

Computer Vision in Convenience Retail

Dennis Francis



The Images

Scraped from the web



Labeled Data

ice cream



cigarettes



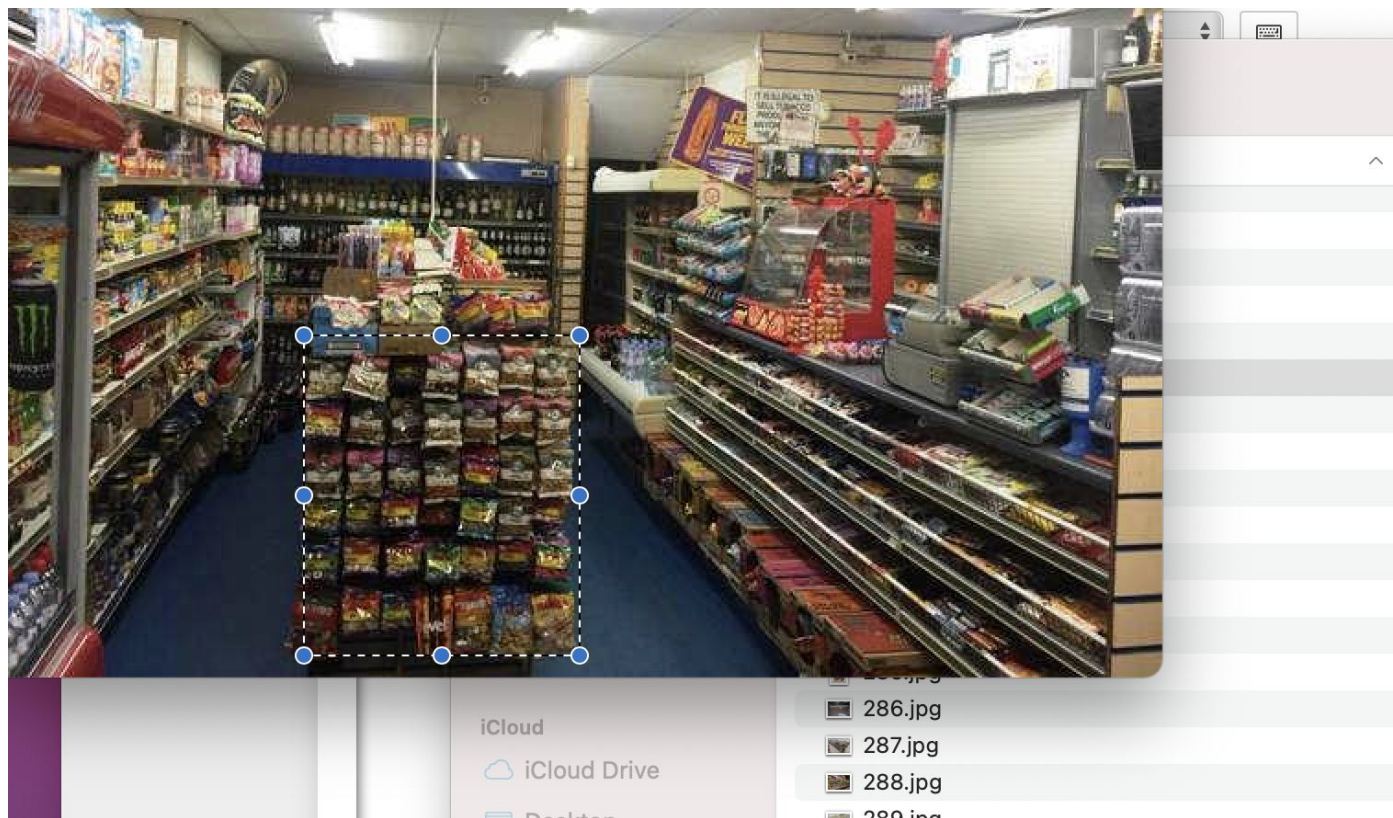
Cropping an image



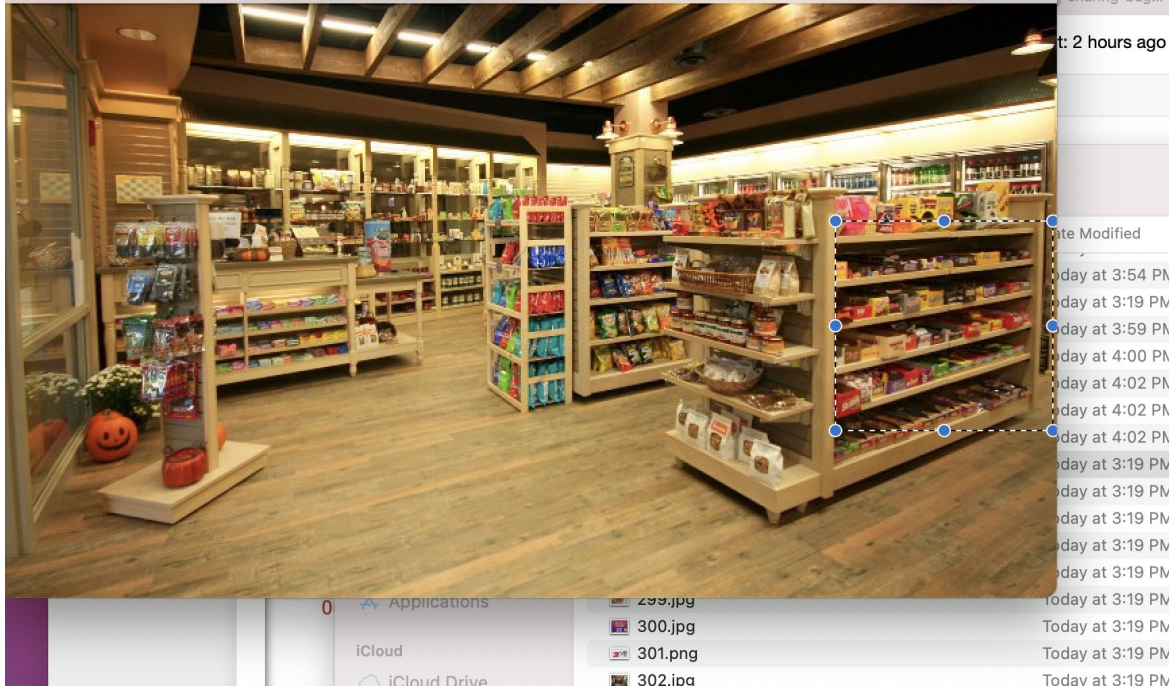
Cropped Image



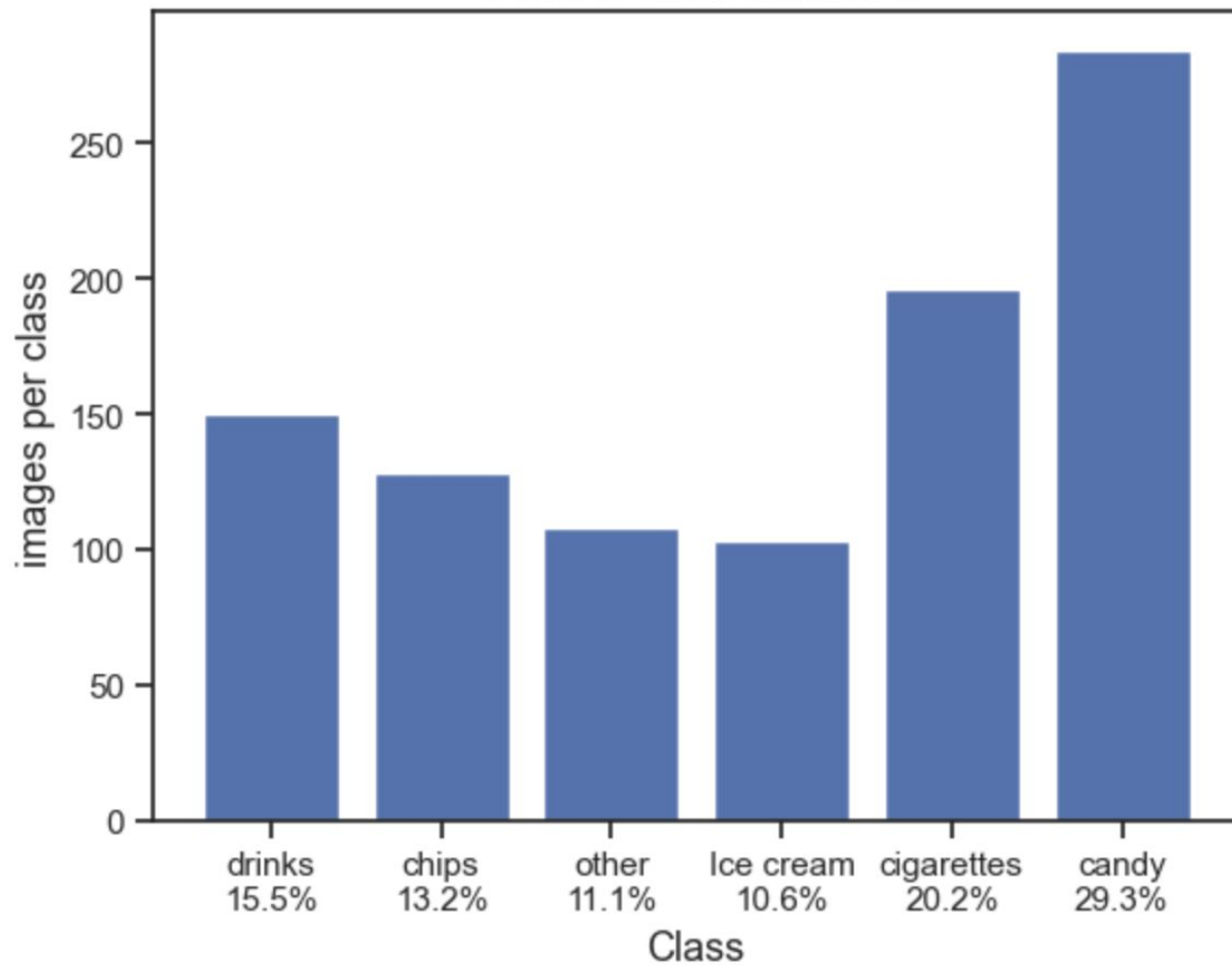
Creating more data



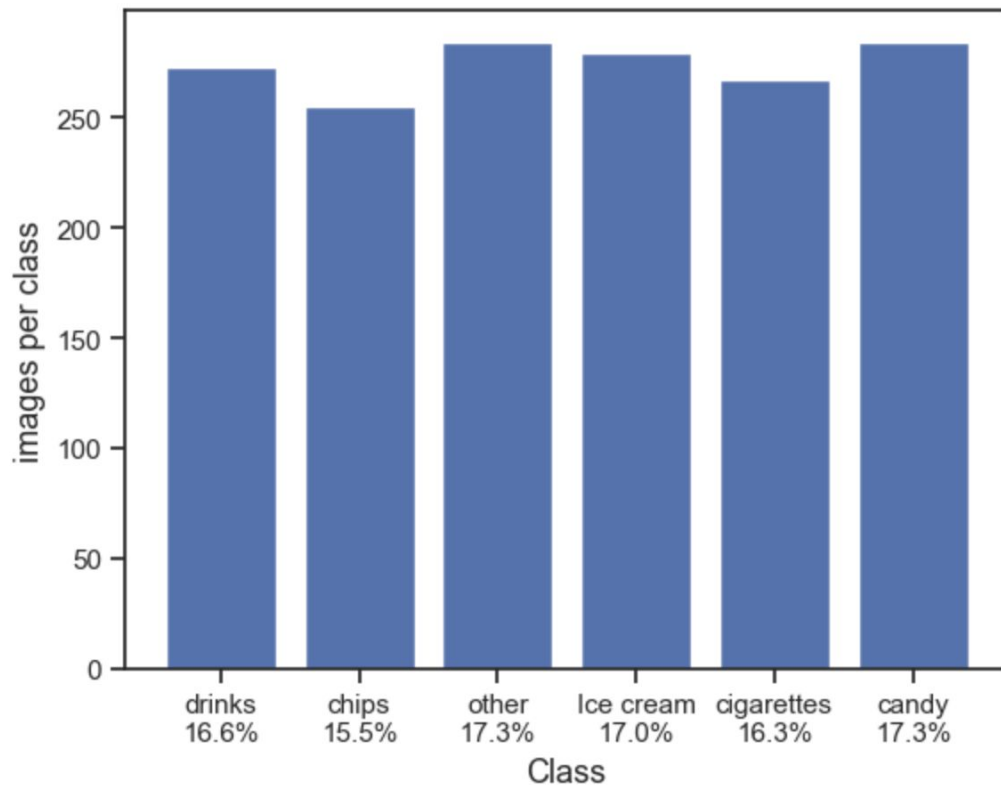
Creating more data



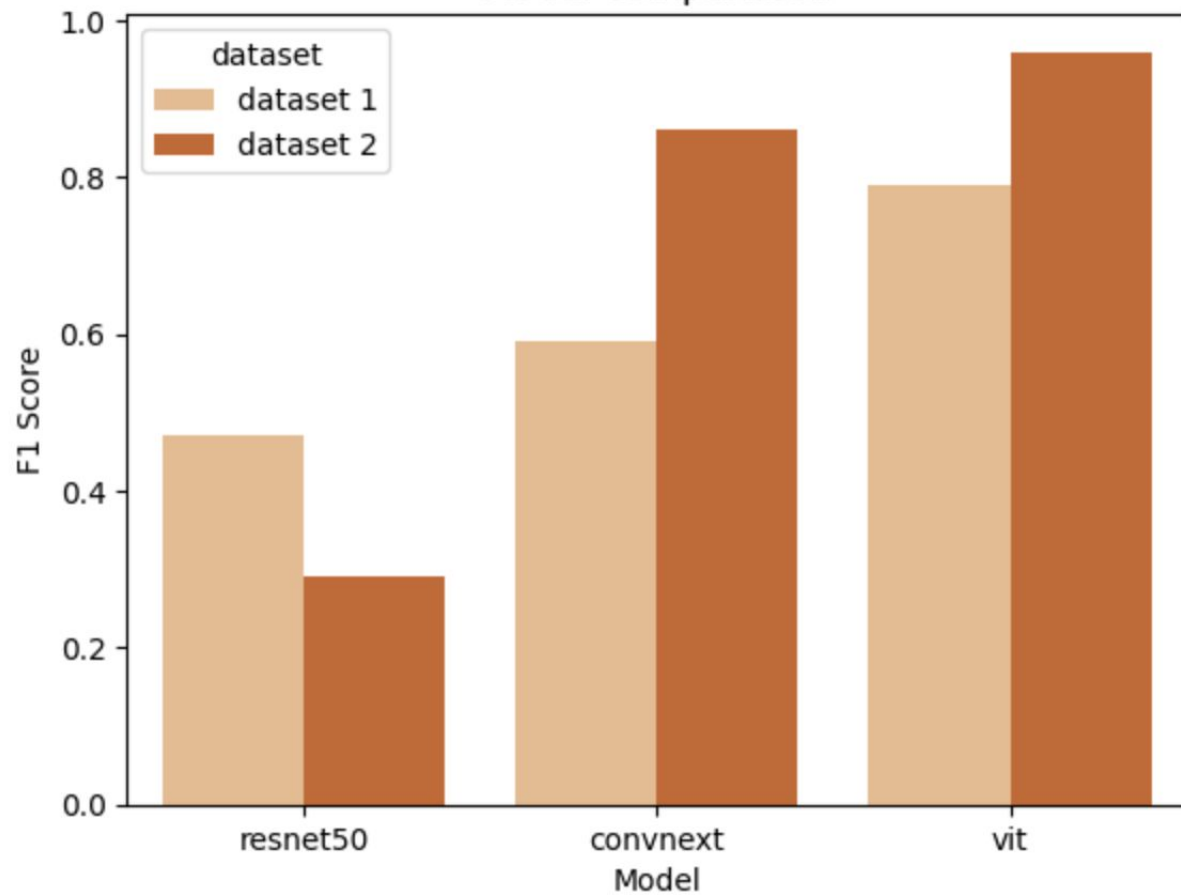
Class Imbalance



The final dataset - 1,321 images



Model Comparisson



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

