FINDING INGREDIENTS OF PIZZA USING DEEP LEARNING

Alim Giray Aytar, Hayati İbiş & Mümin Can Yılmaz
Department of Computer Science and Engineering
Hacettepe University
Ankara,TURKEY
{hayati.ibis12,giray.aytar12,can.yilmaz12}@hacettepe.edu.tr

ABSTRACT

In this paper we give information about progress of our project. We gathered data-set using python web crawler, and now we have 2K images. We realize that is not enough. We decided to support the data-set with manual entries.

1 INTRODUCTION

Our goal with this project is to create a system which will be able to tell you what ingredients are in the pizza. We will use deep learning approach to reach our goal. We will learn our model with our own data set. This is because, traditional machine learning approaches (like logistic regression etc.) can't make good estimations but, distributed representations gives better results. Convolutional layers give us low-mid-high level features. As we move through the layers, the combinations we obtain will ensure that the machine will see more clearly what is inside the pizza.

Finding ingredients from images is about Deep Learning but it is also a Computer Vision problem. For solving this problem we will use Deep Convolutional Neural Networks. For now, we are on the phase of collecting dataset. To train a powerful model, we need to collect a large datasets. This is what we need to do to improve performance. It is difficult to find a large data set that is correctly labeled. We are creating an user interface for collecting more data from people. While continuing the project (e.g.building model), we will continue to collect more data.

Ingredients : Sausage, Mushroom, Cheese, Corn, Tomatoes

The rest of the paper is as follows. We briefly review the related studies in Sec.2. In Sec. 3 we describe the proposed methodology in detail. We discuss future direction in Sec. 4.

2 Related Work

Computer Vision is one of the most important topic of Machine Learning. Thus, the range for researching for studies is huge. For beginners like us, the research part is one of the most important topics of projects. But, Finding the right studies for our model might be hard.
We are about to design a model that includes deep learning methodology. Thus, our model includes convolutional neural network.
Firstly, We started looking and analyzing the architecture of studies of ImageNet Classification. And AlexNet[3] was the first study we come across with. AlexNet[3] contains eight learned layers. Five of these layers is convolutional and three of them is fully-connected. And They were working with ImageNet dataset. But our work is limited with pizzas. So, we also found some studies on food and ingredient recognition.
Deep-based Ingredient Recognition for Cooking Recipe Retrieval[4], in this study; they used a modified version of DCNN and created a structure that has four different deep architectures for multi-task learning of food category and ingredient recognition.
Food Recognition Using Statistics of Pairwise Local Features[5], in this study; the researchers used a pair wise features(PDF). But the model they used is not a network, but standard baseline algorithms specified by PFID(pairwise features): color histogram + SVM(Support Vector Machine) and bag of SIFT features + SVM.
In the process of advancing our design, we will analyze deep into these related works which have been done before us. But, the best way of finding a good model also depends on the experience that we will have with our data set.

3 Methodology

In our project, we are planning to use a (supervised) deep learning-based approach to address the ingredient extraction problem in food recognition. To be more specific, we are going to use a convolutional neural network (CNN), because earlier studies say that they show better performence compared to other methods[1][2]. There are certain advantages of using this method especially for images, for example, features are extracted automatically while training, without extra effort.

Because there are no ready-to-use datasets for our problem, we collected our own data and (for most parts) labeled the images manually. For now, our dataset is pretty limited (2K-3K images), but we are planning to expand it. After we finished building our dataset, we will release it for public use. In addition to that, for our project, we will resize our images 128x128 pixels (but we are not sure of the dimensions for now, other resolutions may provide better results, so we will try some other configurations before deciding final size). We are not going to greyscale our images, because colors can be very helpful when recognizing images.

In our architecture, the input layer will be the size of 128x128x3 and the output layer will be the ingredients appearing in the all images. For example, assume that our model will recognize only 5 ingredients (olives, tomatoes, salami, cheese and mushroom). In this case, our output layer will the size of 5, each one represents the ingredients probability of appearing in the image.

To calculate our program's accuracy, we are going to use k-fold cross validation method. We are planning to split our dataset into 5 sections, 3 for training, 1 for validation and 1 for testing.

We also planning to use some pretrained models like GoogleLeNet and then fine-tune our model with out dataset.

4 Future Work

We are currently working on obtaining the data set from the Web. We are building a data set from nothing. So , there's some hard labeling problems but we will continue gathering image around the web.

After enough number of train sample is gathered, we will start to design a network model. We will try to enhance our model with trials on our data set and with related works.

REFERENCES

1 - https://ai2-s2-pdfs.s3.amazonaws.com/0373/442fb84f7783f3453c46f8cccc12d184dbbc.pdf
2 - Y . Kawano and K. Y anai, "Food image recognition with deep convolutional features," in Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, 2014, pp. 589-593.
3-https://www.cs.toronto.edu/~kriz/imagenet_classification_with_deep_convolutional.pdf
4-http://vireo.cs.cityu.edu.hk/papers/jjchen_mm16_cr.pdf
5-http://homes.cs.washington.edu/~shapiro/cvpr10.pdf