Vegan Ingredient Images Classification using Deep Learning and Transfer Learning

MSc Research Project
Data Analytics

Ketaki Dabekar
Student ID: X22149619

School of Computing
National College of Ireland

Supervisor: Professor Anant Aloka.

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Ketaki Dabekar
Student ID:	X22149619
Programme:	Data Analytics
Year:	2023 - 24
Module:	MSc Research Project
Supervisor:	Professor Anant Aloka.
Submission Due Date:	31/01/2024
Project Title:	Vegan Ingredient Images Classification using Deep Learning and Transfer Learning
Word Count:	7998
Page Count:	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Ketaki Dabekar
Date:	31st Jan 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Vegan Ingredient Images Classification using Deep Learning and Transfer Learning

Ketaki Dabekar
X22149619

Abstract

Nowadays, the rising popularity of veganism and plant-based diets is growing fast. Veganism itself is a growing trend with a lot of business potential. Many people are adopting vegan food, so in the market, new ingredients are being introduced. Many people don't know much about vegan ingredients due to a lack of knowledge about it. The modern food industry offers a diverse range of products. It is difficult to identify vegan ingredients with complex formulations. There is a growing need for advanced technologies that will assist every person in identifying vegan-friendly ingredients in food products. The application of deep learning and transfer learning shows remarkable success in image classification tasks. In this research project, deep learning, and transfer learning techniques will be used to classify vegan ingredient images. There are seven different types of transfer learning techniques available, but in this research, ResNet50, EfficientNet B0, and InceptionV3 are conducted and compared by their testing accuracy, so that an effective and efficient automated system will be built in the future. There are 95 classification classes available within the dataset. A diverse range of images has been collected to train the model well. The dataset is a rich collection of coloured images and textures also contains compositions that are commonly found in vegan food. There are a few classes in the dataset that are imbalanced, so tackle this issue by implementing the data-augmentation technique. According to performance, Inception V3 and ResNet 50 are the top 2 models. The testing accuracy of the Inception V3 model is approximately 89.42%, while the ResNet 50 model shows 81.54% accuracy.

1 Introduction

People are choosing vegan diets for various reasons. Many times, decisions are based on a combination of ethics, health, and environment concerns. Many research papers state the health benefits of choosing vegan food. Basically, vegan food and products exclude all animal products. And in recent years, many people have shifted towards vegan food. Its increasing growth has become a public topic for discussion in many modern countries. Because of these ways of buying, cooking, eating, and socializing, the food trend is changing. The motivation behind adapting vegan food could be that the planet is facing some serious environmental problems, such as increasing temperature and climate change. Orlich et al. (2014) To sustain in life, we must know the alternative sustainable option for food consumption.

Conducting research on vegan food is not only desirable from a social or academic perspective but also well suitable for the current moment. It's an increasing trend in food choices, especially in context with the growing popularity of veganism. Kent et al. (2022) Another reason why people are shifting to vegan food is that some animal species are in danger; some species, especially fish species, are diminishing due to sudden environmental changes and sea pollution, so people are being more cautious to protect them.

In many fields nowadays, deep learning is utilized to perform certain tasks, such as image classification or image identification, in object detection. There are many good uses of deep learning models. The special thing about deep learning is that we can apply it to larger datasets. While doing a literature review, I found that if your dataset is labelled and if you are performing prediction tasks, then a neural networking model must be utilized. Huynh et al. (2022) In many papers, it is stated that transfer learning has the capability to learn from data and retrieve features from images. Using this model, it will predict reasonable results about images. In this research, different transfer learning models were implemented, and comparisons done based on their performance. In order to determine which model, predict accurate predictions about images. For comparison, three things were used: 1. testing accuracy 2. Graph: Accuracy and Loss 3. Prediction based on the image. In addition, the dataset was obtained from the Kaggle website. Dataset is labelled.

1.1 Motivation

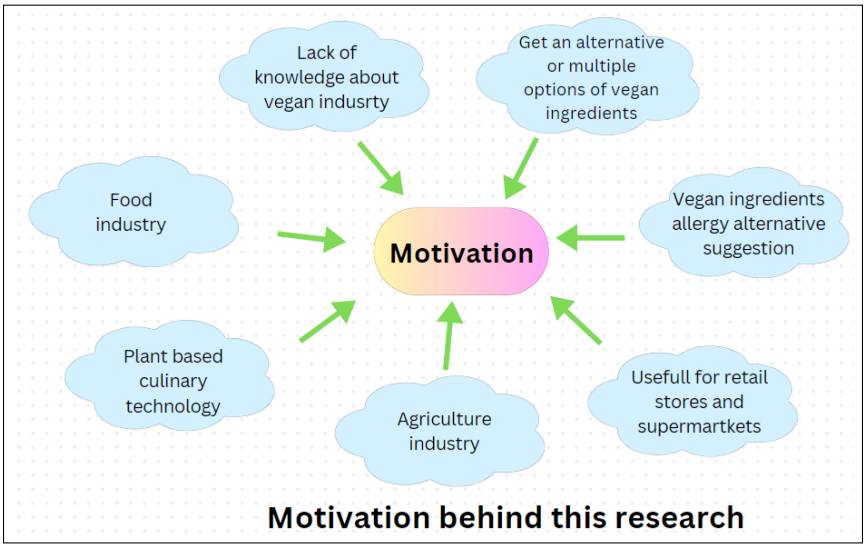

Figure 1: Motivation diagram

Fig. 1 shows the motivation behind this research. And below, we will discuss each topic in detail. If people are eliminating animal meat from their diet to balance all vitamins in the body, they must be aware of alternative plant-based products. Before that, consumers must be aware of the ingredients that will be used in vegan foods. Ardisa et al. (2019) The demand for finding suitable alternatives and developing new cooking skills is that that person needs to have good knowledge about ingredients, like which ingredients are well suited for vegan dishes. If we rely on traditional methods, then those will be time-consuming, and human error must be detected in that, but automated systems offer quick and accurate status of ingredients.

The impact of this study is beneficial for the food industry, especially in plant-based culinary technology. But there is one more industry that will benefit from it: cosmetics, which is beneficial as well as efficient for them and for consumers. In today's world, many cosmetic products come with a vegan tag. Which only includes vegan ingredients in it. And many people purchase those cosmetics and use them in their daily lives. A study shows that from 2014 to 2019, the desired need for vegan and animal-free products has grown considerably; it's increased by 131%. The behavior of consumers in relation to cosmetics is also highlighted. Dos Santos et al. (2023).

The need for classifying and identifying fruits, vegetables, or any ingredients is always important, and it gives a good advantage as well. The agriculture industry or food processing industry, not only this but low-level retail stores and supermarkets where ingredients or this thing are sold out, is efficient and beneficial for them. Because giving them a system like this will perform classification and identification tasks with 1000 elements within a few seconds.Tomar (2020) Mislabeling of ingredients within supermarkets can be harmful for consumers. An appropriate label with vegan content prevents the unintentional consumption of other ingredients.

In general, people who follow veganism still have a lot of barriers due to a lack of understanding about ingredients. Jamnekar et al. (2021) While reading the ingredient list, from product description end user can easily identify whether the product is suitable or not for vegan food. You must have good system for that. This is one of the motivation behind this research. Many times this happens, and the company does not mention that this product can be included in vegan food. Because if a product is vegan, the company usually mentions the tag for the people, and from that, they will get the idea that this product must be included in vegan food. For example, all types of bread may not be included in vegan food. Honey-wheat bread is not included in vegan food because it contains honey as a sweetener. And honey is not included in vegan ingredients, so if you are uncertain about your vegan status, it's advisable to read the list of ingredients.

Some countries, like Indonesia, have a very lower vegetarian index than countries like India. Specifically, among ASEAN countries. Shah and Bhavsar (2021) In the mentioned countries, it is hard for people to gather a information about to vegan food. In those countries, many restaurants are not able to give proper vegan dishes to their consumers. To identify ingredients from dishes or from food, a digital system is needed, which can tell consumers that this is not included in veganism. Nowadays, people are eating and consuming more healthy things, so they try to be innovative and make good dishes using different ingredients. They use different ingredients as well, so the data is becoming more complex day by day. So, identifying appropriate ingredients within the vegan dataset is also a difficult and important task.

While consuming any food allergy is a serious constraint, we need to take it into consideration. Identifying the specific ingredients within food and excluding them from the diet may prevent further causes. Gallagher et al. (2022) For example, a particular person has a soy allergy, and if he or she goes to any restaurant if they serve vegan food, the dish contains soya in it, so using the ingredient identification function, he or she easily gets to know that something allergic is present in the dish. Or it may suggest to that person some alternative beans or lentils, which he or she may tolerate. From a business standpoint, it is critical to understand what consumers admire in vegan food so that offerings can be designed and delivered accordingly.

1.2 Research Question:

"How well do transfer learning models like ResNet 50, Inception V3, and EfficientNet B0 accurately predict a diverse range of vegan ingredients?"

This research major goal is to use a transfer learning model to create an automatic classification application. The study tries to achieve the following goals, which should be considered a subsection of the research question:

• Vegan food contains different types of ingredients, including different types of fruits, vegetables, grains, legumes, nuts, etc. A diverse range means that models should be able to identify not only common items but also less common or nonspecific ingredients.

• Compare the accuracy and performance of the ResNet 50, EfficientNet B0, and Inception V3 models.

1.3 Structure Of Document

This paper is organized according to the following sections: Related works are covered in Section 2. The research methodology mentioned in Section 3 and Section 4 covers design specifications. Section 5 contains an implementation then section 6 contains evaluation of models and a comparison between models. Which gives an idea of which model is performing best on the dataset. Finally, this article is concluded in Section 7. Which contains a conclusion and future work. And last but not least, Section 8 covers all references used in this research paper.

2 Related Work

In this part, we will discuss previous research about deep learning, transfer learning, image classification, and vegan food.

2.1 Vegan food:

In this Gallagher et al. (2022) paper, it shows the types of foods that fall under vegan food. Performed dietary pattern analysis for a vegan diet. Data collected through a survey. And in that 129 vegans voluntarily completed the survey. On the dataset, researchers performed statistical analysis and data screening to identify which food group means ingredients have a higher number of correlations with vegan food. In this paper, researchers gave more information about vegan ingredients. And gave an alternative vegan diet option to the end user. In this paper Romão et al. (2022), they find out about vegan alternative meat products, and this study was conducted in Brazil. Information was collected through vegan products, and then a few steps were performed on the dataset. The initial step was collecting. Then after that step was to repeat the first step, analyze the ingredients, and then search for vegan products. Ingredients were collected from vegan food products. Then, on the collected data, statistical analysis was performed through the IBM SPSS tool. As a result, ingredients were collected from 125 products, and information was gained about ingredients like calorie count and how much energy they were getting from that ingredient.

In this paper Protudjer and Mikkelsen (2020), researchers performed vegan ingredient labeling. Also, they analyzed how these labels are currently used in context with vegan food products. Data was collected from 24 grocery stores, and decided 11 product categories for evaluation. Also, they successfully captured 108 pieces of product information. So, from the study, the end user gets an idea about vegan products. Whether a particular product falls under vegan or not? But at the end, they conclude that the dataset was collected for only a specific region, so it must be a generalized dataset, so there is a future advancement to collect more global data related to vegan products. Also, it contains limited categories, so all ranges must be included, so this is also future advancement.

In a paperZhao et al. (2022), researchers performed calorie measurement on a raw vegan diet using deep learning techniques. The end user needs to capture a photo, and then the system provides a calorie count for that vegan item. The system is able to classify 80 good-quality images for each class using the CNN model. For faster execution, used the RCNN algorithm. In future work, they mentioned about expanding the image dataset and choosing the latest algorithm for testing mixed vegan food images. Vegan dietary patterns have reported a high number of health outcomes in Orlich et al. (2014) studies. The aim of this paper is to study the characteristics and compare the food patterns of several diets. In this study, 68 minor food elements were compared with 17 major food elements. In Cooper et al. (2022) researchers performed text analytics on vegan products that contained vegan ingredients. They took data from social media and performed text mining and artificial intelligence, which will extract data and then process and give information regarding that vegan product, like which ingredients that product contains in it.

2.2 Deep learning and transfer learning:

Deep learning has made good progress in environmental microorganism image classification, according to VijayaKumari et al. (2022). So, at the end, the researcher applied 21 deep learning models. And as a result, researchers found out that using data augmentation with CNN models increases model accuracy, and a few models out of 17 found out to be perfect matches for this task. Such as Inception V3 gave good accuracy. According to Rodrigues et al. (2023), food waste is a major concern in today's world. So we implemented a system that will take and recognize an image with ingredients and give recipe recommendations. For this task, the ResNet 50 model was built on a dataset that contains 36 classes of vegetables and fruits. The model achieved 96% accuracy. In the Lu (2016) paper, researchers focus on the small number of meals classification. Developed one model that recognized the maximum number of foods in a meal. The main aim is to recognize using a transfer learning model that categorizes different food products into their respective categories. The EfficientNet B0 class classified 101 unique food elements from meals with 80% accuracy.

The Chaitanya et al. (2023) paper dataset contains 5822 images, and 10 classification classes are presented for recognition. The RestNet 50 model was constructed to identify images. Then better accuracy was provided by the RestNet 50 model, which was 74% to improve this accuracy. data augmentation technique used and got more than 90% accuracy. In Wang et al. (2019) paper, a model was developed to identify all types of food images as well as their nutritional value, assisting people in maintaining a balanced diet. The Inception V3 CNN model is built on a dataset. Because of the Inception V3 model, the learning process increases. Also, to avoid overfitting, data augmentation is used. Then the extracted information is passed to web crawlers for information retrieval purposes. The model scored 97% for 20 classification classes. According to Jiang (2019) traditional image classification is usually done by extracting low- or mid-level features from images, and then a trainable classifier is used for labeling. But in recent years, feature extraction and classification have both been done together in one network, which is the CNN model. The aim of this study is to classify medical images and give diagnoses based on the images. Classification was done on liver and tumor images. Used the fine-tuning method to increase model accuracy. In Lai (2020) paper, researchers developed food recognition that used CNN and ResNet 50, used 3 different datasets, and applied these 2 algorithms to those 3 datasets. Also performed fine tuning on the pretrained model to achieve good accuracy.

In Rajayogi et al. (2019) paper, researchers tested different types of models against the same dataset for classification problems. Every model was trained from scratch and then applied to four transfer learning models, such as ResNet 50, Inception V3. The good accuracy provided by the Inception V3 model resulted in a total accuracy score of 61.4%. According to Zahisham et al. (2020), it's difficult to train a deep learning model if you don't have lots of datasets available. But using it with a transfer learning model is always a good idea. We can train the transfer model well on task A with a lot of appropriately labeled training datasets. It generalizes well on task A if we shift task A to B, and B doesn't have much data. Using task A, it well predicts the result. So, using this, we can increase the training phase and improve the deep learning model on a small dataset as well. The dataset used in this task for classification was a custom dataset, and the EfficientNet B0 model got 95% classification accuracy. In Cooper et al. (2022), a transfer learning model is used in the X-ray classification task. The image could be classified into three classes, namely normal, pneumonia, and COVID-19. The inspiration behind this research was the delayed availability of the results. Dataset having 1560 chest images. The model used in this task was ResNet 50, which got 97% overall accuracy.

In theÜmit Atila, Uçar, Akyol and Uçar (2021) paper, the EffiecientNet model is used for plant leaf disease classification. Instead of using traditional machine learning algorithms for feature extraction, the EffiecientNet model has been used for extracting features from images. For classification task 39, a number of classes were used. The experiment was carried out on an original and augmented dataset. The accuracy achieved by the EffiecientNet model was around 99.91% on the original dataset, and on the augmented dataset, the accuracy achieved was 99.97%, which is close to 1. In Godishala et al. (2022) paper, performed image classification for breast cancer using deep learning and used an image data generator as a data augmentation technique. And using a pretrained model got better accuracy than other models. In Zhang et al. (2023) paper, the goal of the study was to see how well the EfficientNet family of models classified skin cancer images versus the ResNet architecture. From implementation and evaluation, researchers conclude that the EfficientNet family performed well on skin cancer images with multiple classification classes.In Huynh et al. (2022), this paper performed food category recognition using CNN, transfer learning, and a semi-supervised model. Utilized ResNet 50 and the Inception V3 model to extract features from images.

Based on a literature review, it is evident that previous research in the vegan food industry has primarily relied on manual methods and surveys for ingredient identification, with limited application of deep learning models like CNNs. Major papers which used in this research are relied on traditional methods for classification.Also what difficulty and complexity they faced while capturing images already discussed in literature review above. So this research is helping or fill up gap between traditional and manual ways to automate the system.However, the results from image classification studies show that transfer learning is a highly effective approach. Considering the reviewed literature, I have concluded that leveraging transfer learning models, specifically Inception V3, EfficientNet B0, and ResNet 50, is the optimal strategy to address the research question. These models, as demonstrated in various domains such as environmental microorganism image classification, food waste reduction, and medical image classification, consistently exhibit exceptional accuracy.

3 Methodology

To complete this study, I made use of a framework named as a CRISP-DM. There are six phases, or, we can say, six steps, this framework contains. All steps are covered in Fig. 2. And this architecture is used in many research fields.

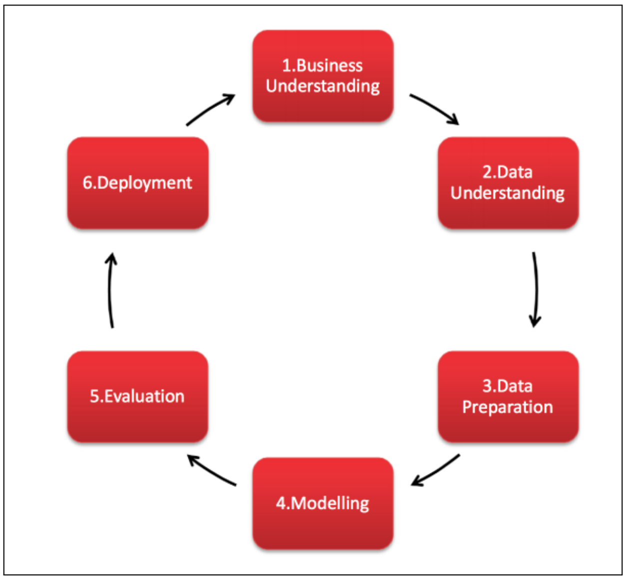

Figure 2: CRISP-DM Framework

3.1 Dataset:

Dataset is an important aspect because of it, we can use data to model. And the selection of the correct dataset is a crucial thing. Whatever data we are feeding to the model based on that model will predict the result. So, it should be aligned with the task. After reviewing a literature review and various papers, I found one dataset. The dataset is available on Kaggle. The dataset that is used to tackle this research question is labelled.The dataset link.[1]

The total number of images in the dataset is 5577. And there are 95 classes in the dataset. Each class shows more than 60 images. The extension of images were JPEG,PNG,JPG within dataset. And size of each image was 224 by 224 applied while pre-processing dataset. Three color channel applied on image which was RGB mean red, green,blue. And resolution of each image set to 800 by 600. Fig. 3 shows the total number of classes and images present within the dataset. also below image shows the

[1]Kaggle website: `https://www.kaggle.com/datasets/ketakidabekar/vegan-ingredients`

```
Total Number of Classes: 95
Total Number of Images: 5577
```

```
['Almonds', 'Apples', 'Avocado', 'Bananas', 'Barley', 'Basil', 'Bay Leaf', 'beetroot', 'Black Beans', 'Black Lentils (Beluga Le
ntils)', 'Black-Eyed Peas', 'Brazil Nuts', 'broccoli', 'Brown Lentils', 'Brussel sprouts', 'Buckwheat', 'cabbage', 'capsicum',
'Carrots', 'Cashews', 'cauliflower', 'Cherries', 'ChiaSeeds', 'Chickpeas', 'Cinnamon', 'Clove', 'Coriander', 'corn', 'courgette
s', 'cucumber', 'cummin seeds', 'dried seaweeds', 'Drum Stick', 'Edamame', 'eggplant', 'Flaxseeds', 'garlic', 'Ginger', 'grapef
ruit', 'Grapes', 'green beans', 'Green Lentils', 'Hazelnuts', 'jackfruit', 'Kale', 'Kidney Beans', 'Kiwi', 'leeks', 'lemon', 'L
emongrass', 'Lima Beans', 'Macadamia Nuts', 'Mangoes', 'Mung Beans', 'Muster Seeds', 'Nutmeg', 'Oats', 'onions', 'Oranges', 'Or
egano', 'Parsley', 'parsnips', 'Peaches', 'Peanuts', 'Pears', 'peas', 'Pecans', 'Pine Nuts', 'Pineapples', 'Pinto Beans', 'Pist
achios', 'Plums', 'Pomegranate', 'potato', 'Pumpkin seeds', 'Quinoa', 'raddish', 'Red Lentils', 'Rice', 'Rosemary', 'Sage', 'Se
same seeds', 'soya', 'spinach', 'strawberries', 'Sunflower seed', 'sweetpotato', 'Tempeh', 'Thyme', 'Tofu', 'tomato', 'Turmeri
c', 'Walnuts', 'Watermelon', 'Yellow Lentils']
```

Figure 3: Dataset decription

names of all classes that will be used for classification, which includes all possible vegan ingredients.

3.2 Data Pre-processing and Data Augmentation:

In the data processing phase, two techniques have been utilized before applying the model on the dataset. I have mentioned a workflow diagram for better understanding about both terms.

3.2.1 Preprocess_input:

Utilized the preprocess_input function, which is provided by the Keras application module. This function is used with transfer learning. Basically, this function preprocesses input images before feeding them to any model. This function does perform normalization of the pixel values of images. Means convert pixel values between 0 and 1. It also performed mean subtraction functionality on each pixel value of the image. It includes channel-wise color normalization. It subtracted the mean values from each color channel. It scaled pixel values by some factor with the digital range. Used original dataset on Inception V3 model so in that just performed basic pre-processing technique such as resize and decode. Then before splitting used label encoder. In dataset class names are in string format but for model wanted to be in integer format so applied using labe encoder for that.

3.2.2 Data augmentation:

It was done by using the ImageDataGenerator library. Originally, ImageDataGenerator was a Keras library before being integrated into TensorFlow. If there are not many images within the dataset and a few classes within the dataset are imbalanced, then we can make use of data augmentation. Using this technique, we can create more images by simply applying rotation, shifting, and flipping operations. Lai (2020) But here in this research, instead of actual data augmentation code, we used the ImageDataGenerator library, which will perform data augmentation operations on datasets and prepare datasets according to model input requirements. And this function is often used with the transfer learning model. Godishala et al. (2022) At the time of model compilation, images are generated and passed to different transfer learning models.

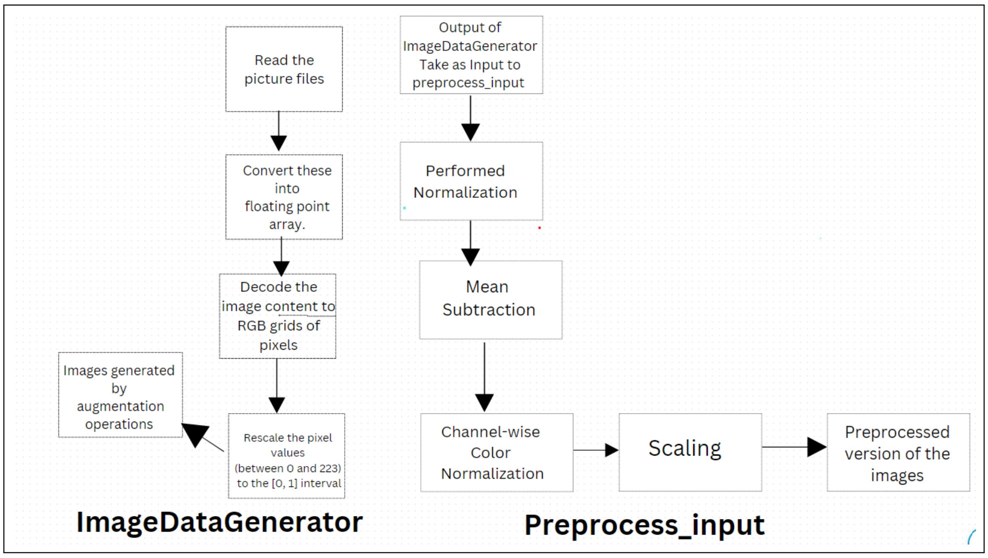

Figure 4: Implemenataion diagram

Below, Fig. 4 shows the workflow diagram of the ImageDataGenerator and Preprocess_input functions.

3.3 Modelling:

After reviewing lots of research papers, I decided to implement three models for this research. As the title suggests, its classification topic has been utilized in many papers by researchers who have utilized a transfer learning model with CNN. Below, Fig. 6 shows the implementation plan. In that, we can see I implemented ResNet 50, EfficientNet B0, and Inception V3 models.

Implementing different models allows researchers to do comparative analysis based on the results they are getting from the models. Various models have different design architectures and capabilities. And based on this, we will evaluate which model will work best on the given dataset. We will also get a capable architecture that will work best for the given classification problem. If we change the architecture of the model, then the results will also differ in the classification problem. Below is the reason behind each model selection:

3.3.1 ResNet 50:

ResNet 50 is also known as deep architecture. And it's well suited for complex image recognition tasks. In this model, we can add as many layers as we want to make the complex image recognition task easier. This model enables the vanishing gradient problem.

3.3.2 EffiecienNet B0:

This model will be used in situations where we have fewer computational resources available. This model will achieve good performance with a small number of parameters. If we wanted to note the difference between model complexity and accuracy, then this model is a perfect example.

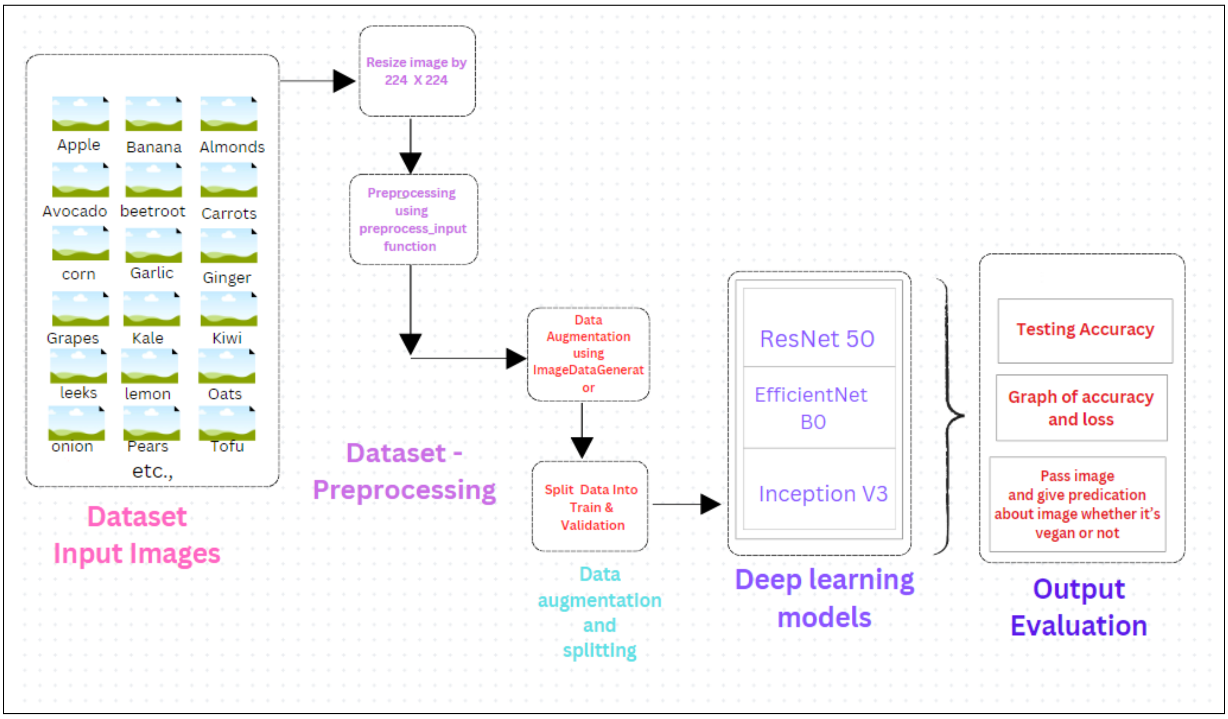

Figure 5: Flow chart of Implementation

3.3.3 Inception V3:

This model contains one special module named "Inception".It will process features at multiple scales within the same layer. This model is good at capturing a wide range of features. It will do feature extraction with minimal computational requirements.

3.4 Evaluation:

Before applying models, I decided to do a comparison between three models. Based on three criteria, I can prove that one model is performing better than another. Also, I get the final model that will perform best on the available dataset. Fig. 6 shows output evaluation factors. Below is a list of evaluation factors in detail:

3.4.1 Testing Accuracy::

When I applied the first model on the dataset, I checked the testing accuracy. In many cases, the model performed well on the training dataset, but it did not generalize well on the validation dataset. As a result, ensuring the accuracy of the testing dataset is critical.

3.4.2 Graph of Accuracy and Loss:

The accuracy graph shows correctly classified instances based on the total number of instances. The accuracy graph should be in increasing order. And it will eventually increase according to epoch values. The loss graph should be in decreasing order. While training the model, the goal is to minimize the loss value. So, it will indicate that the model is making accurate predictions.

3.4.3 Pass the image and give a prediction of that image:

In this evaluation criteria, I am passing an image to a model that I have already trained and tested on a dataset. It will pre-process the image and convert it to the expected model format, and then the model will predict the array of predictions per class. And at the end, it will give us an actual prediction about whether the image is vegan or not.

4 Design Specification

This module shows the actual implementation that was implemented to tackle the research question. Let's discuss in detail the design of each model.

4.1 Transfer learning:

If we compare this research task with the current situation, It may happen that it will generate lots of data per minute. Processing high-quality images for the model is a difficult task, and for that, we may require more computation resources. So, to tackle this transfer learning technique, it is introduced. Transfer learning means transferring information from one place to another, but in a more technical way, it will transfer information after training the dataset on the model. So, the model will acquire features from the dataset and utilize that acquired information in model training for further models. During the literature review, one thing I noticed was that using this technique, model accuracy was also increasing.

4.1.1 ImageNet dataset:

This is a commonly used dataset while training a deep learning model, which almost covers and contains different ranges of object categories. Using the pre-training phase on this dataset, models will gather the ability to recognize and extract features from various sets of images, so if we apply the pretrained model on the actual dataset, it will enhance its performance. Also, using the pretraining phase, we will train the model for the image classification task. As mentioned above, all three models were trained on the ImageNet dataset.

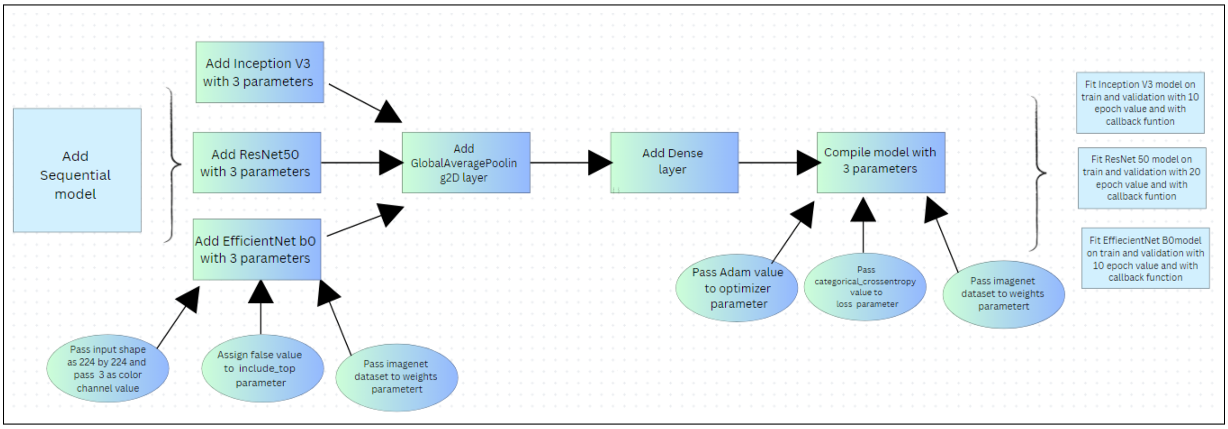

Figure 6: Architecture of tranfer learning model

4.1.2 Architecture :

Above Fig.6. shows the architecture diagram for all models.Here we will discuss the architecture and which layer model was implemented during model building. Used the sequential model first, which was provided by the Keras library. So basically, what this model does is flow data sequentially through layers. And this model is widely used in many deep learning tasks. After initialization, we can add layers one by one in a sequential manner. The top layer is removed in this architecture because it will do classification on the ImageNet dataset, which contains 1000 classification classes, but in this case, I already have 95 classification classes defined in the dataset. This is the reason I have removed the top layer from the architecture.

After loading the pretrained weights, which are gathered from the ImageNet dataset, the model uses GlobalAveragePooling 2D, dense layers. Then the Global Average Pooling 2D layer was applied, which captured spatial hierarchies from the input. This layer will also protect the model from overfitting by reducing the number of parameters. The dense layer will take the final number of classification classes; in my case, I am using 95 classification classes, and then, using the Softmax activation function, it will convert the model output into probabilities. Then we used the Adam optimizer with a 1e-4 learning rate value. And we passed categorical crossentropy as a loss parameter because, here in this study, we are performing multi-class classification. To improve the learning rate The ReduceLROnPlateau callback function is used to adjust the leaning rate if validation loss is not improving. That means it performed fine tuning on the leaning rate based on validation loss. It also monitors the val_loss score. It also adapts the learning rate dynamically. While building this model, I referred Zahisham et al. (2020) this research paper. They almost use a similar structure with minor changes. Below, Fig. 1 shows the design diagram of all models. Basically, I have applied similar building structures to all models.

5 Implementation

We will discuss the overall working of the proposed solution under this section.Because we already know the motivation behind this topic then after literature review I finalised technologies. Also gave previous work that was done on this research topic.

As I was executing my whole code in the jupyter notebook. Then extracted dataset from kaggle website in jupyter notebook directory. In my dataset there are 95 number of folders and in each folder, there are more than 60 images. Then read this image data performed preprocessing and augmentation on those images. For that as I mentioned in methodology I used 2 function for prerprocessing and augmentation.after that splitted the training dataset into training and validation in an 80:20 ratio.

After that,I loaded ResNet 50, EfficientNet B0 and Inception V3 pre-trained architectures. Then I made this architecture sequential and updated the model by adding some layers which I have defined it in detailed in section 4.1.1. To compile this model, I utilized the Adam optimizer and loss as categorical cross-entropy. I passed learning rate parameter and passed value as 1e-4. Then, using the Training and Validation data, I trained all models between 10 and 30 epochs. For verification purpose whether the model was developed appropriately or not, I first examined the model's accuracy for traning and validation dataset. Then checked loss and AUC curve.

All of my code was written in Python. I used kaggle json file to load data and then used zipfile library to extract data from zip file.TensorFlow's Keras library was used to load models.I used the SKlearn, matplotlib and seaborn libraries to create the visualization.

6 Evaluation

In this section, evaluating model that was built based on research questions. I just wanted to check if the model was doing well or not using the dataset. Evaluation will be done on dataset. Here we will discuss how and which model is performing best in the ingredient classification task. In the methodology section above, I have already defined three different evaluation criteria. Accuracy is one of them, and it is the most important evaluation criteria among all. In many research papers, comparisons will be done based on the same evaluation criteria. Let's start with the first model, which is ResNet 50:

6.1 Evaluating the ResNet 50 model:

The accuracy of the training and validation datasets seems to be increasing with each epoch value. Which indicates that the model is learning well from the training dataset. Actually, this is the main motive behind the transfer learning model; it will train the model well in the training phase. Accuracy is increasing for the validation dataset, but it is not showing model overfitting due to the usage of the ReduceLROnPlateau callback function. This function prevents the model from overfitting and improves the accuracy score as well. This model was trained on a 20-epoch value. The model accuracy increased significantly between epoch values 1 and 6. After six epochs, a learning rate reduction was implemented, which resulted in further fine-tuning. The model learns and achieves 81.54% accuracy on the validation dataset by the end of training. Below, Fig.7. shows the accuracy for both datasets.

```
Epoch 16/20
141/141 [==============================] - 1427s 10s/step - loss: 0.0191 - accuracy: 0.9984 - val_loss: 0.8654 - val_accuracy:
0.8072 - lr: 1.6000e-07
Epoch 17/20
141/141 [==============================] - 1599s 11s/step - loss: 0.0198 - accuracy: 0.9987 - val_loss: 0.8291 - val_accuracy:
0.8183 - lr: 1.6000e-07
Epoch 18/20
141/141 [==============================] - 1479s 10s/step - loss: 0.0206 - accuracy: 0.9984 - val_loss: 0.8279 - val_accuracy:
0.8127 - lr: 1.0000e-07
Epoch 19/20
141/141 [==============================] - 1443s 10s/step - loss: 0.0190 - accuracy: 0.9980 - val_loss: 0.8441 - val_accuracy:
0.8090 - lr: 1.0000e-07
Epoch 20/20
141/141 [==============================] - 1386s 10s/step - loss: 0.0183 - accuracy: 0.9984 - val_loss: 0.8315 - val_accuracy:
0.8137 - lr: 1.0000e-07
```

```
34/34 [==============================] - 83s 2s/step - loss: 0.8418 - accuracy: 0.8155
Test accuracy: 0.8154981732368469
```

Figure 7: Accuracy of train and validation

The above graph Fig.8. shows for both datasets, as we can see. The left side graph shows the accuracy graph, and the right side graph shows the loss and loss values within the graph. The accuracy graph should be in increasing order for both data sets, which shows a good sign about model evaluation. And here, in this case, the accuracy curve for both datasets shows an increasing trend. For the loss graph, it should be in decreasing order, which shows the model is a perfect fit. Both curves should be in decreasing order, and here for this model, it is evaluating correctly.

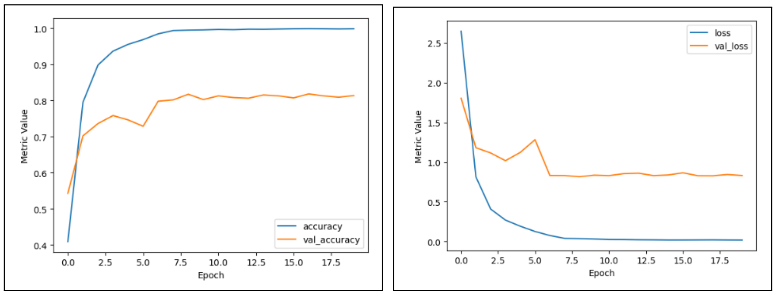

Figure 8: AUC and Loss Graph

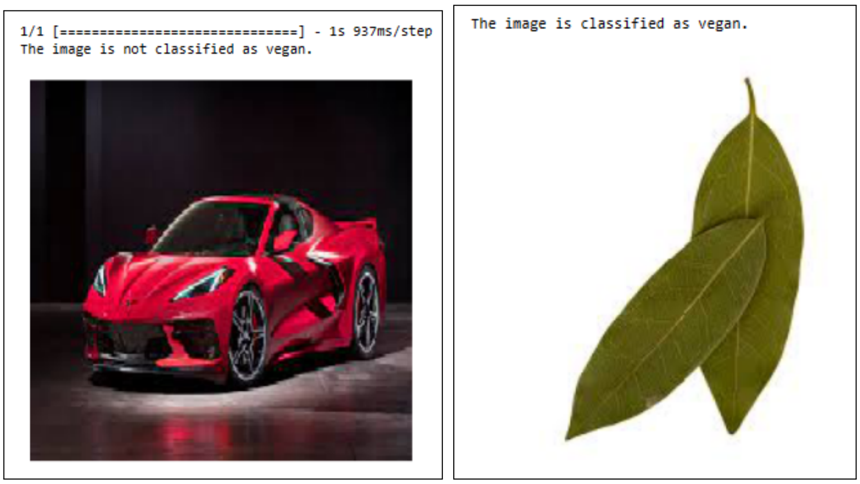

Figure 9: Image Predicatio.

Then we performed a sample image prediction Fig.9.shows that. So, this evaluation proves whether the given image is vegan or not. It is good practice to test the model manually. Also, from manual testing, we get to know how model work, which will confirm that in the future, model will perform well in real-world situations. Below Fig. shows a prediction about the image: First, I pass the car as an image, then the model gave me the output as the image come under vegan ingredients. And in the second image, I passed an image of actual vegan ingredients at that time. The model gave me the correct prediction about the image, which is that the image comes under vegan ingredients.

6.2 Evaluating the EfficientNet B0 model :

The output, which we are discussing below, indicates the training progress of the EfficientNet B0 model over 10 epoch values. And we can see that the model is performing well on both datasets. As the model started training with a high loss (4.1840) value and low accuracy (11.95%) on the first epoch, Fig. 10 shows values per epoch. As the epoch value increases, the scores for both improve significantly. The value for loss decreases and the accuracy increases, which is expected during model training. Exactly this is happening with the AUC and loss graph. Which can be shown in Fig. We can trace the performance of the validation set during model training. The loss for validation decreases from 3.487 to 0.7746, and the accuracy increases from 26.29% to 81.64%. This result on the validation set indicates that the model is generalizing well on the validation dataset. As I mentioned in the design specification of the EfficientNet architecture, because of the ReduceLROnPlateau technique, model coverage is more effective. After training completion, the model was evaluated on the validation dataset and reported 80.81 percent accuracy for the validation dataset. This indicates how the model will perform on new or unseen data.

The above graph(Fig.11.) shows for both datasets, as we can see. The left-side graph shows the accuracy graph, and the right-side graph shows the loss graph. The accuracy graph should be in increasing order for both datasets, which shows good signs about model evaluation. For the loss graph, it should be in decreasing order, which shows the model is a perfect fit. Both curves should be in decreasing order, and here for this model, it is evaluating correctly.

Then I passed the image to this model for prediction. Also, I wanted to make predictions on the process image. In the below Fig., we can see two predication cases, which are making a prediction about whether the image is vegan or not, and, in both cases, a model evaluating the correct answer. First, I passed the car image because it shows that it does not come under vegan, and in the second case, I passed the bay leaf image, which comes under vegan ingredients, for which it evaluates the correct output. shows in Fig.12.

6.3 Evaluating the Inception V3 model:

The accuracy of the training and validation datasets seems to be increasing with each epoch value.shows in Fig.13. Which indicates that the model is learning well from the training dataset. Actually, this is the main motive behind the transfer learning model; it will train the model well in the training phase. Accuracy is increasing for the validation dataset, but it is not showing model overfitting due to the usage of the ReduceLROnPlateau callback function. This function prevents the model from overfitting and improves

```
Epoch 6/10
141/141 [==============================] - 573s 4s/step - loss: 0.6773 - accuracy: 0.8565 - val_loss: 1.0033 - val_accuracy: 0.
7574 - lr: 1.0000e-04
Epoch 7/10
141/141 [==============================] - 622s 4s/step - loss: 0.5200 - accuracy: 0.8963 - val_loss: 0.9356 - val_accuracy: 0.
7703 - lr: 1.0000e-04
Epoch 8/10
141/141 [==============================] - 624s 4s/step - loss: 0.4159 - accuracy: 0.9141 - val_loss: 0.8540 - val_accuracy: 0.
7906 - lr: 1.0000e-04
Epoch 9/10
141/141 [==============================] - 576s 4s/step - loss: 0.3168 - accuracy: 0.9419 - val_loss: 0.8243 - val_accuracy: 0.
7897 - lr: 1.0000e-04
Epoch 10/10
141/141 [==============================] - 571s 4s/step - loss: 0.2586 - accuracy: 0.9517 - val_loss: 0.7746 - val_accuracy: 0.
8164 - lr: 1.0000e-04
```

```
34/34 [==============================] - 44s 1s/step - loss: 0.7880 - accuracy: 0.8081
EfficientNetB0 Test accuracy: 0.8081181049346924
```

Figure 10: Accuracy of Tranining and Validation

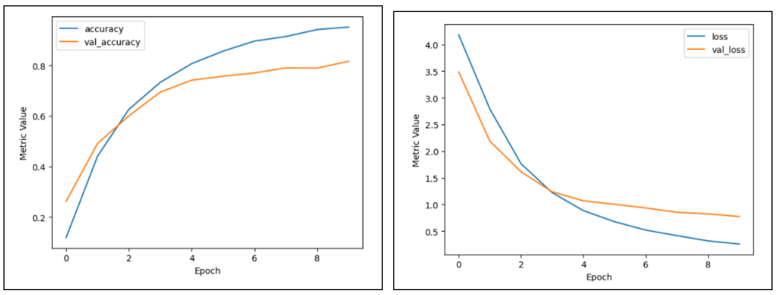

Figure 11: Graph of AUC and loss

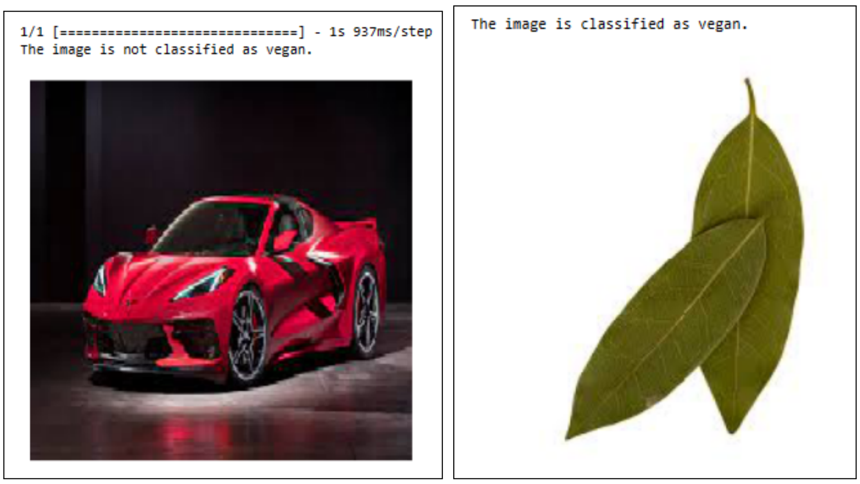

Figure 12: This is a caption

the accuracy score as well. This model was trained on a 20-epoch value. The model accuracy increased significantly between epoch values 1 and 6. After six epochs, a learning rate reduction was implemented, which resulted in further fine-tuning. The model learns and achieves 81.54% accuracy on the validation dataset by the end of training. Below, Fig. shows the accuracy for both datasets.

Implemented Inception V3 model on cases in first used augmented dataset and in second senario used original dataset. Augmented model gave 99% accuracy on training and validation dataset gave 79% accuracy. With original dataset got 99% and 89% accuracy on training and on validation dataset. It is observed that both model performing well on training dataset.

```
Epoch 7/10
141/141 [==============================] - 924s 7s/step - loss: 0.1679 - accuracy: 0.9602 - val_loss: 1.0557 - val_accuracy: 0.
7601 - lr: 1.0000e-04
Epoch 8/10
141/141 [==============================] - 883s 6s/step - loss: 0.1108 - accuracy: 0.9777 - val_loss: 0.9113 - val_accuracy: 0.
7878 - lr: 2.0000e-05
Epoch 9/10
141/141 [==============================] - 854s 6s/step - loss: 0.0756 - accuracy: 0.9855 - val_loss: 0.9405 - val_accuracy: 0.
7832 - lr: 2.0000e-05
Epoch 10/10
141/141 [==============================] - 804s 6s/step - loss: 0.0676 - accuracy: 0.9900 - val_loss: 0.9094 - val_accuracy: 0.
7906 - lr: 2.0000e-05
```

Evaluation of Inception V3 model

Validation accuracy

```
: # Evaluate the model on the test set
eval_result = model_inception.evaluate(validation_generator)
print("InceptionV3 Test accuracy:", eval_result[1])

34/34 [==============================] - 47s 1s/step - loss: 0.8781 - accuracy: 0.7943
InceptionV3 Test accuracy: 0.794280469417572
```

Figure 13: Accuracy of Training and Validation - Augmented model

```
Epoch 8/10
140/140 [==============================] - 1027s 7s/step - loss: 0.0403 - accuracy: 0.9933 - val_loss: 0.4456 - val_accuracy:
0.8978
Epoch 9/10
140/140 [==============================] - 1004s 7s/step - loss: 0.0487 - accuracy: 0.9919 - val_loss: 0.4527 - val_accuracy:
0.8943
Epoch 10/10
140/140 [==============================] - 1054s 7s/step - loss: 0.0379 - accuracy: 0.9937 - val_loss: 0.4361 - val_accuracy:
0.8943
```

Evaluation of Inception V3 model

Validation accuracy

```
7]: # Evaluate the model on the test set
eval_result_inception1 = model_inception1.evaluate(validation_dataset)
print("EfficientNetB0 Test accuracy:", eval_result_efficientnet[1])

35/35 [==============================] - 36s 1s/step - loss: 0.4361 - accuracy: 0.8943
EfficientNetB0 Test accuracy: 0.8942652344703674
```

Figure 14: Accuracy of Training and Validation - Oridinal dataset model

We can see Fig.15. AUC graph both showing almost similar pattern same goes with loss graph as well. There is no discrepancies in all graphs.

Then I performed a sample image prediction. which is shown in Fig.16. So, this evaluation proves whether the given image is vegan or not.I passed image to both datsets

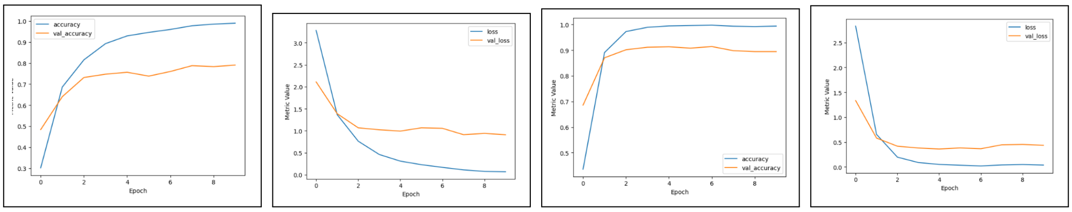

Figure 15: Graph of AUC and Loss

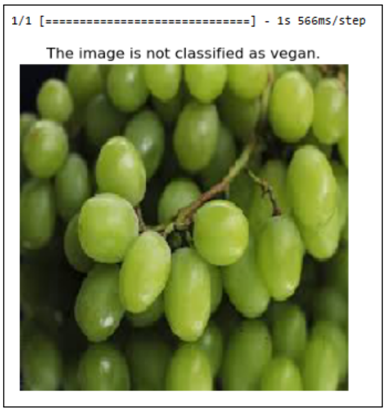

Figure 16: Image Predication

but augmented model classifying really well even if the model don't have good accuracy but In above Fig.16. we can see I passed grape image to original implemented inception V3 model but for few images it is showing correct classification but many times its is not classifying correctly. I changed epoch values as well for both images but 10 it is giving me good accuracy and better classification on images but this case is happened in opposite direction with second model.

6.4 Discussion

In this research project, three transfer learning models—ResNet 50, EfficientNet B0, and Inception V3—were employed for the classification of vegan ingredients. Shows in Fig 17.The training process involved two variation augmenting the dataset using the ImageDataGenerator and applied models on original dataset,then split it into training and validation sets, and utilize pre-trained weights from the ImageNet dataset.

ResNet 50: Implemented as the first model, ResNet 50 demonstrated robust performance with a training accuracy of 99.84% and a validation accuracy of 95.17%. This model consists of residual blocks, enabling effective training of deep networks.Fine-tuning was performed on ResNet 50 to adapt the model to the specific characteristics of the dataset. This step can be crucial for achieving optimal performance on a targeted task.

EfficientNet B0: The addition of EfficientNet B0 to the study provided an alternative architecture. It achieved a training accuracy of 99% and a validation accuracy of 81.54%.

Inception V3: Employed as the third model, Inception V3 exhibited a training accuracy of 99% and a validation accuracy of 80.81%. Inception V3 utilizes inception modules, allowing the network to capture information at different scales.

The ResNet 50 model was trained for 20 epochs for Inception V3 and EfficientNet

Evaluation	ResNet 50	EffiecientNet B0	Inception V3	Inception V3
Training accuracy	99.84 %	95.17 %	99.37 %	99 %
Validation accuracy	81.54 %	80.81 %	89.42 %	79.42 %
Epoch value	20	10	10	10
Data augmentation Applied	Yes	Yes	No	Yes
Perform Fine -tunning	Yes	No	No	No

Figure 17: Evaluation of four experiments

B0, which was trained for 10 epochs. So I changed the epoch value for all models and for ResNet 50 model work best on 20 epoch value like it gave me almost all correct predication on image. We can apply more than 20 epoch value as well but from Fig.7. It's conclude that from 12 Th epoch onwards accuracy is not improving significantly. So I kept 20 as final value. And some what same thing happend with Inception V3 model I gave 10 then changed 25 but from 8 epoch it showed me the same accuracy value but on augmented Inception V3 it is not showing me good accuracy as compared with original dataset but it is giving me better predication's. Also in Zhao et al. (2022) paper showing similar results so we can conclude that data augmentation making big difference while performing classification task. And at the time of predication if we are using augumneted data the model will learn from all angles of image so predication will be easy in that case. Huynh et al. (2022)The evaluation metrics, including accuracy ,AUC and loss curve provided a comprehensive assessment of the models' performance.

7 Conclusion and Future Work

In this study, the motivation behind this research lies in addressing the lack of knowledge about vegan ingredients and providing practical solutions for various industries. The proposed models can be beneficial for retail stores, supermarkets, the agriculture industry, and the food industry by offering suggestions and knowledge for vegan ingredients.

We systematically addressed the research question by implementing and evaluating three deep learning models, namely ResNet 50, Inception V3, and EfficientNet B0, for vegan ingredient classification. Our objectives were to assess their performance, identify strengths and weaknesses, and propose a classification system for the future. The implemented models were validated using the accuracy, AUC, and loss graphs. The ResNet 50 design outperformed well than previous research in this area. This concludes that the ResNet 50 model is best suited for classifying vegan ingredient images.

While the current study achieved promising results, there are avenues for future exploration and improvement. We can use a model ensembling what this does by combining the predictions from different models, which can improve overall accuracy and robustness.We can experiment with different hyperparameter configurations that could lead to further performance improvements. And last, expanding the dataset with more diverse images could help the models generalize better to a wider range of vegan ingredients.

References

Ardisa, K., Prabowo, W. and Rustad, S. (2019). Implementation of convolutional neural network method for detecting vegetables as recommendation for vegetarian food recipes, *Engineering and Computer Science* **46**(1): 83–88.

Chaitanya, A., Shetty, J. and Chiplunkar, P. (2023). Food image classification and data extraction using convolutional neural network and web crawlers, *Procedia Computer Science* **218**: 143–152.

Cooper, K., Dedehayir, O., Riverola, C., Harrington, S. and Alpert, E. (2022). Exploring consumer perceptions of the value proposition embedded in vegan food products using text analytics, *Sustainability* **14**(4): 2075.

Dos Santos, R., de Brito Silva, M., da Costa, M. and Batista, K. (2023). Go vegan! digital influence and social media use in the purchase intention of vegan products in the cosmetics industry, *Social Network Analysis and Mining* **13**(1): 49.

Gallagher, C., Hanley, P. and Lane, K. (2022). Pattern analysis of vegan eating reveals healthy and unhealthy patterns within the vegan diet, *Public Health Nutrition* **25**(5): 1310–1320.

Godishala, A. K., Yassin, H., Veena, R. and Lai, D. T. C. (2022). Breast cancer tumor image classification using deep learning image data generator, *2022 7th International Conference on Image, Vision and Computing (ICIVC)* pp. 418–423.

Huynh, T. T. M., Le, T. M., That, L. T., Tran, L. V. and Dao, S. V. T. (2022). A two-stage feature selection approach for fruit recognition using camera images with various machine learning classifiers, *Computers & Electrical Engineering* **10**: 132260–132270.

Jamnekar, R., Jadhao, A. and Keole, R. (2021). Food classification using image processing techniques.

Jiang, M. (2019). Food image classification with convolutional neural networks.
URL: *https://api.semanticscholar.org/CorpusID:215777485*

Kent, G., Kehoe, L., Flynn, A. and Walton, J. (2022). A review of the definitions and nutritional role in the adult diet, *Proceedings of the Nutrition Society* **81**(1): 62–74.

Lai, K. (2020). Convolutional neural network.

Lu, Y. (2016). Food image recognition by using convolutional neural networks (cnns), *ArXiv* **abs/1612.00983**.
URL: *https://arxiv.org/abs/1612.00983*

Orlich, M., Jaceldo-Siegl, K., Sabaté, J., Fan, J., Singh, P. and Fraser, G. (2014). Patterns of food consumption among vegetarians and non-vegetarians.

Protudjer, J. and Mikkelsen, A. (2020). Veganism and paediatric food allergy: Two increasingly prevalent dietary issues that are challenging when co-occurring, *BMC Pediatrics* **20**(1): 341.

Rajayogi, J. R., Manjunath, G. and Shobha, G. (2019). Indian food image classification with transfer learning, pp. 1–4.

Rodrigues, M., Fidalgo, F. and Oliveira, (2023). Recipeis—recipe recommendation system based on recognition of food ingredients, *Applied Sciences* **13**(13): 7880.

Romão, B., Botelho, R., Nakano, E., Raposo, A., Han, H., Vega-Muñoz, A., Ariza-Montes, A. and Zandonadi, R. (2022). Are vegan alternatives to meat products healthy? a study on nutrients and main ingredients of products commercialized in brazil, *Frontiers in Public Health* **10**: 900598.

Shah, B. and Bhavsar, H. (2021). Overview of deep learning in food image classification for dietary assessment system, *Intelligent Systems, Technologies and Applications: Proceedings of Sixth ISTA 2020, India*, Springer Singapore, pp. 265–285.

Tomar, H. (2020). *Multi-class Image Classification of Fruits and Vegetables Using Transfer Learning Techniques*, PhD thesis, Dublin Business School.

VijayaKumari, G., Vutkur, P. and P., V. (2022). Food classification using transfer learning technique, *Global Transitions Proceedings* **3**(1): 225–229.

Wang, W., Liang, D., Chen, Q., Iwamoto, Y., Han, X., Zhang, Q., Hu, H., Lin, L. and Chen, Y. (2019). Medical image classification using deep learning.
URL: *https://api.semanticscholar.org/CorpusID:209064173*

Zahisham, Z., Lee, C. P. and Lim, K. M. (2020). Food recognition with resnet-50, *2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, Kota Kinabalu, Malaysia, pp. 1–5.

Zhang, Y., Deng, L., Zhu, H., Wang, W., Ren, Z., Zhou, Q., Lu, S., Sun, S., Zhu, Z., Gorriz, J. M. and Wang, S. (2023). Deep learning in food category recognition, *Information Fusion* **98**: 101859.

Zhao, P., Li, C., Rahaman, M., Xu, H., Yang, H., Sun, H., Jiang, T. and Grzegorzek, M. (2022). A comparative study of deep learning classification methods on a small environmental microorganism image dataset (emds-6): From convolutional neural networks to visual transformers, *Frontiers in Microbiology* **13**: 792166.
ΩÜmit Atila et al.

Ümit Atila, Uçar, M., Akyol, K. and Uçar, E. (2021). Plant leaf disease classification using efficientnet deep learning model, *Ecological Informatics* **61**: 101182.