

TÊN ĐỀ TÀI - PHÂN LOẠI NƯỚC GIẢI KHÁT

Trần Đình Khang - 18520896 - CS114.K21.KHTN

Link Github:

<https://github.com/trandinhkhang0279/CS114.K21.KHTN>

ĐẶT VẤN ĐỀ

Sự phát triển mạnh mẽ của đời sống xã hội

- Nước giải khát được yêu thích
- Các nhà sản xuất sản xuất ra các sản phẩm mới, cạnh tranh trên thị trường. Sản phẩm bán nhiều.
- Nhu cầu về đếm số lượng, phân loại sản phẩm, phân tích hiệu quả Marketing tăng
- Đồ án phân loại nước giải khát này ra đời.

MỤC TIÊU ĐỒ ÁN

- 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%.
- Hiểu và áp dụng hoàn thiện được 7 bước xây dựng một dự án machine learning.
- Hiểu rõ về các phương pháp rút trích đặc trưng và các model cung cấp bởi sklearn.
- Từ đó là nền tảng để có thể xây dựng các model deep learning phức tạp hơn. —➡ Con đường học tập AI.

MÔ TẢ BÀI TOÁN

- Bài toán sẽ phân loại 5 loại nước giải khát được yêu thích ở Việt Nam là: Pepsi, Cocacola, Twister, Redbull và Teaplush.
- INPUT: Một ảnh chụp ở tỉ lệ 9:16 có chứa chai hoặc lon nước giải khát.
- OUTPUT: Tên loại nước giải khát trong 5 loại trên



Teaplush



Redbull

MÔ TẢ VỀ BỘ DỮ LIỆU

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

MÔ TẢ VỀ BỘ DỮ LIỆU

Bộ dữ liệu phục vụ training model:

VD:



MÔ TẢ VỀ BỘ DỮ LIỆU

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.

MÔ TẢ VỀ BỘ DỮ LIỆU

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

VD:



MÔ TẢ VỀ BỘ DỮ LIỆU

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.

MÔ TẢ VỀ CÁC ĐẶC TRƯNG

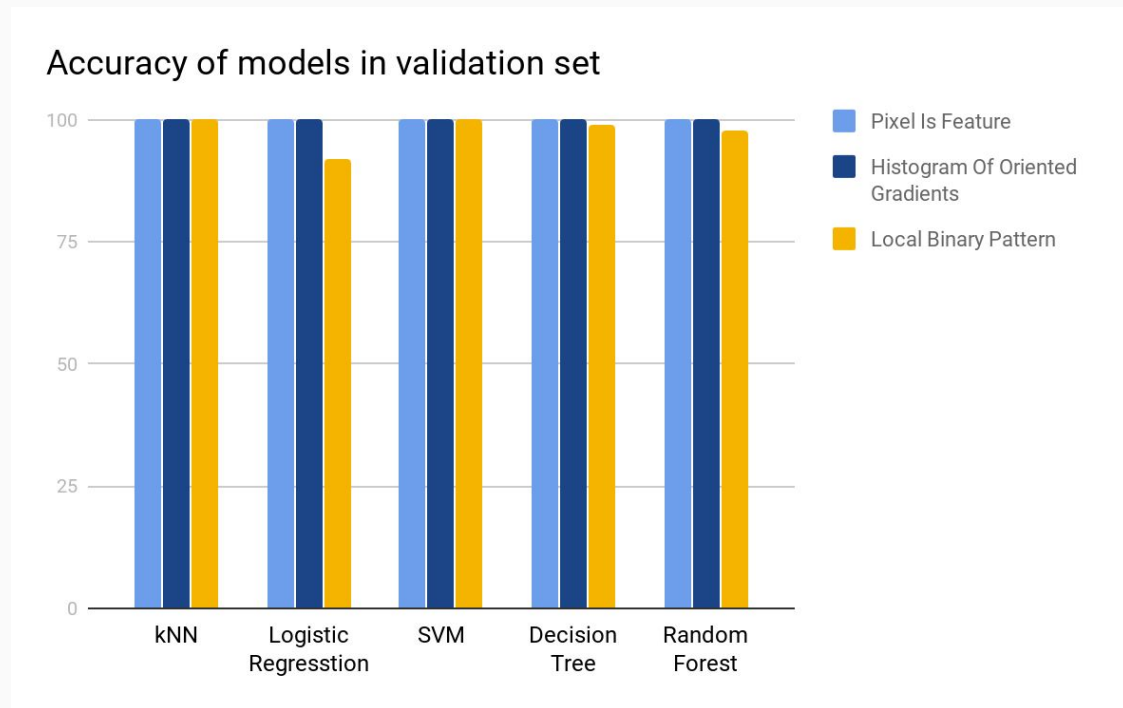
- Quá trình rút trích đặc trưng sẽ thực hiện tải ảnh lên, thực hiện các bước xử lý rồi lưu kết quả vào file định dạng .h5
- Đồ án sẽ thực hiện các phương pháp rút trích đặc trưng riêng biệt rồi so sánh kết quả giữa các phương pháp:
 - Pixel Is Feature: Đây là phương pháp đơn giản nhất, mỗi pixel của ảnh là một đặc trưng
 - Histogram Of Oriented Gradients
 - Local Binary Patterns

MÔ TẢ VỀ THUẬT TOÁN MÁY HỌC

- Import các thư viện cần thiết
- Tải lên các tập dữ liệu .h5 đã rút trích đặc trưng trước đó
- Phân chia dữ liệu training và validation
- Sử dụng 5 model là: KNeighborsClassifier, LogisticRegression, SVM model, DecisionTreeClassifier và RandomForestClassifier.
- 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`

ĐÁNH GIÁ KẾT QUẢ, KẾT LUẬN

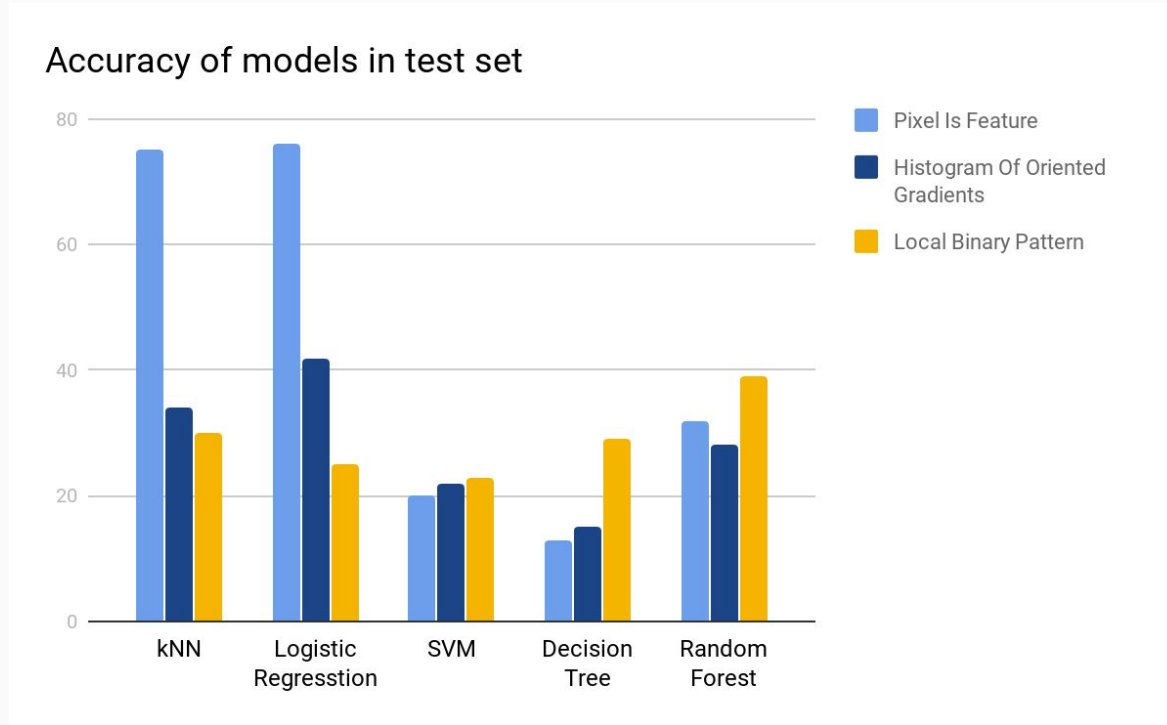
Kết quả của các model trên tập validation



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

ĐÁNH GIÁ KẾT QUẢ, KẾT LUẬN

Kết quả của các model trên tập test



ĐÁNH GIÁ KẾT QUẢ, KẾT LUẬN

- 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
- Có thể việc cắt frame ảnh từ video không làm đa dạng được bộ dữ liệu nên model có thể bị overfitting nhẹ, cần phải thu thập thêm dữ liệu

ĐÁNH GIÁ KẾT QUẢ, KẾT LUẬN

- 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.

HƯỚNG CẢI THIỆN, PHÁT TRIỂN

- 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.