

Classifying groceries items by image using Convolutional Neural Networks

Derek Farren

The Freiburg Groceries Dataset

Recent announcements from Amazon regarding their "Just walk out" technology used in Amazon Go has brought attention in Computer Vision techniques used in brick and mortar retail.

This research intends to achieve better performance than previous work in classifying grocery items by using convolutional neural networks.

The Freiburg Groceries Dataset

This dataset contains images from real-world settings at different stores and apartments. In contrast to existing groceries datasets, it includes a large variety of perspectives, lighting conditions, and degrees of clutter.

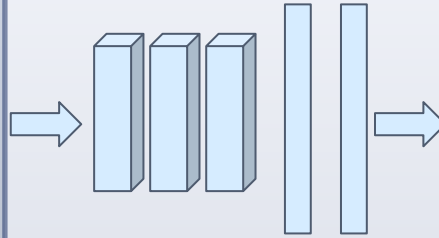
5,000 images covering 25 different classes of groceries, with at least 97 images per class.

The state of the art classifier on this dataset achieved a mean accuracy of 78.9% [1].



[1] The Freiburg Groceries Dataset, Philipp Jund, Nichola Abdo, Andreas Eitel, Wolfram Burgard

Network Architecture

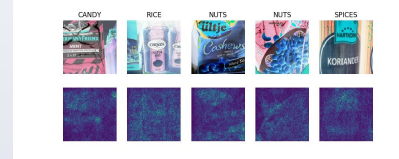


Input: 256x256x3 image

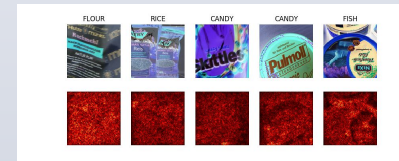
1. 11x11x3 CNN layer + Max Pool + Batch Norm
2. 5x5x16 CNN layer + Max Pool + Batch Norm
3. 3x3x96 CNN layer + Max Pool
4. 2048 Fully connected layer
5. 2048 Fully connected layer

Output: 25 logits

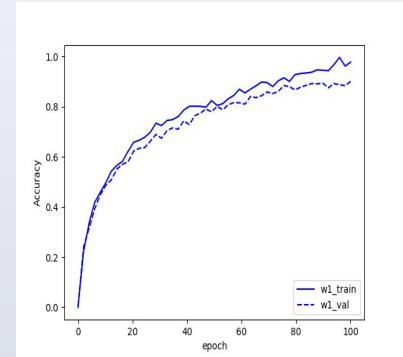
Saliency maps



In general, text and edges seems to be the main feature defining classes.



Model results



Accuracy: 89.12%

The state of the art classifier on this dataset achieved a mean accuracy of 78.9% [1].