Classifying plastic resins codes

GA DSI 38 Capstone Project by Ben

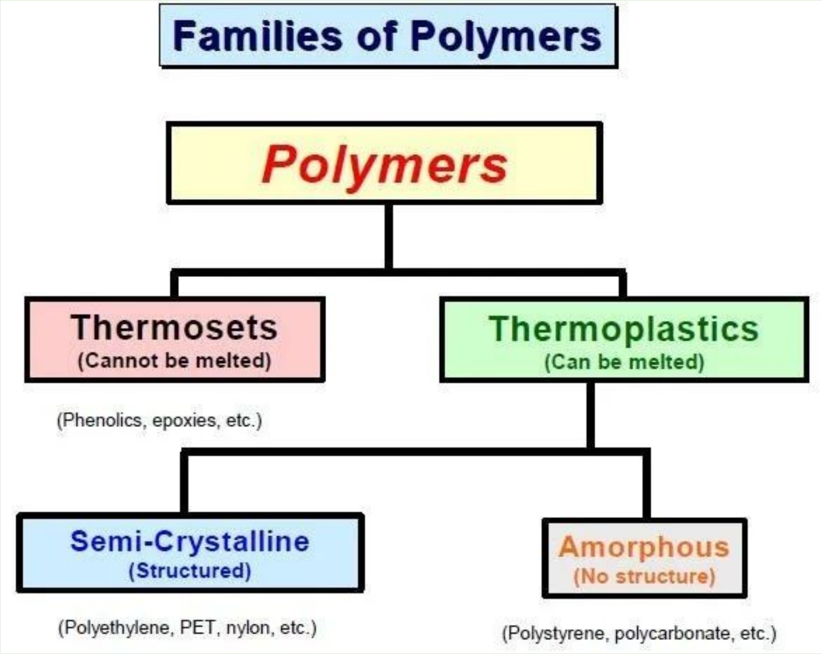
Contents

Background

- Singapore's National Recycling Programme
  - Launched in 2001
  - **Blue recycling bins** and recycling collection services
  - Encourage household recycling - 3Rs
- Sustainable Singapore Blueprint
  - **30%** domestic recycling rate **by 2030**
- Total amount of plastic waste generated for 2022
  - 1 millions tones
  - Only about 6% is recycled..

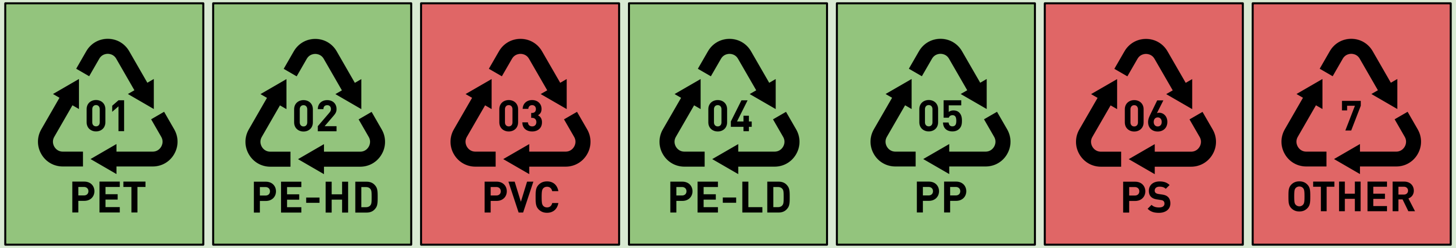
Background

- Different plastic resins have **different properties**
- Due to economy of scale, actual recycling process does not happen in Singapore
- By international regulation (Basel Convention)
  - limit of **< 0.5% contamination**

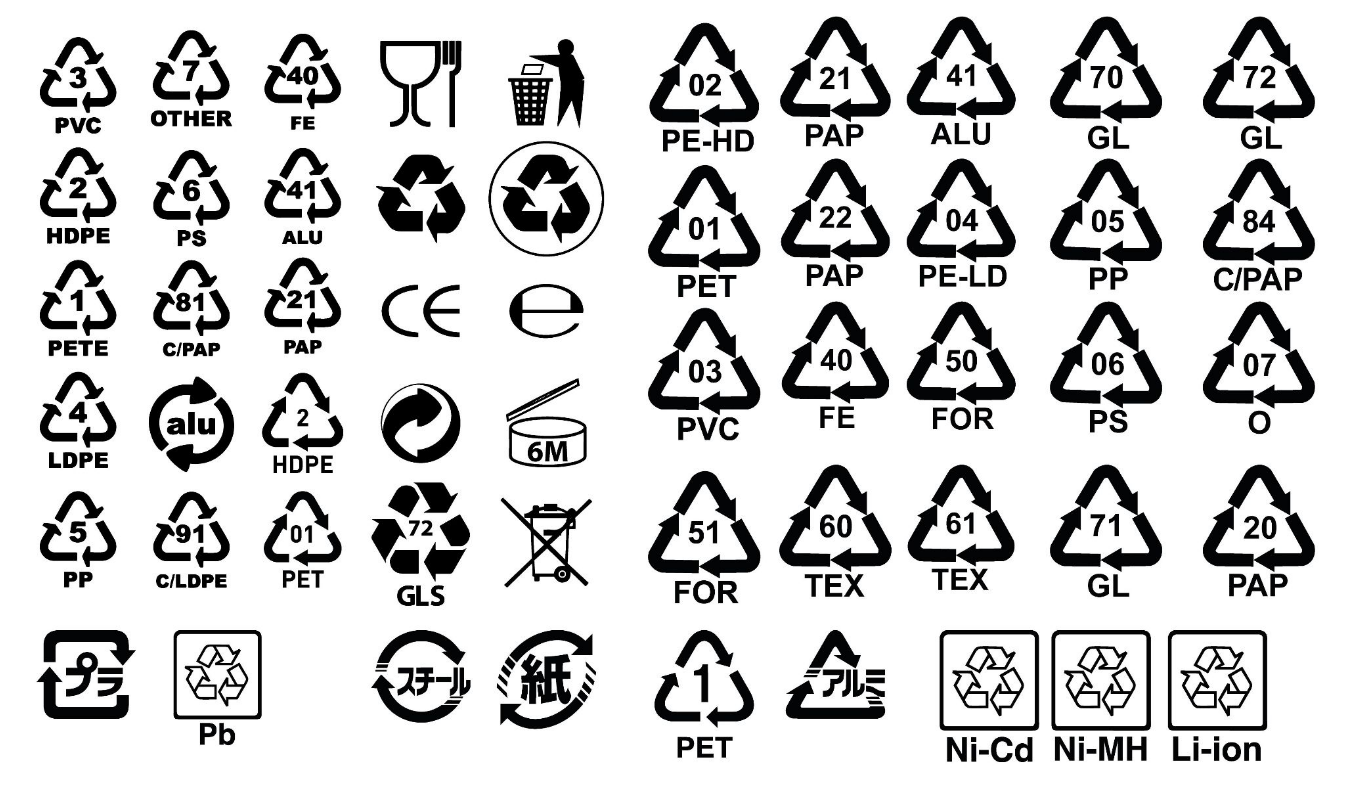


Background

Background



Background


3
PVC
7
OTHER
40
FE
2
HDPE
6
PS
41
ALU
1
PETE
81
C/PAP
21
PAP
CE
e
4
LDPE
alu
2
HDPE
6M
5
PP
91
C/LDPE
01
PET
72
GLS
プラ
Pb
スチール
紙
02
PE-HD
21
PAP
41
ALU
70
GL
72
GL
01
PET
22
PAP
04
PE-LD
05
PP
84
C/PAP
03
PVC
40
FE
50
FOR
06
PS
07
O
51
FOR
60
TEX
61
TEX
71
GL
20
PAP
1
PET
アルミ
Ni-Cd
Ni-MH
Li-ion

Problem statement

Context

- Some **knowledge necessary** for individuals **to recycle effectively**
- To nudge singapore towards **30% recycling rate by 2030**
- NEA is **exploring new ways to encouraging singapore to recycle** on top of existing campaigns and posters
- **NEA Information Technology Division** has been tasked with developing a proof of concept project.

Purpose

- Determine if it there is a feasible way to **assist the general public** in their recycling activities in particular, determining **which plastics are recyclable**

Problem statement

<u>How</u>

- By utilising **computer vision** in a **binary classification model** to classify whether a plastic material is **recyclable or not** by recognising the Resin Identification markings on the plastic material, thereby reducing the confusion and knowledge requirement for the layperson to recycle effectively.

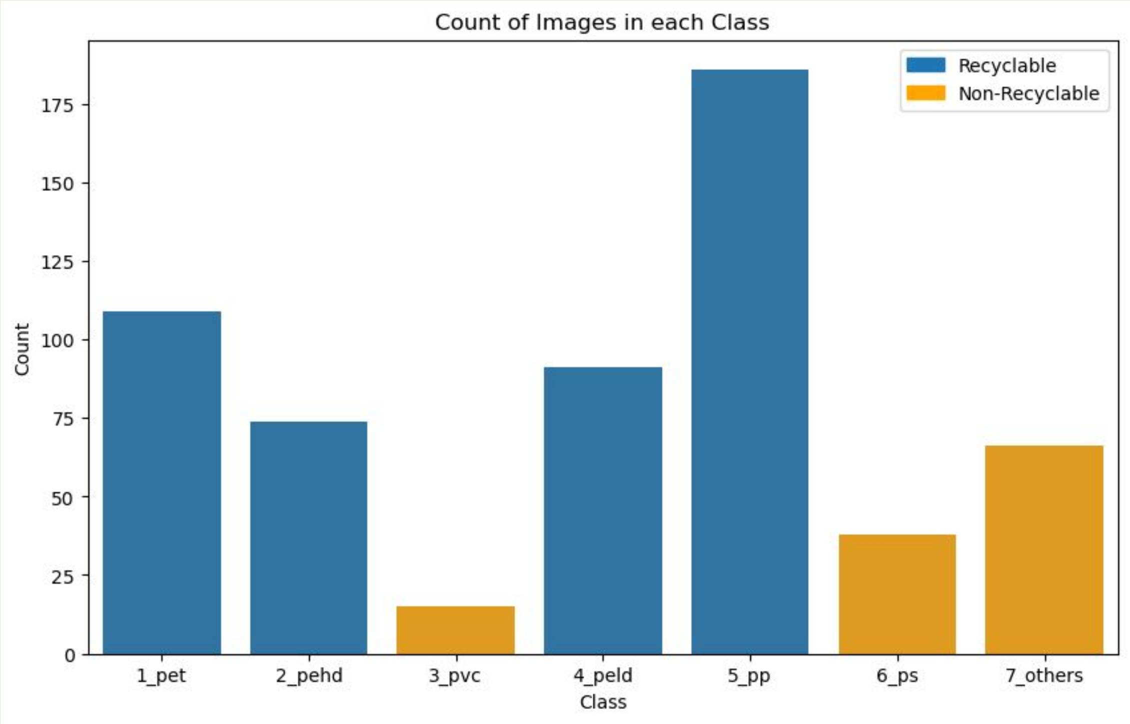
<u>Metric</u>

- Scoring metric used in classification is **F1 score with at least 80%** being used as as a **threshold of success for the project** to be considered feasible for further pursuit.

Data

- Obtained on kaggle & mixed in with personally taken images.
    - **Separated in individual class**
- Visualised within the environment for verification
    - Logos have varying colours, shadings and styles
    - Resolution is 200 x 200 pixels.

Exploratory Analysis

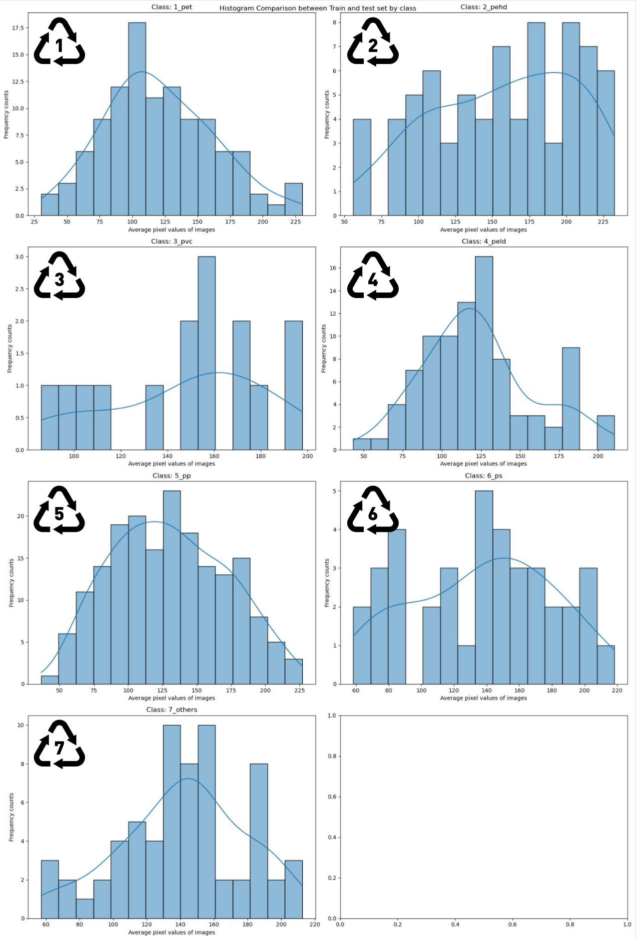


- **Imbalanced** dataset
- Number of images per class fluctuates
- **Less non-recyclables** images in the data set

Exploratory Analysis

Average Pixel count

- Distribution of images within each class
- Some classes (2,3,4, 6,7) are not normally distributed due the the lower image counts



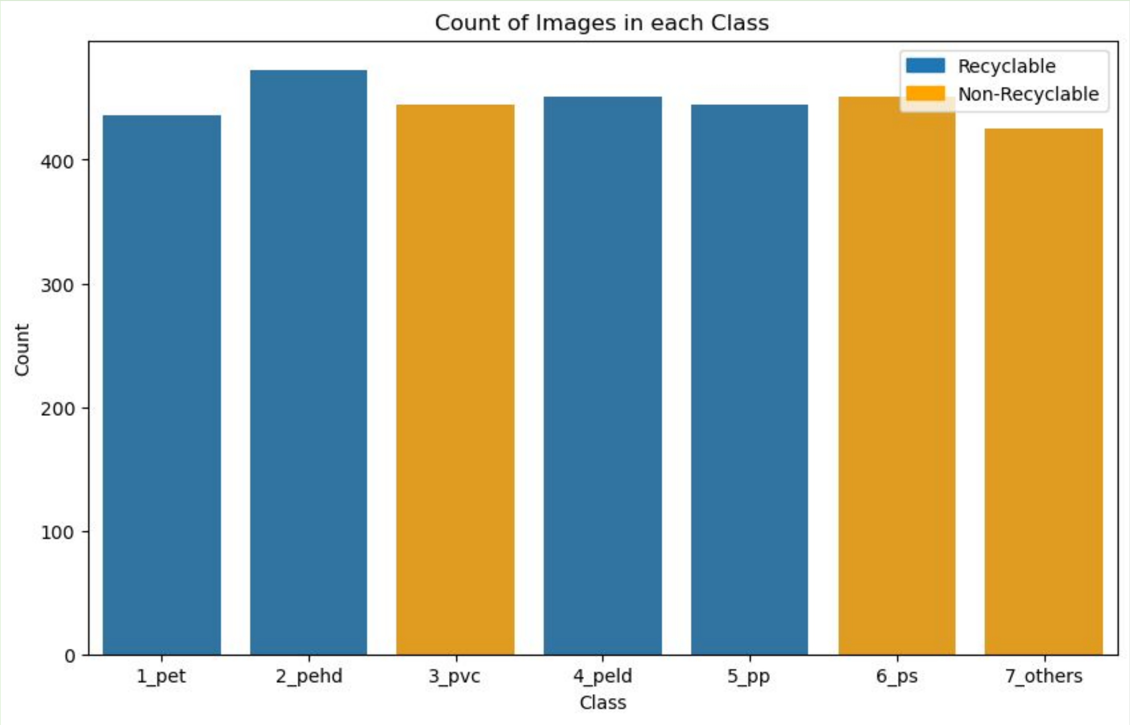
Preprocessing Steps

Data Augmentation

- To **increase the number of images**
- Introduce more variations and help the **model generalise better**
  - Brightness, Contrast
  - Shift, Scale, Rotate
  - **Better reflects conditions** that may occur in life.

Original | Augmented | Augmented

Preprocessing Steps

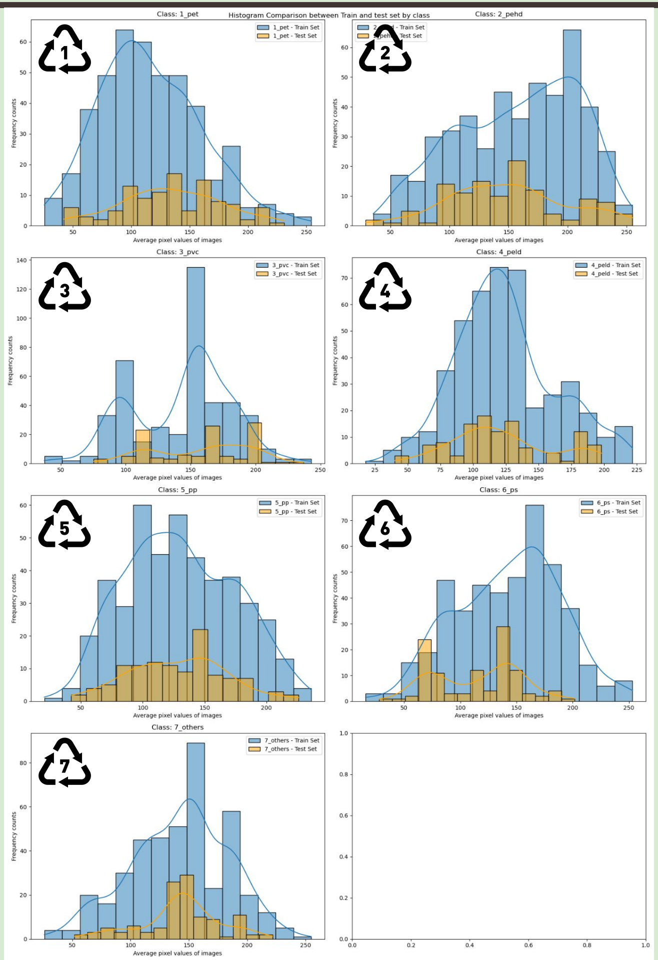


- Data is augmented
- Classes balanced
- Number of images per class similar
- Still less non-recyclables images overall

4 Preprocessing Steps

Average Pixel count

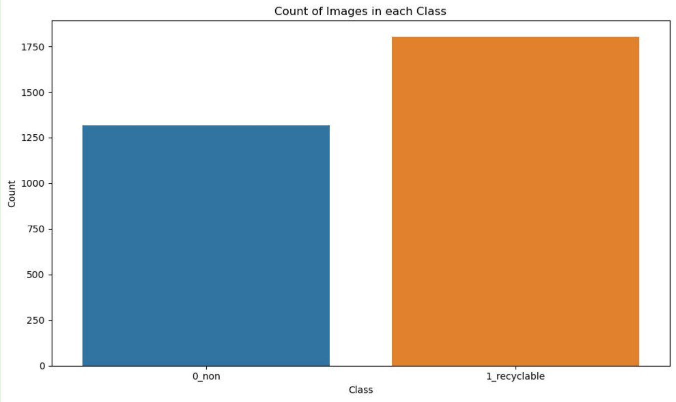
- Augmentation has helped to fill in some gaps
- classes 3, 4 &,6 still show quite a significant difference from a normal distribution

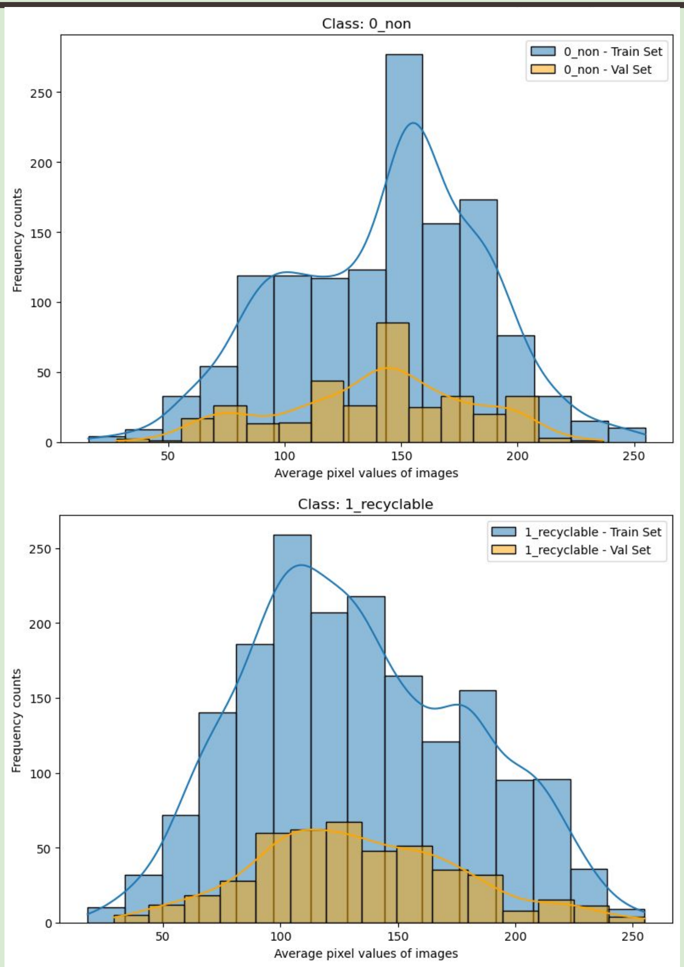


♳ Preprocessing Steps

Collapsing back into binary form

- Data imbalance of lesser non recyclables images.





♵ Modelling

Pretrained models from 3 types of network architecture used

- MobileNet
  - **Computationally efficient** -> Suitable for mobile devices
  - May sacrifice some accuracy
- EfficientNet
  - Offers better accuracy than MobileNet
  - Requires more computational resources than MobileNet
- ResNet
  - Higher accuracy from deeper learning from residuals(errors)
  - Requires significantly more computational resources

Modelling results

No.	Model	Training F1 Score	Validation F1 Score
1	MobileNetV2 + dropout added	82.4%	81.0%
2	MobileNetV3 + dropout added	81.3%	80.9%
3	EfficientNetV2M + dropout added	81.2%	80.2%
4	EfficientNetV2M	82.5%	79.4%
5	ResNet50 + dropout added	82.7%	79.1%

Modelling results

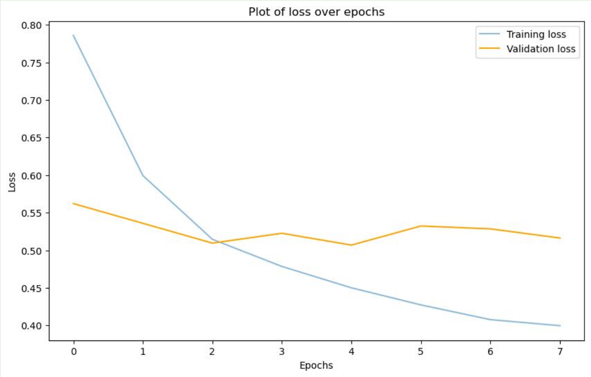
No.	Model	Training F1 Score	Validation F1 Score
1	MobileNetV2 + dropout added	82.4%	81.0%
2	MobileNetV3 + dropout added	81.3%	80.9%
3	EfficientNetV2M + dropout added	81.2%	80.2%
4	EfficientNetV2M	82.5%	79.4%
5	ResNet50 + dropout added	82.7%	79.1%

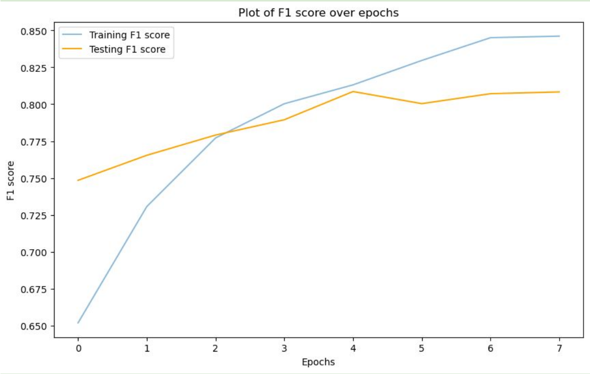
Modelling results

No.	Model	Training F1 Score	Validation F1 Score
1	MobileNetV2 + dropout added	82.4%	81.0%
2	MobileNetV3 + dropout added	81.3%	80.9%
3	EfficientNetV2M + dropout added	81.2%	80.2%
4	EfficientNetV2M	82.5%	79.4%
5	ResNet50 + dropout added	82.7%	79.1%

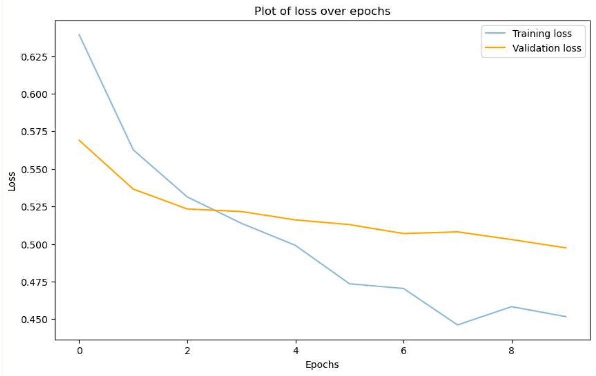
Modelling results

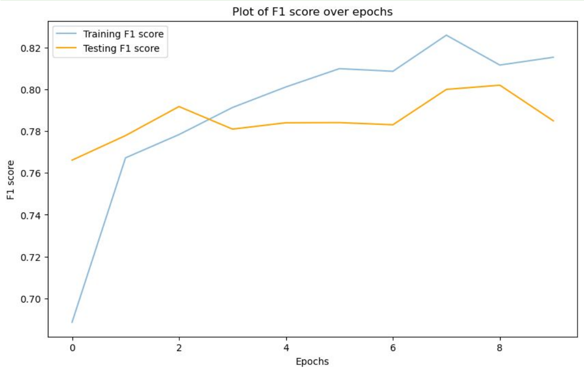
Model 2: Score 80.9%
(MobileNetV3 + dropout added)



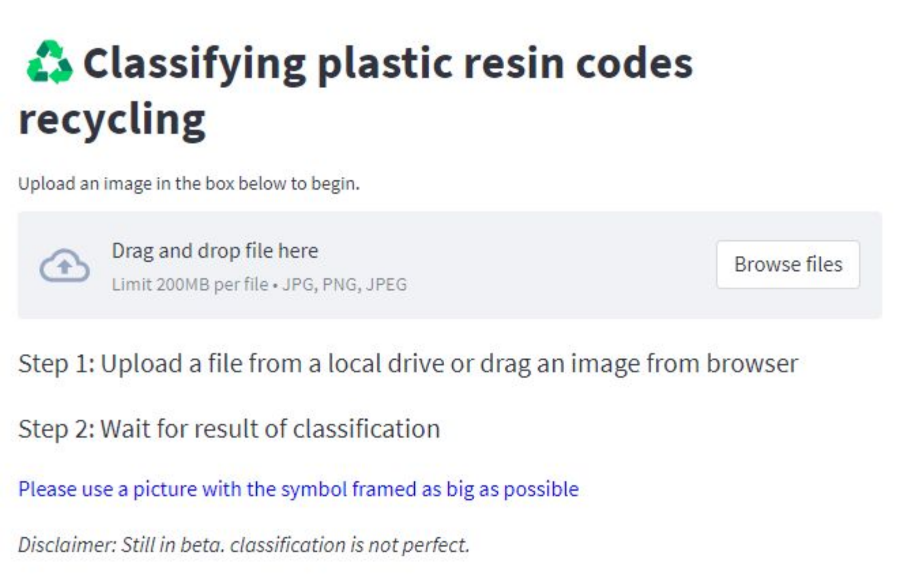


Model 3: Score 80.2%
(EfficientNetV2M + dropout added)





Proof of concept: local deployment



Conclusion & Future works

Conclusion

- Selected model able to classify the “recyclability” fairly well.
- Modelling score of 80% meets the threshold of success
- More images have to be sourced to improve the dataset

Future works

- Multiclass classification,
  - 7 symbols
  - Plastics and all the other symbols
- Localisation of the symbols on images.
  - Facilitate image/video capture
- Explore SSD models for faster processing for multiple items.

Thanks!