Deep Learning Based On Food Image Recognition

Shreyaa Sridhar
School of Computing and Engineering
University of Missouri-Kansas City
ssn2m@mail.umkc.edu

Geovanni Nicque West
School of Computing and Engineering
University of Missouri-Kansas City
gwrn2@mail.umkc.edu

Khushbu Kolhe
School of Computing and Engineering
University of Missouri-Kansas City
kkvcm@mail.umkc.edu

Naga Venkata Satya Pranoop Mutha
School of Computing and Engineering
University of Missouri-Kansas City
nmgfg@mail.umkc.edu

Abstract

Good nutrition is an important part of leading a healthy lifestyle. Food choices that we make each day affect our health - how we feel today, tomorrow and in the future. In addition to that, predicting calories also plays an important role because it allows for you to have a tight handle on how many calories you are presumably taking in. We propose an application which recognizes food and predict calories, also a comparison of shallow and deep learning model's accuracy based on the image taken.

Keywords: Deep learning, food image recognition, Food-101 dataset

I.INTRODUCTION

Accurately recognizing food images is a particularly challenging task based on the nature of images, which is why approaches like Spark API models in the field have achieved a low classification accuracy. Deep neural networks have outperformed such solutions, and we present a approach where we are comparing the accuracy of different shallow and deep learning models. Here we have retrained the models like Clarifai , Inception V3 , MobileNet models with our dataset and comparing their predictions.

Considering the food recognition problem, it has been broadly investigated for many years and there have been several systems as well as datasets related to this problem. In FoodAI[8], authors study a Singaporean food recognition and health care system by using a deep learning approach. Similarly, CLARIFAI[9] is constructed for recognizing Western foods.

II.RELATED WORK

NutriNet model [2] was built to recognize food and drinks. NutriNet+ was built on 6 convolutional layers when compared to AlexNet it has one extra layer.This convolutional layer was added to gain additional knowledge about the features in the higher resolution images.NutriNet model uses UNIMIB 2016 Food image dataset. NutriNet with the NAG solver was the best-performing model with a classification accuracy of 94.47%. This

model majority of images are correctly classified in the recognition dataset, but not necessarily all of them. As a consequence, this could lower the classification accuracy for real-world images. Finally, since image segmentation was not performed, irrelevant items present in the images made the recognition difficult.When comparing NutriNet to NutriNet+, we can see that the extra convolutional layer did not yield any performance increase, as NutriNet+ models achieved results that are almost identical to the results by NutriNet models.

A food image recognition system that uses the multiple kernel learning method was introduced, which tested different feature extractors, and their combination, on a self-acquired dataset [5]. This proved to be a step in the right direction, as the authors achieved an accuracy of 26% to 38% for the individual features they used and an accuracy of 61.34% when these features were combined; the features include color, texture and SIFT information. Upon conducting a real-world test on 166 food images taken with mobile phones, the authors reported a lower classification accuracy of 37.35%, which was due to factors like occlusion, noise and additional items being present in the real-world images. The fact that the combination of features performed better than the individual features further hinted at the need for a more in-depth representation of the food images.

Pairwise local features method, which applies the specifics of food images to their recognition [6], analyzes the ingredient relations in the food image, such as the relations between bread and meat in a sandwich, by computing pairwise statistics between the local features. The authors performed an evaluation of their algorithm on the PFID dataset and achieved an accuracy of 19% to 28%, depending on which measure they employed in the pairwise local features method. However, they also noted that the dataset had narrowly-defined food classes, and after joining them into 7 classes, they reported an accuracy of 69% to 78%. This further confirmed the limitations of food image recognition approaches of that time: if a food image recognition algorithm achieved a high classification accuracy, it was only because the food classes were very general (e.g., “chicken”).

III. PROPOSED WORK

PROPOSED MODELS

In our project, we have retained models like Clarifai API , Tensorflow (uses Inception V3 Model) and TensorflowLite Model (uses MobileNet Model). Spark API works well on small dataset , but throws an error when we work on large dataset. Clarifai API Model uses its own dataset which recognizes Western foods, so we retrained the model with our dataset of 10 foods. Tensorflow uses ImageNet dataset which recognizes the food images well. In our model, we have retrained the model based on our dataset and it recognizes well. In case of TensorflowLite model uses MobileNet model, also uses ImageNet dataset which recognizes the images well. Here again we have retrained the model based on our dataset and then converted to .tflite format. After conversion, we are predicting the food image considered.Our model works best with Clarifai API Model when compared with all other models. We are predicting the calories of our food image using Nutritionix API.

A. Analytical Tools and IDE Used

- Apache Spark

- Android Studio
- Tensorflow
- IntelliJ
- PyCharm

B. Expected Inputs/Outputs

- **Input** - Images from Food-101 dataset

- **Output** - Image Categorization, Image Prediction, Statistical parameters like accuracy, precision, confusion matrix, etc.

C. Feature Specification

- Food Image Recognition
- Predict Calories using Nutritionix API

ALGORITHMS

For analysis we have used :

- Clarifai API Model.
- Shallow Learning : Spark API Model.
- Deep Learning : Softmax Regression Model, CNN Model, Inception V3 Model and MobileNet Model.

Training and testing data is mostly considered as 70% for training data and 30% for testing.

IV. IMPLEMENTATION AND EVALUATION

SYSTEM DESIGN AND IMPLEMENTATION

A. SOFTWARE ARCHITECTURE

The architecture diagram in figure 1 explains how our model works. Images from Food-101 dataset is provided as input to our model. Here we have compared how our dataset works with different models such as Clarifai API model , Spark API model and Deep Learning models(SoftMax Regression Model, Convolutional Neural Network Model and Inception V3 Model) . The next step is running our model to obtain annotations for each image, check accuracy and obtain confusion matrix.

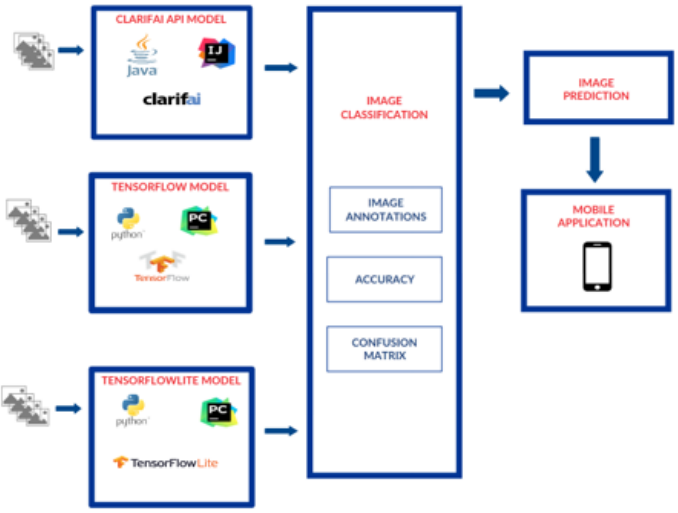

Figure 1 : Architecture Diagram

B. UML DIAGRAMS

Activity Diagram

The activity diagram figure 2 shows the workflow of our model.

Sequence Diagram

The sequence diagram figure 3 shows the interaction or collaboration within our model.

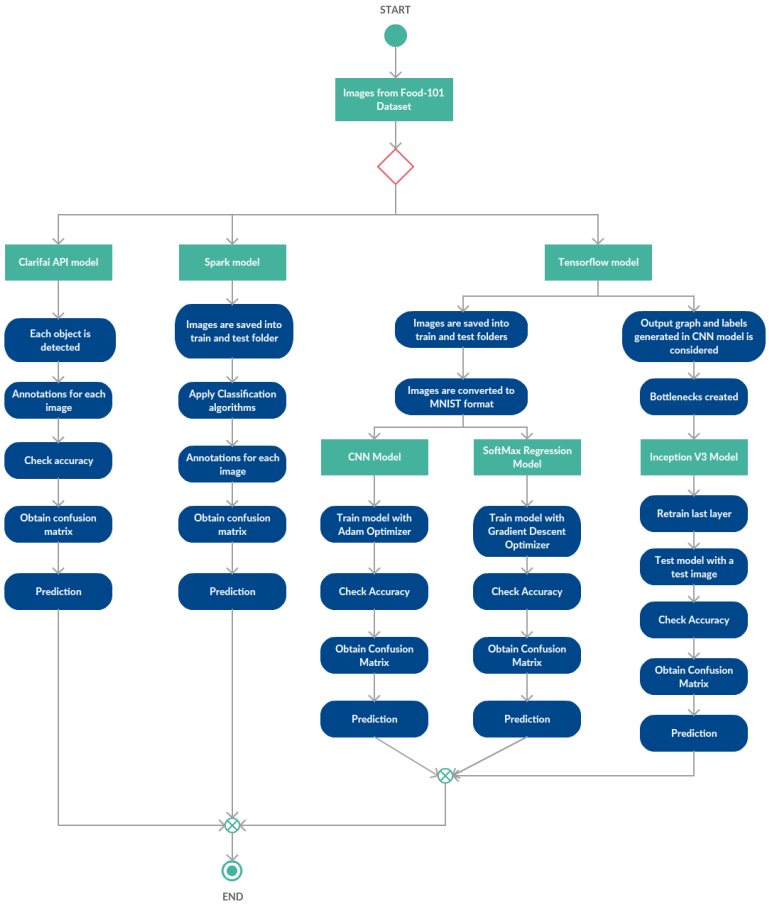

Figure 2 : Activity Diagram

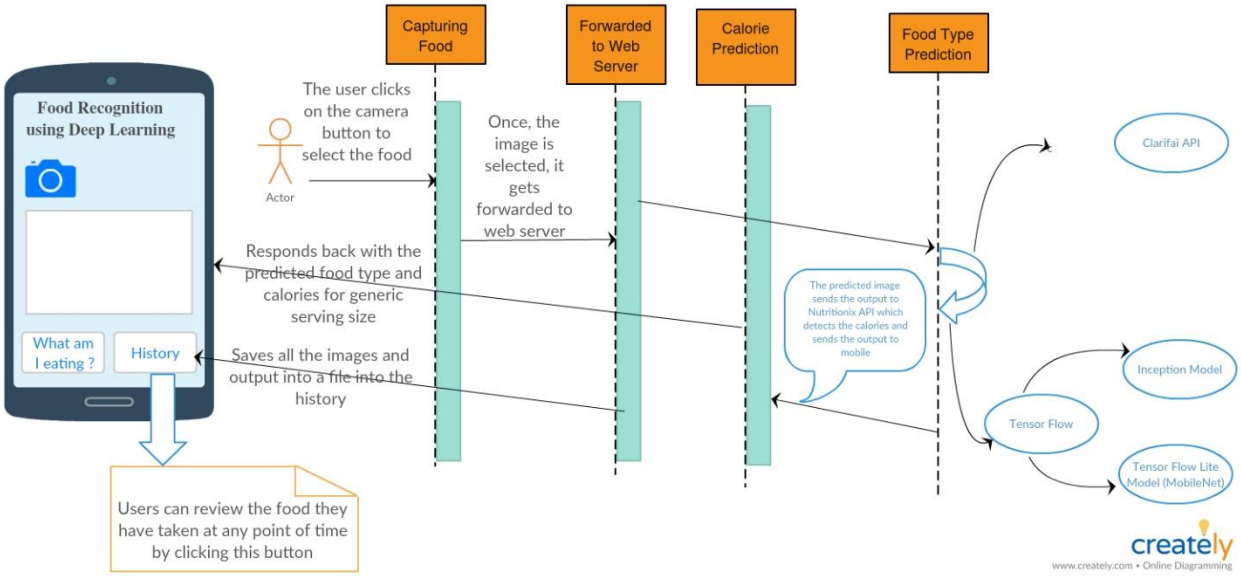

Figure 3 : Sequence Diagram

C. IMPLEMENTATION DETAILS

User selects the camera icon to take picture of the food. The food image is sent to the server which handles the image recognition . Here we have included history which displays all the food history the user has checked previously. Calories are obtained by Nutritionix API considering any food image as for one serving.

Source Code Link :
https://github.com/shreyaabadri/Big-Data-Analytics-And-Applications-Project-Spring-2018/tree/master/Project%20Increment%203/Source%20Code

D. APPLICATIONS

The figure shows the basic Mobile UI of our application. Camera function is included to take real time food images and predict them. **'What I am eating?'** allows you to see the predictions for food images and **'History'** shows the food images predicted with Food name mentioned in Clarifai since it has better accuracy and calories using **Nutritionix API.**

Figure 4 : Launch Page

Figure 5 : After taking image

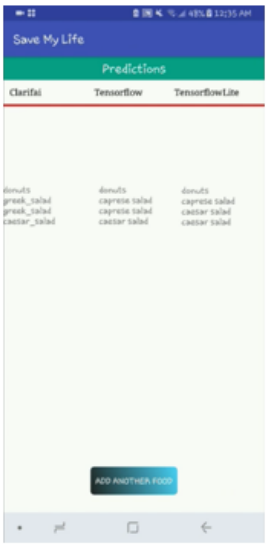

Figure 6 : Predictions for Clarifai , Tensorflow and Tensorflow Lite models.

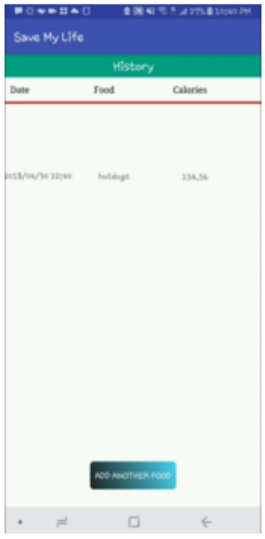

Figure 7 : Food History

Video Link :

https://www.youtube.com/watch?v=b-ZRdeyZFuE&feature=youtu.be

Presentation Link :

https://docs.google.com/presentation/d/1RaRwBdeyimgsn-cWrkjxZCJF6E1f-OgiFxmHGiwNWzo/edit?usp=sharing

EVALUATION AND RESULTS

A. EVALUATION PLAN

Datasets

We have taken images of food from **Food-101 dataset**

https://www.vision.ee.ethz.ch/datasets_extra/food-101/

This dataset contains 101 food categories and 1000 images in each category. We have considered 10 food image categories for our project - **caesar salad, caprese salad, donuts, dumplings, french fries, greek salad, guacamole, hotdog, risotto and sushi**. We have considered similar categories like three different salads and our models predicts well all the categories. In our project, we have retrained our models with our dataset and obtained confusion matrix and predictions .

Figure 8 : Sample of our Dataset

System Specification

- Processor – Intel Core i7 Processor
- Operating System – Windows OS
- RAM – 8GB

Measurements

Our dataset used consists of 300 images in each image category. Images are close up images of food and are preprocessed with before training the model.

B. EVALUATION AND RESULTS

Comparative evaluation for three different approaches- Clarifai, TensorFlow, TensorFlow Lite

Model	Accuracy
Clarifai API Model	95.2%
Tensorflow Model	88.4%
TensorflowLite Model	83.8%

Figure 9 : Accuracy of our models used in application

1. Clarifai API Model

In our project, we are using Clarifai API Model for food image recognition We have retrained the Clarifai model with our 10 image categories. It has predicted an accuracy of 95.2%. In general, Clarifai API prediction is based on what model you run the input through. For example, if you run your input through the 'food' model, the predictions it returns will contain concepts that the 'food' model knows about.

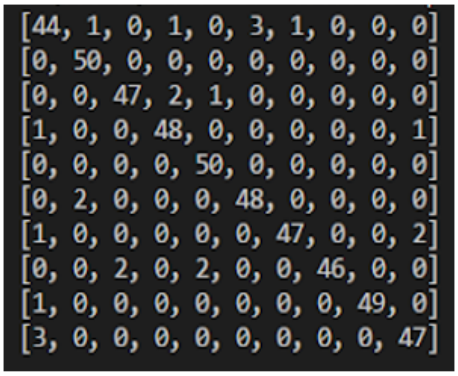

Figure 10 : Confusion Matrix of Clarifai Model

2. Tensorflow Model (Inception V3 Model)

In our project, we are also using Inception V3 Model for food image recognition. We have retrained the Inception V3 model with our dataset of 10 image categories. Our Inception V3 model has predicted an accuracy of 88.4%.

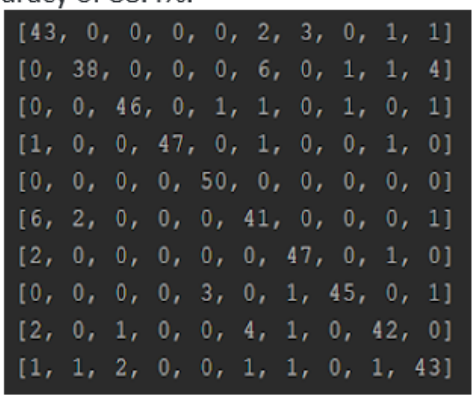

Figure 11 : Confusion Matrix of Inception V3 Model

3. TensorflowLite Model (MobileNet Model)

In our project, we are also using MobileNet

Model for food image recognition. We have retrained the MobileNet model with our dataset of 10 image categories and converted to .tflite format to use for prediction. Our model has predicted an accuracy of 83.8%.

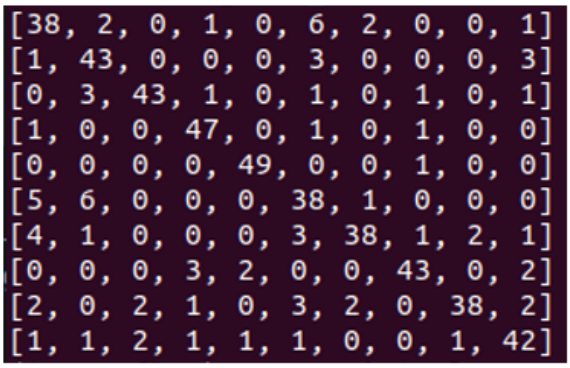

Figure 12 : Confusion Matrix of MobileNet Model

Comparative evaluation with others

Model	Accuracy
Spark API Model	44.04%
SoftMax Regression Model	41.85%
CNN Model	51.3%

Figure 13 : Accuracy of models

1. **Spark API**

Spark API works well on small dataset and by using the dataset on different models like Random Forest Model, Decision Tree Model and Naive Bayes Model, we found that Random Forest Model works well for our model. The accuracy of our model using Spark API is 44.04%

2. **SoftMax Regression Model**

SoftMax Regression was tested on our dataset of 10 image categories. Our Model has predicted an accuracy of 41.85%.

3. **CNN Model**

Convolutional Neural Network works well on image dataset. It was tested on our dataset of 10 image categories. Our Model has predicted an accuracy of 51.3%.

V. DISCUSSION AND LIMITATION

This data set contains of 101 food types, each food type consisting of 1000 images. In comparison among different models used for our dataset, we have got the best accuracy for Clarifai API Model and least for Spark API model. Though, the accuracy looks very low except on Clarifai, it is mainly because we restricted input to only ten food types, which is minute compared to the data set we have taken for reference. We tested with multiple food types, but taking five or six food types at a time. We observed that the accuracy varies for the models (i.e not one model stays low everytime, they keep altering). We have used similar types of foods like three different types of salads to check the efficiency of our model. In addition, we also try to display the calories present in each food type, and the accuracy comparison for each of the models. The solution to lost data is still under study. The provisions for editing or making changes in the app are not yet to be figured out. As future work, we plan to determine portion size of food images and to train our model based on different lighting conditions.

VI. CONCLUSION

In this paper, we have described our study of developing a food image recognition mobile application. Our dataset was trained on across different shallow and deep learning models. After comparing the accuracies of different models, we inferred that Inception V3 model works best and provides accurate result.

VII. REFERENCE

1. Yuzhen Lu, “Food Image Recognition by Using Convolutional Neural Networks (CNNs)”, Cornell University Library, Computer Vision

and Pattern Recognition,Dec 2016, arxiv:1612.00983
2. Simon Mezgec and Barbara Korousic Seljak, "NutriNet: A Deep Learning Food and Drink Image Recognition System for Dietary Assessment" NCBI, Nutrients, v.9(7), 2017 July, ,PMC5537777
3. Chen M., Dhingra K., Wu W., Yang L., Sukthankar R., Yang J. PFID: Pittsburgh Fast-Food Image Dataset; Proceedings of the ICIP 2009; Cairo, Egypt. 7–10 November 2009; pp. 289–292.
4. 5. Lowe D.G. Object Recognition from Local Scale-Invariant Features; Proceedings of the ICCV'99; Corfu, Greece. 20–21 September 1999; pp. 1150–1157.
5. 6. Joutou T., Yanai K. A Food Image Recognition System with Multiple Kernel Learning; Proceedings of the ICIP 2009; Cairo, Egypt. 7–10 November 2009; pp. 285–288.
6. 7. Yang S., Chen M., Pomerlau D., Sukthankar R. Food Recognition using Statistics of Pairwise Local Features; Proceedings of the CVPR 2010; San Francisco, CA, USA. 13–18 June 2010; pp. 2249–2256.
7. Artifact from the Future: GOTTA EAT 'EM ALL http://www.iftf.org/future-now/article-detail/artifact-from-the-future-gotta-eat-em-all/
8. http://foodai.org
9. http://blog.clarifai.com