**Machine Learning – Case Study**

**Title: Fruits & Vegetable Classification & Calories Counter Web-App**

# Enrollment Number: 19012022001 Batch: DL2

**Type:** Web-Application

**Model:** MobilenetV2 Model

# Introduction

Dự án này về phân loại trái cây và rau củ. Đây là một ứng dụng web đơn giản mà mọi người dùng có thể sử dụng. Người dùng cần tải lên hình ảnh của bất kỳ loại trái cây hoặc rau củ nào. Hệ thống của chúng tôi sẽ tự động phân loại hình ảnh và đưa ra dự đoán về tên của trái cây hoặc rau củ, và chúng tôi đã thêm một module khác để tính lượng calo của đối tượng đã dự đoán. Đây là ứng dụng web, do đó người dùng có thể sử dụng trực tiếp trên bất kỳ trình duyệt nào.

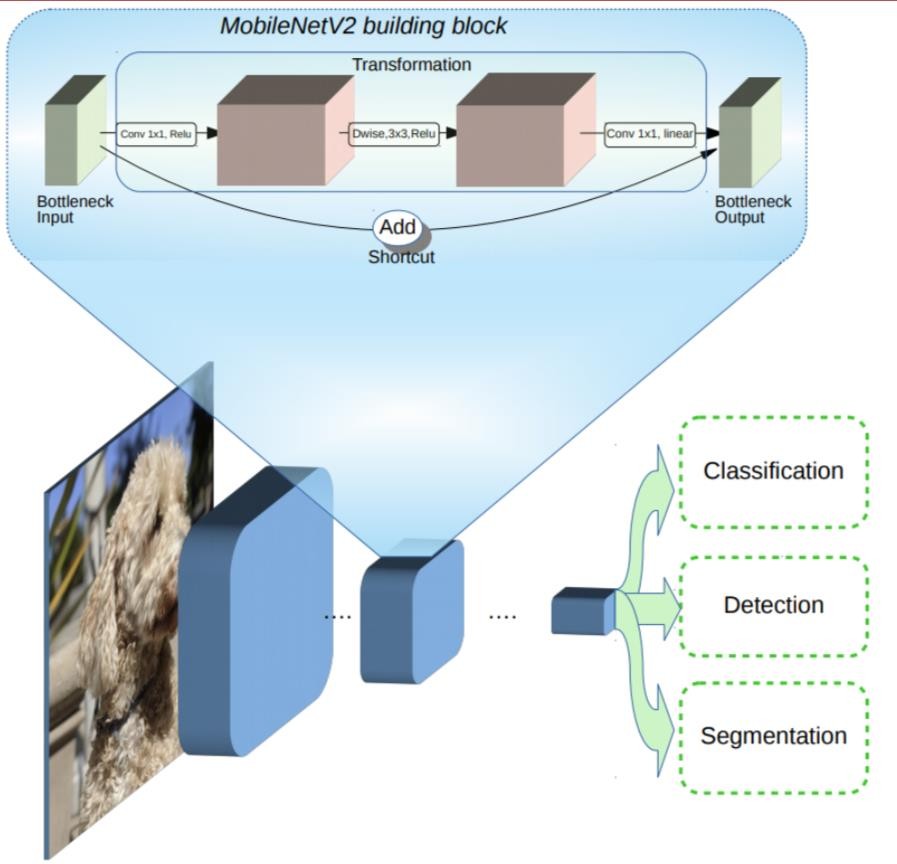
# Tools & Libraries

Chúng tôi sử dụng các công cụ và thư viện sau để phát triển dự án:

|  |  |  |
| --- | --- | --- |
| **No** | **Tools & Library Name** | **Usage** |
| 1 | Keras | Dùng cho các tác vụ học sâu như tạo mô hình, dự đoán đối tượng, v.v. |
| 2 | Pillow | Dùng để tiền xử lý hình ảnh trong tập dữ liệu. |
| 3 | Streamit | Là khung nền tảng backend để phát triển ứng dụng web. |
| 4 | Beautifulsoup, Requests | Dùng để trích xuất dữ liệu calo từ internet cho đối tượng đã dự đoán. |
| 5 | Numpy | Dùng để xử lý ma trận hình ảnh. |

# Kiến trúc, Tập dữ liệu & Quy trình làm việc

* **Kiến trúc:** Chúng tôi sử dụng kiến trúc MobilenetV2. MobileNetV2 là một kiến trúc mạng nơ-ron tích chập nhằm hoạt động hiệu quả trên các thiết bị di động. Nó dựa trên cấu trúc dư ngược, nơi các kết nối dư được thiết lập giữa các lớp cổ chai. Mobilenet hỗ trợ mọi kích thước đầu vào lớn hơn 32 x 32.
  + Trong MobileNetV2, có hai loại khối. Một là khối dư với bước chuyển, loại khác là khối với bước chuyển 2 để giảm kích thước.
  + Có 3 lớp cho cả hai loại khối.
  + Lớp đầu tiên là tích chập 1×1 với ReLU6.
  + Lớp thứ hai là tích chập sâu.
  + Lớp thứ ba là tích chập 1×1 khác nhưng không có tính phi tuyến.



*Figure 1. MobilenetV2 Architecture*

* **Tập dữ liệu:** Trong dự án này, chúng tôi sử dụng tập dữ liệu “Nhận dạng hình ảnh trái cây và rau củ”. Tập dữ liệu này có 36 lớp, với gần 100 hình ảnh cho mỗi lớp, do đó chúng tôi có khoảng hơn 3600 hình ảnh để huấn luyện. Chúng tôi có 10 hình ảnh cho mỗi lớp trong tập Train/Validation.

* **Quy trình làm việc:** Trong phần này, chúng ta sẽ xem cách ứng dụng web của chúng ta hoạt động. Chúng tôi đã chia các module để công việc của chúng tôi trở nên dễ dàng hơn. Phần frontend-backend sẽ được Streamlit xử lý. Người dùng sẽ truy cập vào ứng dụng của chúng tôi thông qua URL. Sẽ có nút tải lên để người dùng tải lên hình ảnh. Sau khi tải lên hình ảnh, hệ thống của chúng tôi sẽ thực hiện nhiệm vụ tự động. 
  + - Người dùng sẽ tải lên hình ảnh. Hình ảnh đó sẽ được lưu vào hệ thống cục bộ.
    - Sau đó, Pillow sẽ thay đổi kích thước hình ảnh theo kích thước mô hình của chúng tôi và chuyển đổi nó thành vector.
    - Vector này sẽ được truyền vào mô hình của chúng tôi, mô hình sẽ phân loại lớp của danh mục.
    - Chúng tôi sẽ nhận được ID của danh mục, và sau đó chúng tôi cần ánh xạ các nhãn theo ID.
    - Hệ thống của chúng tôi sẽ trích xuất dữ liệu calo từ web cho đối tượng đã dự đoán. Ứng dụng của chúng tôi sẽ hiển thị kết quả và lượng calo trên ứng dụng.

# Code

Mã nguồn mở và được công bố trên GitHub kèm theo tệp read-me.

https://github.com/Yuurei001/Fruit\_Vegetable\_Recognition

# Kết quả

Một số ảnh chụp màn hình của ứng dụng đang hoạt động.

