Shop Image Recognition

Springboard Capstone 3

Nicholas Mai, September 2022

Business Problem

- Shopee is a major e-commerce brand in South East Asia with a lowest-price guarantee program
- Image recognition is used to improve the lowest-price guarantee program
- This is done using image classification models to detect label groupings

Data

- Features of importance
  - 32,412 images
  - 11,014 image classes / label groups
- Other features
  - Positing id
  - Image phash
  - Title

Data Wrangling

- As this project focuses on building a functioning neural network for image classification, the number of label groups at 11,014 classes requires investigation
  - 63% of the label groups contained 2 images
  - The high number of classes compounded with the limited images per a class led to the decision to decrease the number of images
  - Top 21 label groups with 827 images images
    - Largest label group has 51 images

Data Wrangling

- Using DataLoader images and label groups are added to PyTorch
- An augment image toggle function is included
  - When on an images has a 50% change of being augmented as:
    - Random Vertical flip, Random Horizontal flip, Random Grayscale
- Data is split into train (80%), validation (10%), and test (10%)

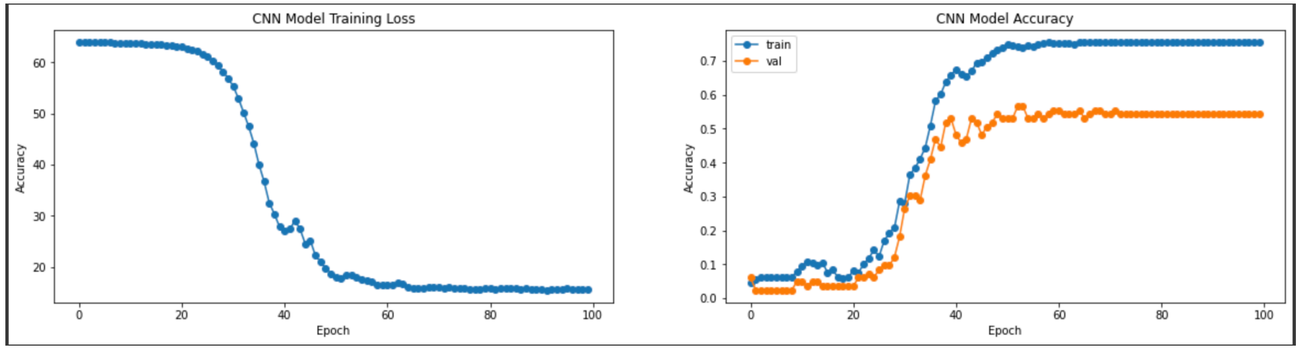
Model Fundamentals

- 3 models are used and tested against each other to improve prediction performance
  - Convolution Neural Network (CNN)
    - Personally made CNN model with: 3 convulsion layers, 1 flatten layer, and 2 linear dense layers
  - EfficientNet without pre-trained weights
    - A more complicated CNN model architecture
  - EfficientNet with pre-trained weights
    - Pre-trained model enabling transfer learning

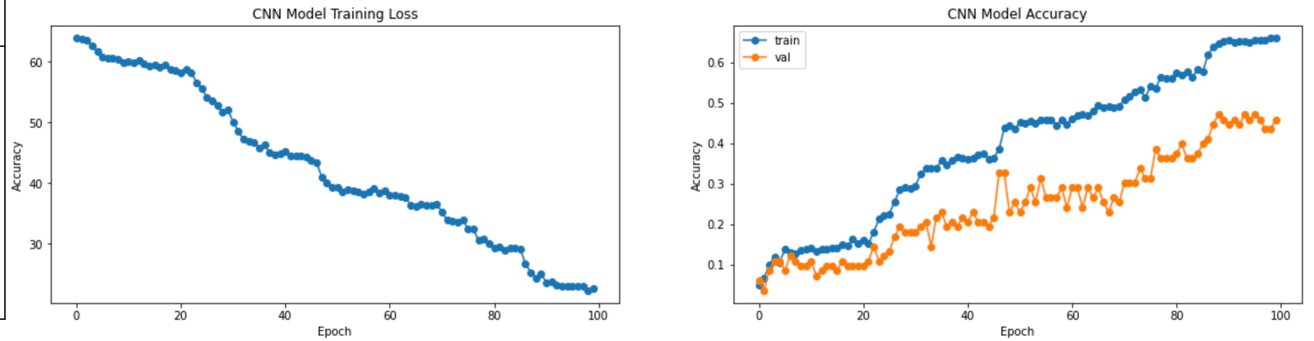
CNN model

	Train Accuracy	Validation Accuracy	Final Test Prediction Accuracy
Basic Model	80%	60%	27%
Augmentation & weight-decay	60%	40%	30%

Basic Model



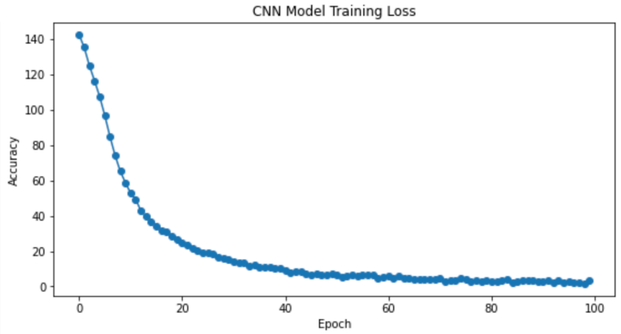
With Augmentation added Weight-decay = 0.001

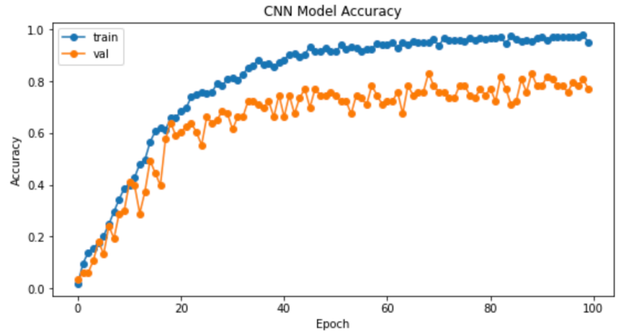


EfficientNet

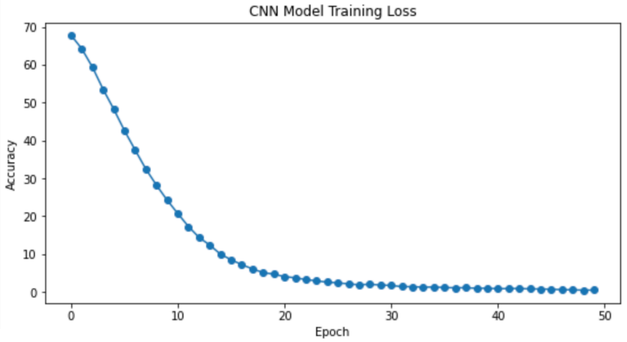
	Train Accuracy	Validation Accuracy	Final Test Prediction Accuracy
Basic Model	90%	70%	32%
Pre-trained weights, batch size = 74	95%	92%	89%

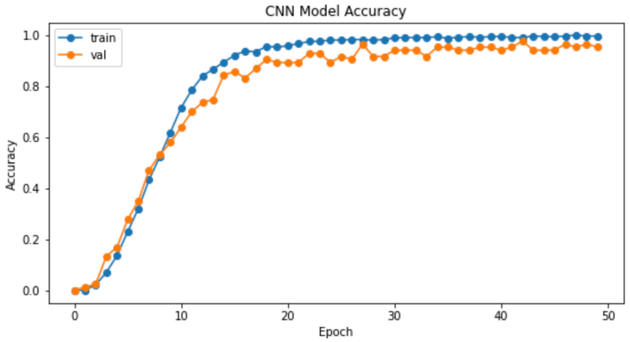
Basic Model





With Pre-trained weights





Recommendations

- Using pre-trained models increases the prediction accuracy at the greatest scale
- Results from the EfficientNet can still mimic simple CNN model, and requires repeated testing and understanding of the parameters involved
- Further testing with other models and transfer learning would aid in accommodating the full image data set