# Computer Vision in Convenience Retail

Dennis Francis

IAMS
IAMS
IAMS

# The Images

Scraped from the web

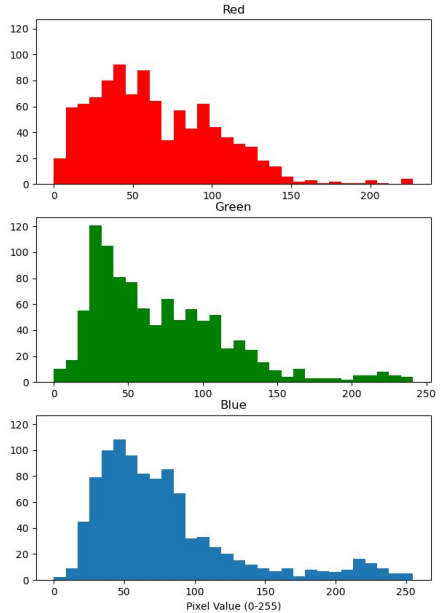
# Labeled Data

ice cream

cigarettes

# EDA

Histograms of RGB



# EDA

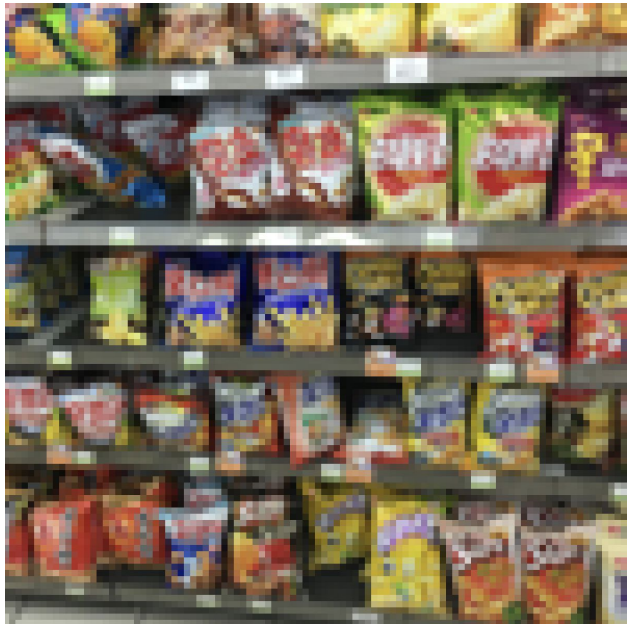
Observe the difference

in image features of

1600 x 1600 vs

100 x 100 image

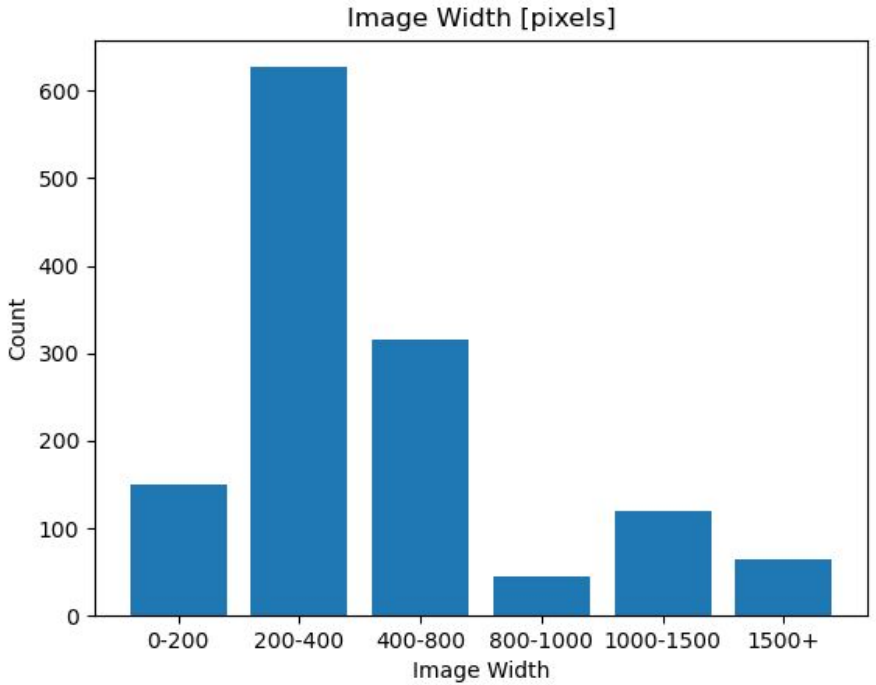
```
tensor([[[255, 210, 213,  ..., 217, 215, 200],
         [252, 205, 213,  ..., 217, 216, 210],
         [232, 203, 213,  ..., 209, 208, 213],
         ...,
         [190, 191, 191,  ..., 181, 181, 181],
         [191, 191, 191,  ..., 181, 181, 181],
         [191, 191, 191,  ..., 181, 180, 180]],

        [[178, 103,  90,  ..., 166, 168, 153],
         [154,  91,  86,  ..., 163, 165, 159],
         [120,  77,  78,  ..., 150, 152, 157],
         ...,
         [193, 194, 194,  ..., 176, 176, 176],
         [194, 194, 194,  ..., 176, 176, 176],
         [194, 194, 194,  ..., 176, 175, 175]],

        [[190, 113,  95,  ...,  61,  54,  35],
         [167, 101,  93,  ...,  63,  58,  50],
         [136,  91,  85,  ...,  58,  57,  62],
         ...,
         [176, 177, 177,  ..., 136, 136, 136],
         [177, 177, 177,  ..., 136, 136, 136],
         [177, 177, 177,  ..., 136, 135, 135]]]
```



# EDA

224 x224 images are the

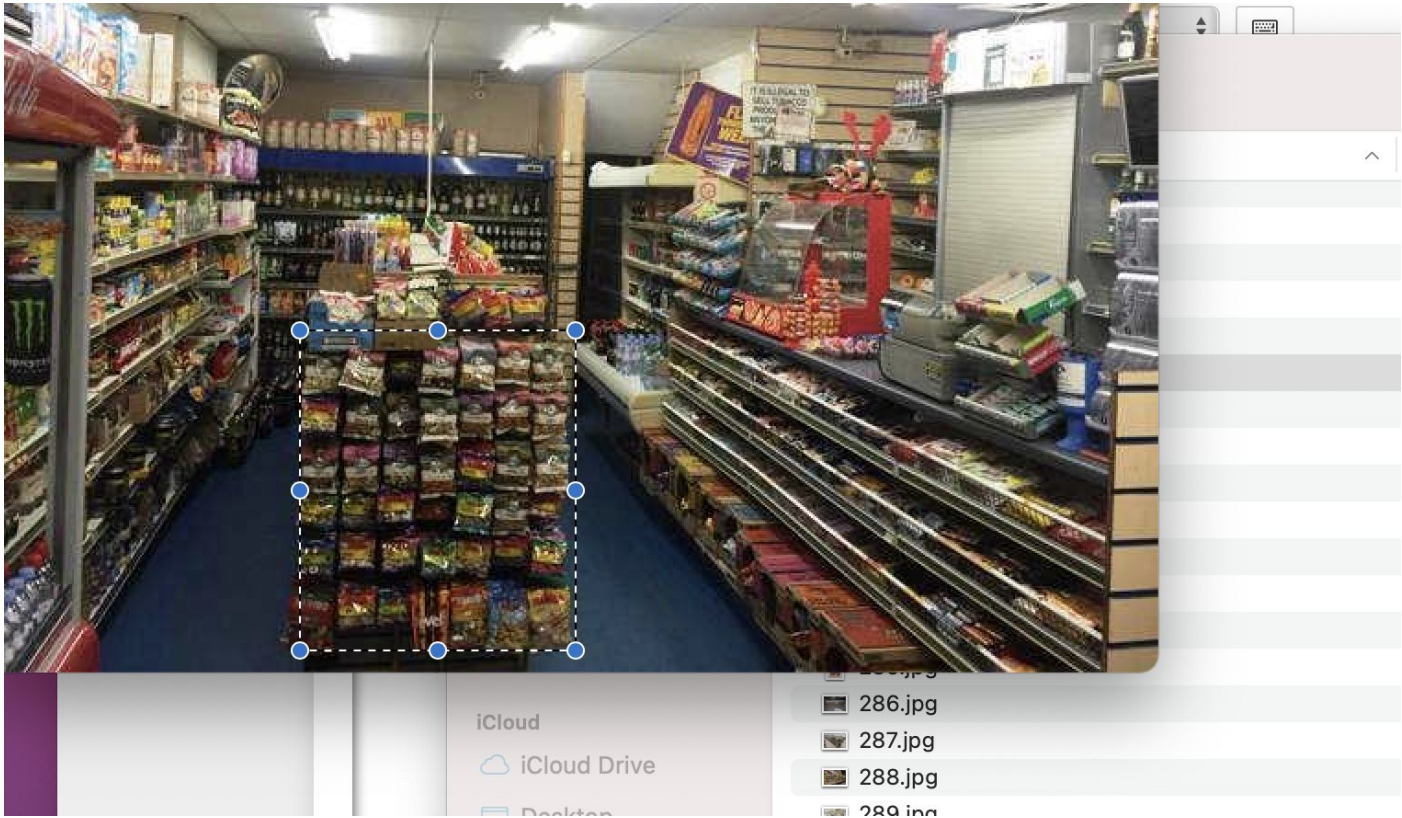
most common



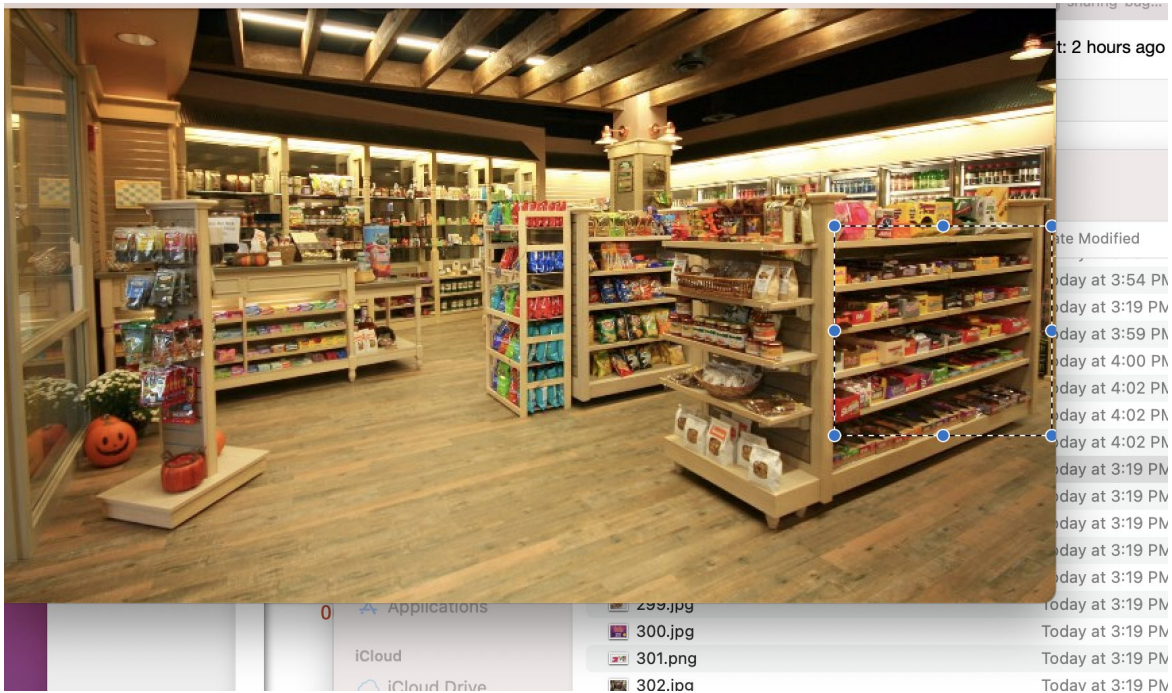
# Cropping an image

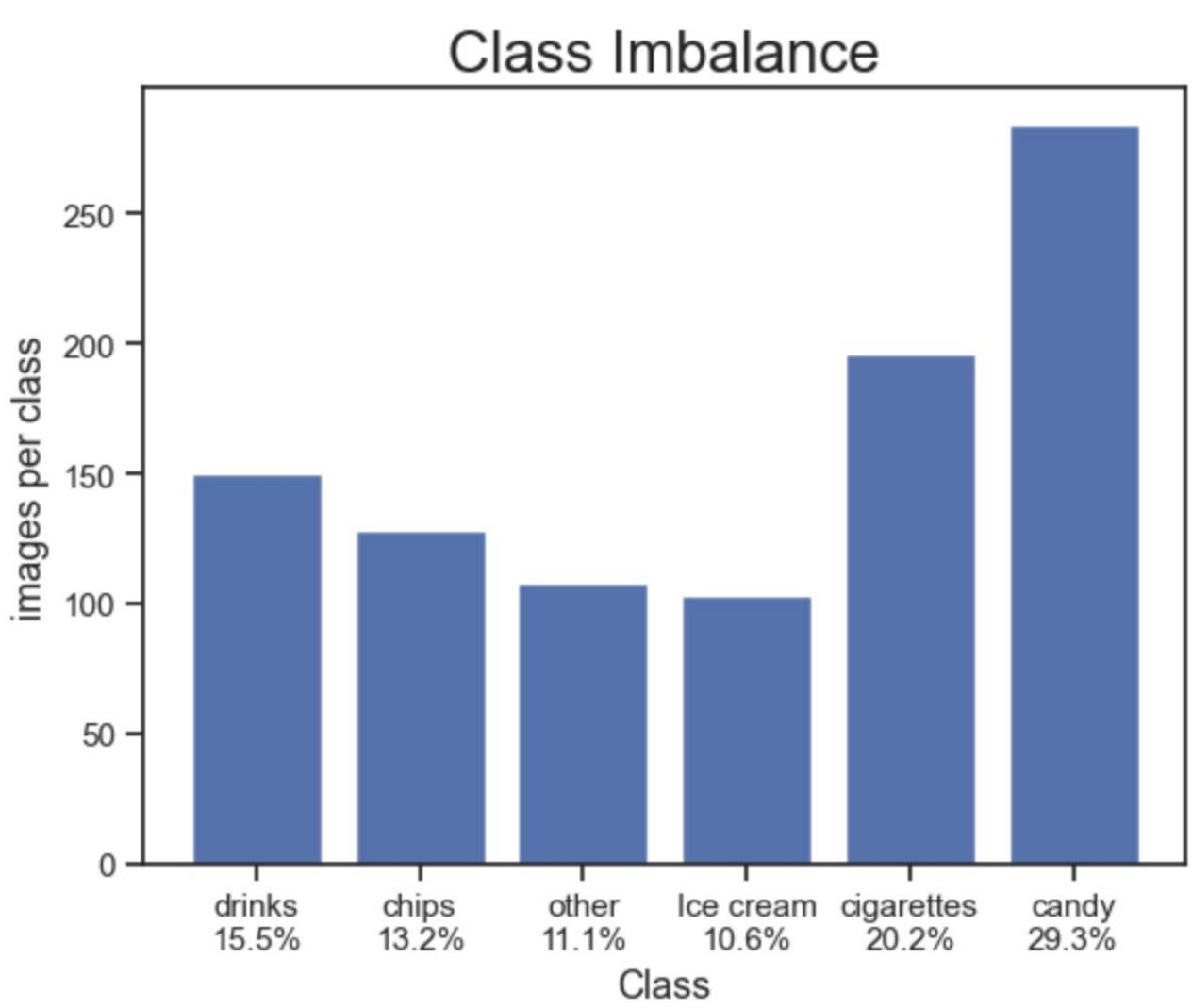
# Cropped Image

# Creating more data

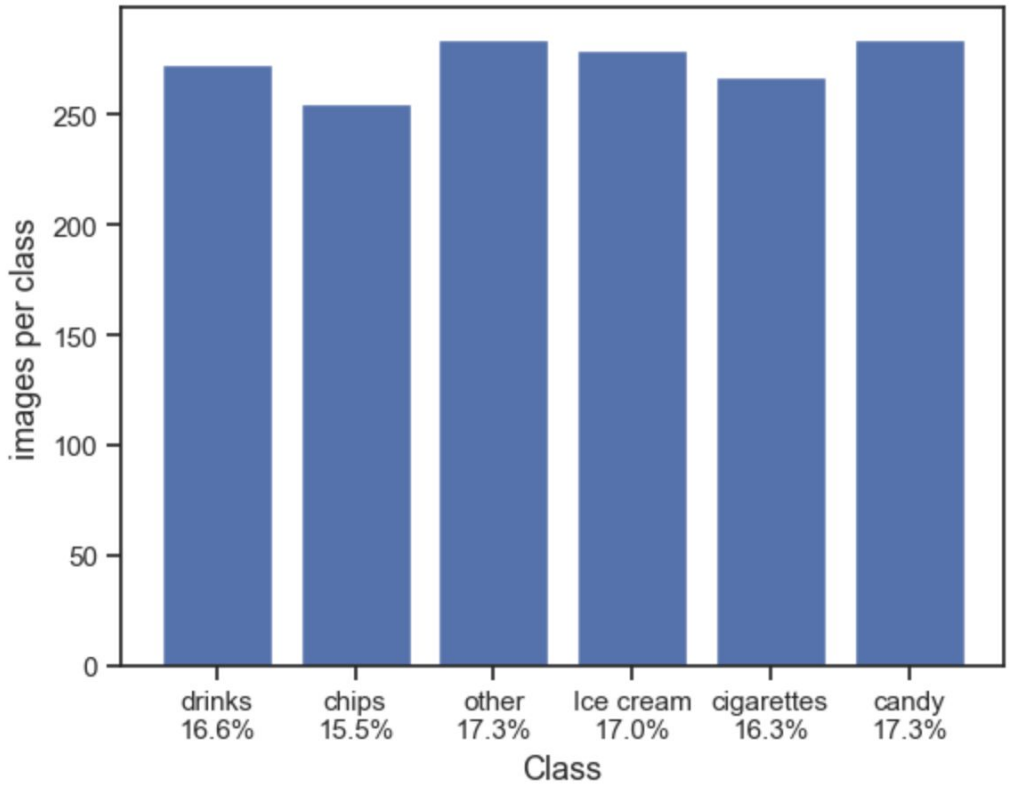


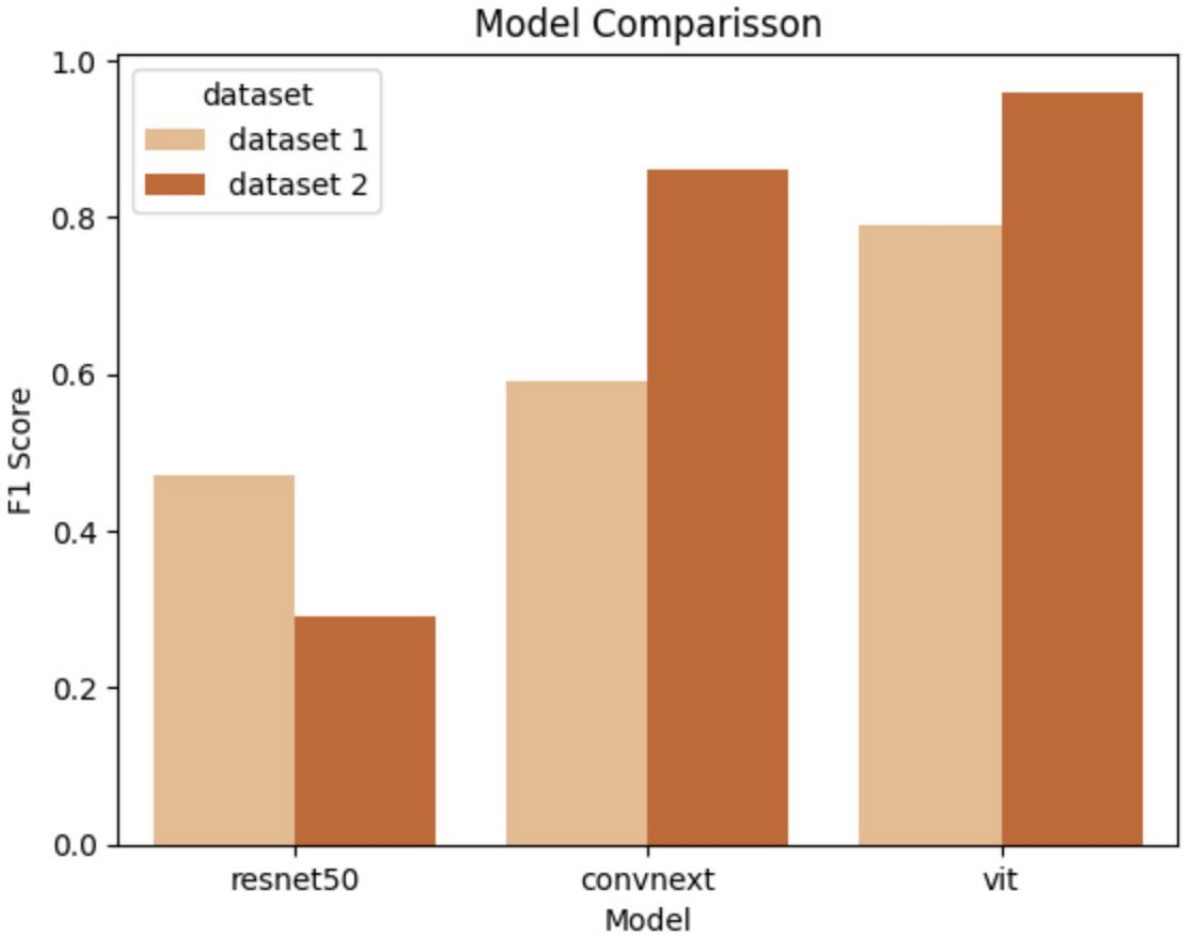
# Creating more data




Class Imbalance
images per class
0
50
100
150
200
250
drinks
15.5%
chips
13.2%
other
11.1%
Ice cream
10.6%
cigarettes
20.2%
candy
29.3%
Class

# The final dataset  - 1,321 images




Model Comparisson
dataset
dataset 1
dataset 2
1.0
0.8
0.6
0.4
0.2
0.0
F1 Score
resnet50
convnext
vit
Model

# Model results

| model | accuracy | precision | recall | f-1score |
|---|---|---|---|---|
| ResNet50 | 0.33 | 0.44 | 0.22 | 0.29 |
| ConvNeXt | 0.85 | 0.84 | 0.9 | 0.86 |
| ViT | 0.95 | 0.94 | 0.92 | 0.96 |