ĐẠI HỌC QUỐC GIA TPHCM

TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

ĐỒ ÁN MÔN HỌC

MÁY HỌC

ĐỀ TÀI:

PHÂN LOẠI THƯƠNG HIỆU NƯỚC GIẢI KHÁT

SINH VIÊN THỰC HIỆN:

Trần Đình Khang - 18520896

GIẢNG VIÊN HƯỚNG DẪN:

TS. Lê Đình Duy

ThS. Phạm Nguyễn Trường An

TP.HCM , ngày 15 tháng 8 năm 2020

MỤC LỤC

PHỤ LỤC ẢNH

I. ĐẶT VẤN ĐỀ

Với sự phát triển mạnh mẽ của đời sống xã hội, các loại nước hiện nay ngày càng được yêu thích và ngày càng đa dạng về thương hiệu, chủng loại, quy cách đóng gói. Các nhà sản xuất đều mong muốn tìm tòi để sản xuất ra các sản phẩm mới, cạnh tranh trên thị trường. Do đa dạng về loại, cạnh tranh ngày càng gay gắt, các sản phẩm bán chạy nhiều nên yêu cầu về việc đếm số lượng, phân loại sản phẩm cũng như phân tích tính hiệu quả Marketing cũng tăng lên. Do đó em tiến hành thực hiện đồ án "Phân loại nước giải khát được yêu thích ở Việt Nam" này để đáp ứng các nhu cầu trên.

II. MỤC TIÊU ĐỒ ÁN

Do đồ án được yêu cầu để phục vụ mục đích thương mại nên cần độ chính xác cao, tuy nhiên đây là đồ án nền tảng nên cũng là một thách thức lớn. Độ chính xác kỳ vọng đạt được là 85%.

Qua đồ án này mong muốn có thể hiểu và áp dụng hoàn thiện được 7 bước xây dựng một dự án machine learning, hiểu cơ bản về các phương pháp đặc trưng, các model được cung cấp bởi sklearn. Từ đó là nền tảng để có thể xây dựng các model deep learning phức tạp hơn cũng như góp nhặt được những kinh nghiệm qua quá trình làm đồ án, kinh nghiệm truyền đạt từ các giảng viên để vững bước trên con đường học tập Trí tuệ nhân tạo.

III. MÔ TẢ BÀI TOÁN

1. Mô tả về bài toán

- Tên bài toán: Phân loại nước giải khát được yêu thích ở Việt Nam.
- Bài toán sẽ phân loại 5 loại nước giải khát sau:
  - Pepsi
  - Cocacola
  - Twister
  - Redbull
  - Teaplus

2. Input và Output bài toán

- INPUT: Một ảnh chụp ở tỉ lệ 9:16 có chứa chai hoặc lon nước giải khát.

  VD như một trong các ảnh sau:

Ảnh 1

Ảnh 2

Ảnh 3

- OUTPUT: Tên loại nước giải khát trong 5 loại trên.

  VD: Pepsi

IV. MÔ TẢ VỀ BỘ DỮ LIỆU

1. Bộ dữ liệu phục vụ training model

- Dữ liệu dùng để xây dựng model được thu thập bằng cách cắt frame từ video quay thực tế.
- Video được quay với tỉ lệ khung hình 9:16, chất lượng full HD, 30 fps. Mỗi video dài khoảng 20s với 5 class, sau khi cắt frame ta thu được 600 ảnh cho mỗi class, tổng cộng có 3000 ảnh.
- Được chia theo tỉ lệ 80% training và 20% validation
- Dữ liệu được lưu tại đây: Link dữ liệu training
- VD:

Ảnh 4

Ảnh 5

Ảnh 6

2. Bộ dữ liệu phục vụ đánh giá model (test set)

- Nguồn dữ liệu này được thu thập bằng cách crawl dữ liệu từ internet, cụ thể là bing image search bằng 1 một cụ có tên google_download_image được chia sẻ trên github.
- Dữ liệu sau khi thu thập sẽ được xử lí bằng tay, loại bỏ đi cắt dữ liệu rác, cắt ảnh về tỉ lệ 9:16.
- Mỗi class sẽ có 60 ảnh, tổng là 300 ảnh.

- Dữ liệu được lưu tại đây: [Link dữ liệu test]
- VD:

Ảnh 7

Ảnh 8

Ảnh 9

Ảnh 10

Ảnh 11

3. Quá trình xử lí dữ liệu

- Các tập dữ liệu thu thập, phân loại và xử lí bằng tay nên khá chuẩn, do đó trong quá trình tiền xử lí sẽ được thực hiện kết hợp trong phần rút trích đặc trưng.
- Quá trình tiền xử lí dữ liệu sẽ thực hiện load ảnh từ đường dẫn thư mục, resize về 36x64 pixel (hệ số này để đảm bảo tỉ lệ ratio 9:16), đưa vào biến ma trận numpy và gán nhãn tương ứng.

V. MÔ TẢ VỀ CÁC ĐẶC TRƯNG

- Quá trình rút trích đặc trưng sẽ thực hiện tải ảnh lên, thực hiện các bước xử lí rồi lưu kết quả vào file định dạng .h5
- Đồ án sẽ thực hiện các phương pháp rút trích đặc trưng riêng biệt rồi so sánh kết quả giữa các phương pháp:
  - Pixel Is Feature: Đây là phương pháp đơn giản nhất, mỗi pixel của ảnh là một đặc trưng
  - Histogram Of Oriented Gradients
  - Local Binary Patterns
- Các file đã được trích xuất được lưu trong thư mục h5: Link đặc trưng
- Các phương pháp rút trích đặt trưng được cài đặt trong các file: simple_dataset_loader.py, simple_preprocessor.py, hog.py, lbps.py. Các file này được lưu trong thư mục đồ án và share trên github. Link github

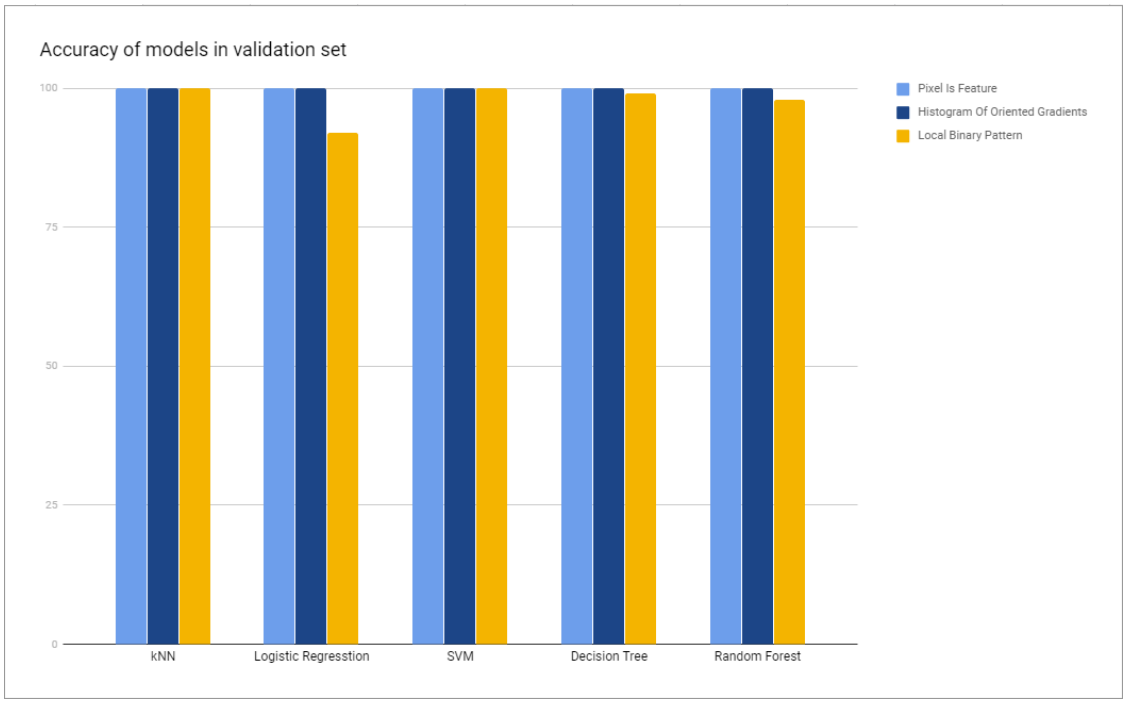
VI. MÔ TẢ VỀ CÁC ĐẶC TRƯNG

Phần này sẽ bao gồm:

- Import các thư viện cần thiết
- Tải lên các tập dữ liệu .h5 đã rút trích đặc trưng trước đó
- Phân chia dữ liệu traning và validation
- Sử dụng 5 model là:
  - Kneighbors Classifier
  - Logistic Regression
  - SVM
  - Decision Tree Classifier
  - Random Forest Classifier
- Với 3 phương pháp rút trích đặc trưng, có tất cả 15 model, thực hiện đánh giá kết quả của các model trên tập validation
- Lưu lại các model vào file, các file được đặt trong thư mục extract_model. Link model

VII.ĐÁNH GIÁ KẾT QUẢ, KẾT LUẬN

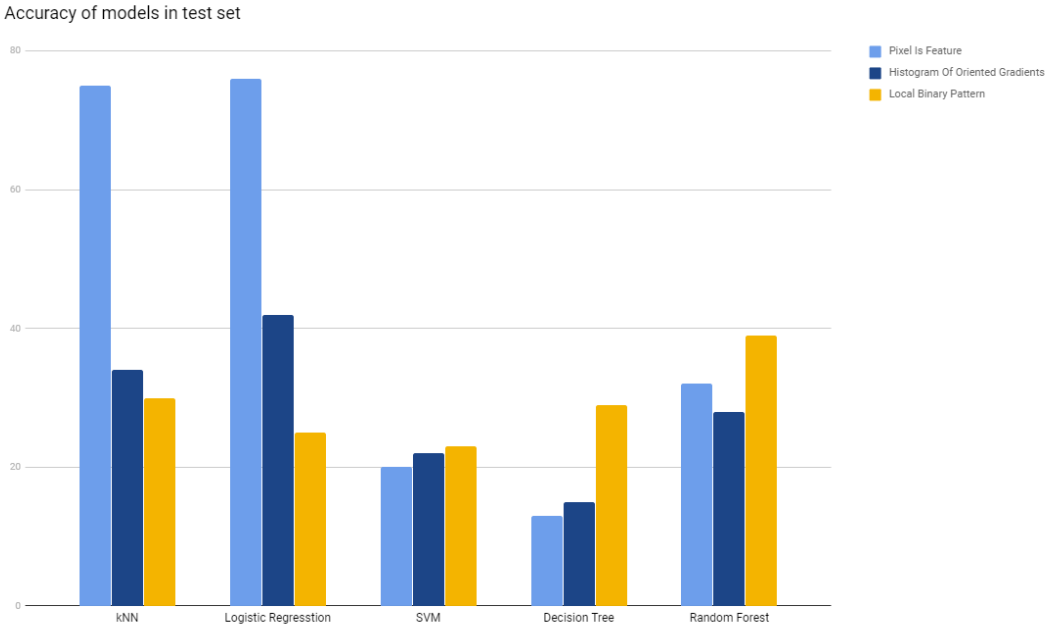
1. Kết quả của các model trên tập validation:



Ảnh 12: Kết quả của các model trên tập validation

Do cắt frame ảnh từ video, các ảnh trong tập dữ liệu khá giống nhau nên kết quả độ chính xác của các model trên tập validation set rất cao.

2. Kết quả của các model trên tập đánh giá (test set):



Ảnh 13: Kết quả của các model trên tập test

3. Nhận xét, kết luận:

a) Nhận xét về các thuật toán cùng các phương pháp rút trích đặc trưng

- Tổng quan: Phương pháp rút trích đặc trưng Pixel Is Feature cho kết quả cao nhất trong hầu hết các thuật toán. Thuật toán Logistic và kNN cho kết quả cao nhất trong các thuật toán.
- Với phương pháp rút trích đặc trưng Pixel Is Feature, các thuật toán Logistic và kNN cho kết quả độ chính xác lần lượt là 76% và 75% trên tập test
- Phương pháp rút trích đặc trưng Histogram Of Oriented Gradients và Local Binary Patterns cho kết quả xấp xỉ bằng nhau. Tuy nhiên, độ chính xác của chúng trên 5 thuật toán đều thấp (dưới 50%) nên không thể ứng dụng được, cần được cải thiện thêm.

- 2 model có khả năng ứng dụng được là kNN (Pixel Is Feature) và Logistic (Pixel Is Feature) nhưng với tỉ lệ 76% là một tỉ lệ tương đối, không cao lắm nên nếu muốn áp dụng cho mục đích thương mại cần độ chính xác cao như đếm, phân loại... thì cần phải cải thiện thêm.
- Có thể việc cắt frame ảnh từ video không làm đa dạng được bộ dữ liệu nên model có thể bị overfiting nhẹ, cần phải thu thập thêm dữ liệu

b) Các hướng cải thiện, phát triển thêm trong tương lai

- Thu thập thêm data để cải thiện độ chính xác của các model
- Dùng test set để turning model, thi thập thêm dataset
- Xây dựng thêm model deeplearning để đánh giá, so sánh độ chính xác
- Phát triển thành API hoặc xây dựng thành ứng dụng có thể dùng trên smartphone hoặc các phần cứng khác dùng trong thương mại.
- Xây dựng thêm model nhận dạng vật thể để cắt khung ảnh chứa vật thể rồi dùng khung ảnh đó làm input của model này để phân loại sản phẩm, có thể cho độ chính xác cao hơn.