Vegetable Classifier Project Report

Your Name

June 2, 2025

Abstract

The Vegetable Classifier project leverages deep learning to classify images of 15 different vegetables with high accuracy. Using a pre-trained MobileNetV2 model fine-tuned with PyTorch, the project achieves a validation accuracy of 99.50% on a dataset of 21,000 images. A user-friendly web interface built with Streamlit allows users to upload images, view predictions, explore nutritional information, and see confidence scores via an interactive bar chart. This report details the dataset, methodology, model performance, features, and setup instructions, along with placeholders for key screenshots.

1 Introduction

The Vegetable Classifier project aims to develop an efficient and user-friendly system for identifying vegetables from images. Built as a distinct alternative to a similar project, it focuses on simplicity, effectiveness, and added value through features like nutritional information. The project uses a deep learning model (MobileNetV2) to classify 15 vegetable classes and integrates a Streamlit-based web app for interactive predictions. Key objectives include achieving high classification accuracy, providing nutritional insights, and ensuring ease of use through a modular codebase.

2 Dataset

The dataset comprises 21,000 images across 15 vegetable classes, organized into training, validation, and test sets:

- **Classes**: Bean, Bitter Gourd, Bottle Gourd, Brinjal, Broccoli, Cabbage, Capsicum, Carrot, Cauliflower, Cucumber, Papaya, Potato, Pumpkin, Radish, Tomato.
- **Total Images**: 21,000 (1,400 per class).
- **Split**:
  - Training: 14,700 images (70%, 980 per class).
  - Validation: 3,150 images (15%, 210 per class).
  - Test: 3,150 images (15%, 210 per class).
- **Image Specs**: 224×224 pixels, .jpg format.
- **Location**: `D:\College Projects\VegetableClassifier\dataset\{train,validation,test}`.

Each subfolder contains class-specific directories (e.g., `train/Carrot/`).

3 Methodology

3.1 Model Architecture

The project uses MobileNetV2, a lightweight convolutional neural network pre-trained on ImageNet. The final classification layer was modified to output 15 classes, and the model was fine-tuned using PyTorch.

3.2 Training Process

The model was trained for 10 epochs with the following setup:

- **Optimizer**: Adam (learning rate = 0.001).
- **Loss Function**: Cross-Entropy Loss.
- **Data Augmentation**: Random horizontal flips and rotations.
- **Batch Size**: 32.

Training was performed on the dataset at `D:\College Projects\VegetableClassifier\dataset`.

3.3 App Development

A web interface was developed using Streamlit, featuring:

- Image upload for predictions.
- Display of predicted vegetable, confidence score, and nutritional info.
- Interactive bar chart showing the top 5 prediction probabilities using Plotly.

4 Model Performance

The model achieved excellent performance over 10 epochs:

- **Training Accuracy**: 98.87% (Epoch 10).
- **Validation Accuracy**: 99.50% (Epoch 10).
- **Training Loss**: 0.0374 (Epoch 10).
- **Validation Loss**: 0.0193 (Epoch 10).

Training logs:

```
Epoch 1/10, Train Loss: 0.2382, Train Acc: 0.9290, Val Loss: 0.1319, Val Acc: 0.9587
Epoch 2/10, Train Loss: 0.1370, Train Acc: 0.9583, Val Loss: 0.0930, Val Acc: 0.9713
...
Epoch 10/10, Train Loss: 0.0374, Train Acc: 0.9887, Val Loss: 0.0193, Val Acc: 0.9950
```

The model shows strong generalization, with validation accuracy slightly higher than training accuracy, indicating no overfitting.

5 Features

The Vegetable Classifier app includes:

- **Classification**: Predicts the vegetable class with high accuracy.
- **Nutritional Info**: Displays calories and vitamins per 100g (e.g., Carrot: 41 kcal, Vitamins A, K).
- **Confidence Scores**: Interactive bar chart showing the top 5 prediction probabilities.
- **User Interface**: Streamlit app with Home, Prediction, and About pages.

6 Setup Instructions

1. **Verify Dataset**: Ensure the dataset is at `D:\College Projects\VegetableClassifier\dataset\`.
2. **Set Up Virtual Environment**:

```
cd D:\College Projects\VegetableClassifier
python -m venv venv
venv\Scripts\activate  # Windows
```

3. **Install Dependencies**:

```
pip install -r requirements.txt
```

4. **Train the Model** (if not already done):

```
python train.py
```

5. **Run the App**:

```
streamlit run app.py
```

   Access at `http://localhost:8501`.

7 Results and Screenshots

7.1 Prediction Example

The app accurately predicts vegetable classes. For example, uploading a carrot image yields: "Prediction: Carrot (Confidence: 92.34%)" with nutritional info and a confidence chart.

7.2 Training Curves

The training curves show the model's performance over 10 epochs, with steady improvements in accuracy and loss reduction.

7.3 Confusion Matrix (Optional)

A confusion matrix on the test set highlights any misclassifications (e.g., Carrot vs. Radish).

Figure 1: Streamlit App Prediction Page: Showing a prediction for a carrot image, nutritional info, and confidence chart.

Figure 2: Training Curves: Training and validation accuracy/loss over 10 epochs.

8 Conclusion

The Vegetable Classifier project successfully achieves its objectives, delivering a high-accuracy classification system (99.50% validation accuracy) with a user-friendly interface. The integration of nutritional information and confidence scores enhances its practical utility. Future improvements could include webcam support, more detailed nutritional data, and fine-tuning for specific classes if needed.

Figure 3: Confusion Matrix: Showing model performance on the test set.