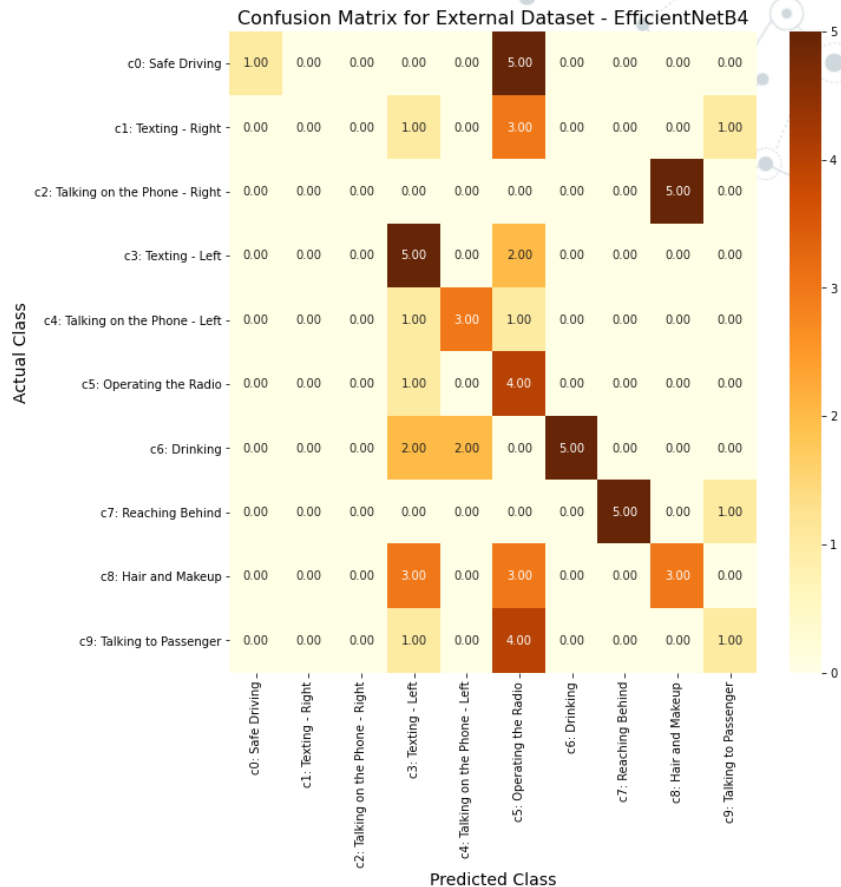
## Test on Unseen Data

- ◎ Images outside the experiment
- ◎ Different lighting, car, location, camera angle



## Test on Unseen Data

- Does not perform well
  - 43% accuracy
- Some classes 100% wrong



A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines. The nodes are represented by small circles, some of which are larger and have concentric circles, suggesting different levels of connectivity or importance. The lines are thin and gray, creating a mesh-like structure.

# Conclusion

## Conclusion

- ◎ Successfully created model that predicted well for test set
- ◎ Not robust enough to be useful in real life
- ◎ Shows that detecting distraction with dashboard cameras is feasible
- ◎ Can be applied to improve road safety
  - Useful information for insurance companies
  - Real time feedback to drivers

## Recommendations for Future Development

- ◎ Improve performance on unseen external data
  - Grayscale photos to reduce effect of driver features
  - Image augmentation
  - Gather more varied data
- ◎ Further development
  - Apply model to videos

A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines. The nodes are represented by small circles, some of which are larger and have concentric circles, suggesting different levels or types of connectivity. The lines are thin and gray, creating a mesh-like structure.

# Thank You