IMAGE ANALYSIS OF PLANT BASED MEAT PRODUCTS

CS 499: Report 1

Authors
Badal Chowdhary - 20110034
Hitesh Jain - 20110077
Under the guidance of **Prof. Shanmuganathan Raman**

1 Introduction

In the ever-evolving culinary landscape and with changing dietary preferences, plant-based meat products have undoubtedly initiated a revolution in the food industry. These innovative products offer a sustainable and environmentally conscious substitute for conventional meat, catering to the discerning preferences of today's consumers. The goal of our study is to comprehensively explore the realm of plant-based meat products, with a focus on understanding the complexities of their relationship with traditional meat products. Ultimately, our aspiration is to develop a discerning model that can classify plant-based meat products uniquely when compared to their counterparts in the world of conventional meat.

2 Previous Work

2.1 Dataset Creation

The previous research included a creation of dataset of images encompassing 14 classes of plant-based meat product patties. The patties were obtained from two sources: 1) the IITGn Food Lab product, and 2) the commercial Tata product. Two distinct cooking methods, 1) air frying and 2) deep frying, were utilized to prepare the patties. Standardized cooking durations were followed, along with an additional duration to create overcooked patties, depending on the type of patty and cooking method. Ultimately, 14 distinct classes were established based as listed in Table 1 in the types of patties and cooking methods used.

Table 1: Dataset Class Description

Commercial Tata Product	IITGN FoodLab Product	IITGN FoodLab Product Degraded 15 days & Starch Coated
• Unbaked • Deep Fried – Normal Cooked – Over cooked, 4min extra • Air Fried – Normal Cooked – Over cooked, 10min extra	• Unbaked • Deep Fried – Normal Cooked – Over cooked, 2min extra • Air Fried – Normal Cooked – Over cooked, 20min extra	• Deep Fried – Normal Cooked – Over cooked, 2min extra • Air Fried – Normal Cooked – Over cooked, 12min extra

2.2 Methodology and Results

The dataset consisted of a total of 10,900 images, with around 800 images obtained in each class, captured using a mobile phone camera under different lighting conditions and settings. Different backgrounds were utilized to reduce the impact of background noise and bias. The approach followed a three-stage process involving pre-processing the images, selecting an appropriate model, and training and fine-tuning the model. In the first stage, the Lanczos re-sampling technique was used to standardize the image size and resolution. In the second stage, the ResNet50 model, a deep neural network implemented widely in image classification tasks was chosen, and utilized transfer learning to leverage the pre-trained weights on the Food101 dataset - a vast collection of food images from diverse categories. In the third step, the ResNet50 model was fine-tuned with the sampled images to improve its accuracy in classifying the given inputs. To ensure that the model was not biased, the dataset was partitioned into 70% for training, 15% for testing, and 15% for validation. As a result, the method achieved a train accuracy of 97.6%, validation accuracy of 98.56%, and a test accuracy of 94.62% for categorizing plant-based meat analogous products.

3 Objectives and Proposed Strategy

3.1 Creation of new Dataset for meat products

The current dataset exclusively consists of images of plant-based meat products. Consequently, the goal is to construct a new dataset that includes non-vegetarian meat products to enable a comprehensive comparison and quantify the similarities and disparities between the two. Given that a significant practical application of this research involves industrial quality testing and assurance, a consistent approach to image acquisition is essential. Specifically, we will standardize the background, camera setup, lighting, and overall conditions during the image capture process.

3.2 Curation of Existing Dataset

The current dataset comprises images of products taken against varying background colors. However, to ensure the industrial quality testing and assurance, it is necessary to employ a standardized imaging technique, and the choice of background colors becomes irrelevant. With this in mind, our objective is to refine the present dataset by removing the background colors and extraneous noise.

The dataset under consideration comprises of a total of fourteen classes. However, we have identified that certain classes among them are redundant and could be excluded without affecting the quality of the dataset. Specifically, in the current dataset, two different cooking methods, namely Deep Frying and Air Frying, are used for every three types of products. However, according to research presented in [1], the cooking method employed does not have a significant impact on critical parameters such as appearance, color, and texture of the plant-based meat products. To streamline the dataset, we therefore propose to employ only one cooking method for all three types of products, which would reduce the number of classes from fourteen to eight.

3.3 Zero Shot inference on Foundational Models

Classifying plant-based meat patties presents numerous challenges. Firstly, multiple types of plant-based meat patties exist, posing difficulties in the classification process. Secondly, given their novelty, acquiring a substantial sample size for training can be challenging. Additionally, there may be inconsistencies in the products' texture, flavor, and nutrition. To address these challenges, we propose leveraging state-of-the-art models such as CLIP[2] and BLIP[3], which have produced exceptional results in similar tasks. In fact, these models have achieved remarkable accuracy in the realm of image classification, demonstrating a significant increase from the prior state-of-the-art of 12% to an impresive 76% for zero-shot ImageNet classification.

3.4 Quantifying Texture as a way to classify Image

Texture plays a crucial role in determining the appearance and tactile experience of meat alternatives derived from plants. Although the RGB color space is commonly employed to describe color variations, it may not adequately capture the subtle nuances of texture. To enhance classification accuracy when it comes to texture, researchers have turned to the CIE-Lab color space[4], which is based on human perception of color. In the context of texture-based classification for plant-based meat products, one statistic that proves useful is the Kullback-Leibler (KL) divergence[5]. This measure serves to quantify the disparity between a given probability distribution and a reference probability distribution. In our case, it can be utilized to compare the texture of a plant-based meat product with a reference distribution of textures. The Histogram of Gradients (HoG)[6] algorithm can also be employed in the classification of plant-based meat product images based on texture. This method involves texture feature extraction utilizing HoG and the use of support vector machine (SVM) for image classification. HoG has demonstrated effectiveness in capturing critical texture information, thereby resulting in precise product classification. This technique counts gradient orientation occurrences in a localized image area, facilitating texture-related characteristic detection. HoG technique is different from other feature descriptors, such as edge orientation histograms and shape contexts because it is specially designed to recognize and represent texture information.

References

[1] G. Vu, H. Zhou, and D. J. McClements, "Impact of cooking method on properties of beef and plant-based burgers: Appearance, texture, thermal properties, and shrinkage," *Journal of Agriculture and Food Research*, vol. 9, p. 100355, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666154322000886

[2] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever, "Learning transferable visual models from natural language supervision," 2021.

[3] J. Li, D. Li, C. Xiong, and S. Hoi, "Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation," 2022.

[4] J. Martínez Sandoval, E. Sandoval, M. Martinez-Rosas, and M. Miranda, *Color Analysis and Image Processing Applied in Agriculture*, 01 2018, pp. 61–78.

[5] S. Fekri-Ershad and F. Tajeripour, "Impulse-Noise Resistant Color-Texture Classification Approach Using Hybrid Color Local Binary Patterns and Kullback–Leibler Divergence," *The Computer Journal*, vol. 60, no. 11, pp. 1633–1648, 04 2017. [Online]. Available: https://doi.org/10.1093/comjnl/bxx033

[6] H. Demir, "Classification of texture images based on the histogram of oriented gradients using support vector machines," *Electrica*, vol. 18, pp. 90–94, 2018. [Online]. Available: https://electricajournal.org/Content/files/sayilar/28/90-94.pdf