Predicting Meal Vegetariansm Using Convolutional Neural Networks

Melvin van 't Holt, 2047886

1 Introduction

Recognition of food ingredients enable applications to detect individual ingredients using a single image. The article *Food Ingredients Recognition Through Multi-label Learning*[1] shows us how a Convolutional Neural Network (CNN) can be used to predict such ingredients with high accuracy. Because of the increase of vegetarianism in the 21st century[2] an application of such models in order to detect whether a meal is vegetarian or not could prove to be valuable for vegetarians around the world.

Our research question reads as follows:

To what extent can we predict whether a meal contains non-vegetarian ingredients using a single image?

2 Experimental procedure

For this project we used the Ingredients101 and Recipe5k datasets that were constructed for the above mentioned paper. The 406 unique ingredients in the Ingredients101 dataset were labeled to be either 'Vegetarian' or 'Not-Vegetarian' with the help of the FoodClassification dictionary. All 4826 unique recipes in the Recipes5k were then labeled by setting any recipe to be 'Non-Vegetarian' if it contained any 'Non-Vegetarian' ingredient and 'Vegetarian' otherwise. The same train/validation/test sets as the article were used. For the train set, the data was balanced by equalling the number of rows for each class by generating blurred images of the minority class using the Python Pillow package. This required the generation of 555 images of the minority class which were randomly selected. The resulting sets and label divisions are as follows:

Set	Label	Nr. of recipes	Percentage
Train	Vegetarian	1982	50%
Train	Non-Vegetarian	1982	50%
Validation	Vegetarian	386	61%
Validation	Non-Vegetarian	248	39%
Test	Vegetarian	462	59%
Test	Vegetarian	320	41%

To answer our research question, we attempted three different approaches: (i) directly predicting whether a meal is vegetarian or not using a binary classifier, (ii) predicting the 101 food types in the dataset using a muticlass classifier and assigning the most common class to the prediction, (iii) using the multi-label classifier from the article and converting the predicted ingredients to our two classes. Approach (iii) uses the labeled ingredients by setting the prediction to 'Not-vegetarian' if any of the predicted ingredients is labeled non-vegetarian. Like in the article each model used Adam optimizer with learning rate 0.001

Approach	Loss function	Output function
(i) Binary Classifier	Binary Cross Entropy	Sigmoid
(ii) Multiclass Classifer	Categorical Cross Entropy	Softmax
(iii) Multi-label Classifier	Binary Cross Entropy	Sigmoid

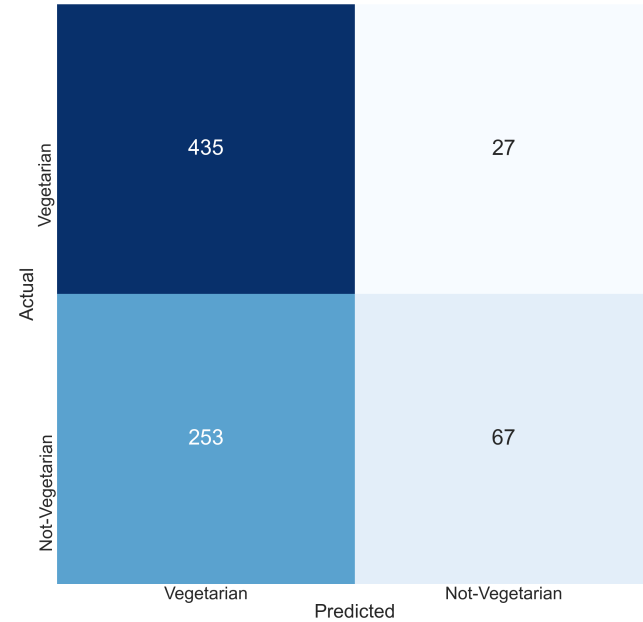
We used three different models for each of our approaches: (1) a simple CNN from a Fruit recognition article[3], (2) A pre-trained Resnet50 network from the Keras library, (3) A pre-trained InceptionV3 network from the Keras library.

The performance was evaluated using a simple two by two confusion matrix for our two classes, in which we aim to increase the accuracy. As a baseline, we attempt to find an accuracy score higher than simply predicting the majority class in our test set (i.e. higher than 59% accuracy).

Code from the Fruit Recognition article[4] and the Ingredient Recognition article[5] were used to create our models. However, in order to be able to run the models on the newest Keras and Tensorflow versions, the code was rewritten and can be found on my Github page[6]

3 Results

(i): The first approach did not provide a higher accuracy score than our baseline as it predicted only the majority class in our test set. This is likely due to the fact that one first needs to have knowledge of the meal in order to predict whether it is vegetarian. Also, non-vegetarian ingredients can be in many different forms (e.g. fish vs meat), or even unseen (e.g. empanadas).



(ii): Like our first approach, the second approach did not provide a higher accuracy score than our baseline as it failed to accurately predict the category of each meal. This is likely due to the fact that the 3409 images in our train set were not enough to reliably predict over 101 different classes.

(iii): Using the third approach, the pre-trained Resnet50 network (2) as our starting point proved to find higher accuracy scores than our baseline. From the confusion matrix above we can see that the model was able to predict the right label 502 out of 782 times, giving us an accuracy score of 64%. The model is only marginally better than our baseline as it still predicts non-vegetarian recipes wrongfully 253 out of 320 times (79%). The gain in recall on the 'Not-Vegetarian' (0% to 21%) outweighs the loss in recall on the 'Vegetarian' label (100% to 94%). Some of the other metrics are shown below.

Label	Precision	Recall	F1-score	Support
Vegetarian	0.63	0.94	0.76	462
Not-Vegetarian	0.71	0.21	0.32	320
macro average	0.67	0.58	0.54	782
weighted avg	0.67	0.64	0.58	782

The model was further trained on the train set. However, this only resulted in over-fitting on the 3964 images and provided no better accuracy scores. The simple CNN (1) and the pre-trained InceptionV3 network (3) failed to provide higher scores than our baseline.

4 Discussion and conclusion

Predicting whether a meal is vegetarian or not is not a simple task, as it requires knowledge of recipes as well as the fact that non-vegetarian ingredients can come in many different forms. Our model was only slightly better than predicting the majority class. Possible solutions to increase the models' performance would be to use more data such as the AIFood[7] dataset to train our model as it contains 372.095 images compared to the 4.826 images in the Recipes5k dataset or using other models created by more recent papers.[8]

If we want to apply our model in a real life scenario, the accuracy score might not be a suitable metric to evaluate the performance. Since false negatives will most likely have a higher cost (i.e. the model predicts vegetarian whilst the meal contains non-vegetarian ingredients). In this case, a better approach would be to evaluate on the F1 and/or Recall scores, or adjust weights in favor of the non-vegetarian classification.

We conclude that we can create a model that predicts whether a meal contains non-vegetarian ingredients using a single image that is slightly better than predicting only the majority class. The performance of the model will likely be inferior to humans in its' predictions but can provide value as starting point for creating a better model.

[1] Food Ingredients Recognition Through Multi-label Learning
[2] No meat today: The rise of vegetarianism and veganism
[3] Fruit Recognition Article
[4] Fruit Image Code
[5] Food Ingredients Code
[6] Github page containing the code
[7] AIFood Dataset
[8] Deep-based Ingredient Recognition for Cooking Recipe Retrieval