Testing Real-Time Food Recognition Using Convolutional Neural Networks

Edgar De La Torre, Fabian Gonzalez
{edgar.delatorrebazan01, fabian.gonzalez03}@utrgv.edu
02/14/19

Summary of the Proposal

Sometimes when traveling across the world we come across dishes that we cannot identify. Identifying these dishes computationally is complicated due to the different variations across different locations. We believe that using machine learning techniques for this type of subject will allow users to more appropriately identify what meal they are eating. Our proposal is developing an application that uses a convolutional neural network to classify different meals. We also plan to allow users to train the model by allowing them to take pictures of their meals and identify their attributes. At the end we intend to test other image classification algorithms to test for any improvement in classification.

Background

Digital media has grown exponentially due to the growth of the internet and technology, and with their growth many images and videos are stored, shared, and analyzed by many in order to pursue some goal they may have. Focusing on the analytical portion of digital media comes the usage of technology to help in detecting and recognizing objects, whether they be inanimate or animate. Object detection, recognition, and tracking has been a field researched by many due to how it helps advance the technology used in many industries and its help extends from automatically tagging images to help people with eyesight-impairment, to security to identify people through surveillance. For both of those reasons and any of the others extending, while we may high speed and high resolution cameras, the technology used to recognize what is shown in those images is still in need to catch up.

The rise of research behind deep learning methods has now played a key role in the field of computer vision. Through the use of CNN (convolutional neural networks), image detection and analysis speeds have increased and continue to do so every other year. From the original proposed R-CNN (which focuses on generating custom regions to try and detect or recognize an object in) to the newer Faster R-CNN, object detection based on region has grown to be the fastest, and most efficient, method used. According to Microsoft Research, their method can detect objects at the speed of 5 through 17 frames per second, which makes it very efficient based on what it’s trying to accomplish per frame. Aside from that, the growth in image datasets, such as COCO (Common Objects in Context), provide researchers with plenty of training and test images which aid in the efficiency in testing out new ideas.

Using the newly done research, we will test out Microsoft’s Faster R-CNN in combination with other learning networks for object recognition to try and improve the accuracy and speed in trying to detect food in real-time.

Goal and Objectives

The goal of this research is to implement a recent method proposed by Microsoft for region-based convolutional neural networks in order to provide with near real-time object detection. Through that object detection we will implement different methods to recognize the objects (which
in this case will be food) detected and continue to try and speed up the detection, prediction, and accuracy of the objects presented.

**Data and Methods**

Using our own custom data and existing datasets, we will train the Faster R-CNN alongside multiple versions of object recognition algorithms and networks through TensorFlow (and possibly OpenCV) to see if we could either speed-up or increase the accuracy of trying to detect food present. We would also try and train our data using smaller images (256 x 256) in order to decrease the time and power necessary to finish and run our tests. We still have more learning and research to do before knowing what exactly would be needed for our training and research.

**Image Data:**
- **Dimensions:** 256x256 pixels
- **Content:** Various Foods
- **Amount Per Class:** Unknown

**References**