XLA – Xử lý ảnh

# MỞ ĐẦU

Nhằm bắt kịp với những ứng dụng thực tiễn mà XLA đã và đang mang lại cho đời sống và trong nền công nghiệp của nước ta, nhóm tụi em quyết định thực hiện đề tài ***Hệ thống phân loại Cam / Chanh sử dụng công nghệ Xử lý ảnh*** làm nền tảng cho thấy được ứng dụng rõ nét của XLA trong các ngành công nghiệp hàng hóa (ở đây là phân loại sản phẩm). Đề tài có thể đươc mở rộng và sử dụng rộng rãi với các biến thể như: Hệ thống phân loại hàng hóa dựa dựa trên kích thước mà màu sắc từ đó đưa đúng loại hàng vào khu vực lưu trữ, Hệ thống phân loại các loại hạt, củ quả kém chất lượng, không đạt yêu cầu về mặt hình ảnh bên ngoài…. Vì vậy, đề tài được chúng em xây dựng nhằm mô phỏng lại toàn bộ quy trình xử lý và phân loại hàng hóa, giúp mọi người hiểu rõ hơn về cách thức hoạt động và biết thêm được những ứng dụng thực tế mà nó có thể mang lại.

Nội dung đề tài được tổ chức thành 5 chương:

**Chương 1: Tổng quan đề tài**

Trong chương này chúng em sẽ giới thiệu về đề tài, lý do chọn đề tài, đưa ra các hướng hoàn thiện đề tài và ứng dụng của nó khi hoàn thiện sản phẩm.

**Chương 2: Cơ sở lý thuyết**

Trong chương này chúng em sẽ đi tìm hiểu về mức độ cần thiết của đề tài, đề tài sẽ được áp dụng vào đâu, những bài toán cần giải quyết trong vấn đề phân loại hàng hóa, sản phẩm. Từ đó xây dựng lộ trình phát triển sản phẩm trong đề tài này.

**Chương 3: Phân tích yêu cầu**

Sau khi đã xác định được bài toán cần giải quyết trong đề tài này, chúng em sẽ cùng nhau phân tích xem cần phải thực hiện điều gì trong đề tài này. Phân tích chi tiết chức năng thực tiễn của từng phần trong bài toán, đưa ra được hướng phát triển và hoàn thiện sản phẩm.

**Chương 4: Triển khai ứng dụng**

Khi đã có được hướng đi cụ thể ở Chương 3, trong Chương này chúng em sẽ bắt tay vào giải quyết từng phần nhỏ của đề tài đã đặt ra, bắt đầu thiết kế và xây dựng một hệ thống hoàn chỉnh và ổn định.

**Chương 5: Kết luận**

Chương này sẽ nói sâu vào kết quả đạt được của đề tài, nói lên những điều làm được và chưa làm được để rút kinh nghiệm phát triền để đưa đề tài này tới được tầm với của nhiều người hơn nữa.

# CHƯƠNG 1. TỔNG QUAN ĐỀ TÀI

## TÊN ĐỀ TÀI

***” Hệ thống phân loại Cam / Chanh sử dụng công nghệ Xử lý ảnh”***

## LÝ DO CHỌN ĐỀ TÀI

Trí tuệ nhân tạo hiện đang làm mưa làm gió và nắm một phần thiết yếu trong cuộc đua của các nước khi tiế tới cuộc cách mạng công nghiệp 4.0, nó làm thay đổi gần như mọi mặt của cuộc sống từ vấn đề giao thông, ý tế cho đến giáo dục. Xử lý ảnh (Computer Vision) hiện đang là một nhánh của Trí tuệ nhân tạo đang được phát triển mạnh mẽ do tính ứng dụng cao trong đời sống thường nhật của mọi người như: Xe tự hành, Bãi giữ xe thông minh, Trạm thu phí không dừng… và những ứng dụng khác trong công nghiệp như các hệ thống phân loại hàng hóa, đánh giá chất lượng sản phẩm, hệ thống thu hoạch củ quả tự động bằng robot ….

Tự động hóa trong phân loại hàng hóa đã là một tiêu chuẩn hướng đến của các công ty phân phối, nó giúp tiếp kiệm được chi phí nhân công, năng suất làm việc cao hơn nhiều so với người bình thường và với sự phát triển của AI thì việc một ngày nào đó máy móc sẽ thay thế con người trong công đoạn phân phối này sẽ không còn xa. Tuy mạnh mẽ là vậy nhưng ứng dụng này của XLA vẫn chưa được nhiều người biết đến hoặc là đã biến đến nhưng chưa biết cách ứng dụng vào trong công việc sản xuất của mình, gây ra thiệt hại khá đáng kể về mặt giá thành nhân công và hiệu suất làm việc củ quy trình.

Dựa vào tầm quan trọng của XLA trong việc phân phối hàng hóa, chúng em muốn xây dựng một mô hình mô phỏng lại hệ thống phân loại sản phẩm ứng dụng trong nông nghiệp thực phẩm, xem xét và phân loại loại thực phẩm trên băng truyền.

## MỤC TIÊU CỦA ĐỀ TÀI

Xây dựng mô hình phân loại Cam/ Chanh sử dụng công nghệ Xử lý ảnh với deep learning.

Mô hình hoạt động có độ chính xác cao, tốc độ xử lý hình ảnh của máy tính nhúng Raspberry Pi 3 từ 5-6 FPS .

## PHẠM VI NGHIÊN CỨU

Phạm vi của đề tài dừng ở mức mô phỏng hoàn chỉnh quá trình phân loại sản phẩm (Cam/ Chanh).

## PHƯƠNG PHÁP THỰC HIỆN

### Nghiên cứu sơ bộ

Trước khi bắt đàu một đề tài, ta cần phải tìm hiểu xem đề tài đó đã có ai thực hiện hay chưa, nhưng điểm mạnh, yếu mà những đề tài trước đã gặp phải mà mình cần phải lưu ý kỹ.

Ta cần phải tìm hiểu xem những công nghệ nào đã được áp dụng vào trong các đề tài liên quan, cách các chủ đề tài đó đưa ra vấn đề và bắt tay vào giải quyết chúng. Ngoài ra cần tìm được công nghệ phù hợp với đề tài mà mình thực hiện nhằm đạt được kết quả thu được cao nhất có thể.

### Chọn công nghệ phù hợp

Công nghệ nền tảng của sản phẩm luôn là mấu chốt cho sự thành công của sản phẩm đó, nó ảnh hướng tới hướng phát triển sản phẩm, cách tiếp cận nghiên cứu và cách áp dụng sản phẩm vào trong đời sống. Biết được sức nặng mà việc lựa chọn công nghệ đang nắm giữ trong đề tài, đây sẽ là công đoạn khá khó khăn do phải lựa chọn công nghệ phù hợp với đề tài vừa khái quát được chức năng và ý nghĩa của đề tài, vừa giúp sản phẩm khi hoàn thành đạt được hiệu suất hoạt động cao nhất có thể.

### Thiết kế mô hình

Trong phần này, chúng em sẽ cùng nhau thiết kế để tạo ra được mô hình hoàn chỉnh có độ hiệu quả cao nhất. Các vấn đề liên quan tới vật liệu sử dụng, kết cấu giá đỡ băng chuyền, kết cấu lắp đặt camera và cần gạt cũng ảnh hướng rất nhiều tới quá trình hoạt động của mô hình.

### Lăp đặt và kiểm thử

Sau khi đã hoàn thiện mô hình thì đây là bước khá quan trọng, ta phải kiểm xem mô hình đã hoạt động hiêu quả hay chưa, những trục trặc nào có thể xảy ra trong quá trình sử dụng từ đó đưa ra được những biện pháp phòng ngừa để mô hình luôn được hoạt động hiện quả.

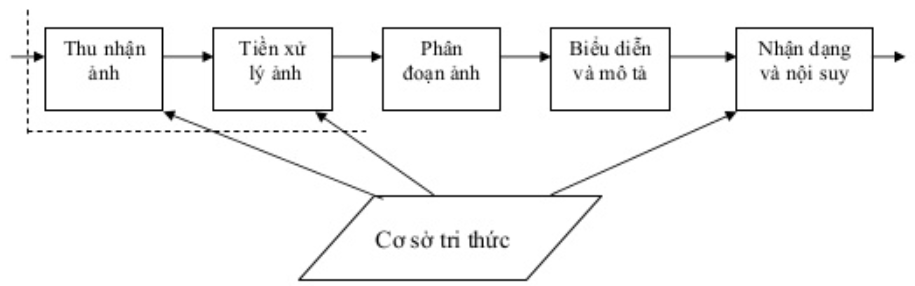
# CHƯƠNG 2. CƠ SỞ LÝ THUYẾT

## 2.1. CỞ SỞ LÝ THUYẾT THỊ GIÁC MÁY TÍNH

### 2.1.1. Định nghĩa Thị giác máy tính

Theo wikipedia: Thị giác máy tính (*Computer Vision*) là một lĩnh vực bao gồm các phương pháp thu nhận, xử lý ảnh ký thuật số, phân tích và nhận dạng các hình ảnh, nói chung là dữ liệu đa chiều từ thế giới thực để cho ra các thông tin số hoặc biểu tượng, ví dụ trong các dạng quyết định. Việc phát triển lĩnh vực này có bối cảnh từ việc sao chép các khả năng thị giác con người bởi sự nhận diện và hiểu biết một hình ảnh mang tính điện tử.

[1] Thị giác máy tính là một môn học khoa học liên quan đến lý thuyết đằng sau các hệ thống nhân tạo có trích xuất các thông tin từ các hình ảnh. Dữ liệu hình ảnh có thể nhiều dạng, chẳng hạn như chuỗi video, các cảnh từ đa camera, hay dữ liệu đa chiều từ máy quét y học. Thị giác máy tính còn là một môn học kỹ thuật, trong đó tìm kiếm việc áp dụng các mô hình và các lý thuyết cho việc xây dựng các hệ thống thị giác máy tính.



*Các bước cơ bản trong xử lý ảnh [1]*

1. **Thu nhận ảnh (Image Acquisition)**

[1] Ảnh có thể nhận qua camera màu hoăc đen trắng. Thường ảnh nhận qua camera là ảnh tương tự (loại camera ống chuẩn CCIR với tần số 1/25, mỗi ảnh 25 dòng), cũng có loại camera đã số hóa là loại photodiot tạo cường độ sáng tại mỗi điểm ảnh.

Camera thường dùng là loại quét dòng, ảnh tạo ra có dạng hai chiều. Chất lượng một ảnh thu nhận được phụ thuộc vào thiết bị thu và môi trường (ánh sáng, phong cảnh).

1. **Tiền xử lý ảnh (Image Processing)**

[2] Hình ảnh tiền xử lý là hình ảnh chưa được chỉnh sửa ở bất kỳ phương diện nào. Ở bước này hình ảnh sẽ được cải thiện về độ tương phản, khử nhiễu, khử bóng, khử độ lệch… và với mục đích làm cho ảnh trở lên tốt hơn nữa và thường được thự hiện bởi những bộ lọc. Có rất nhiều phương pháp để xử lý ảnh ở giai đoạn này:

* **Nhị phân ảnh:** là thao tác chuyển ảnh từ ảnh đa cấp xám (ảnh màu) về ảnh nhị phân.
* **Hiệu chỉnh độ nghiêng của trang văn bản:** thường được sử dụng trong các trường hợp ảnh đầu vào là những trang tài liệu dạng văn bản thường bị lệch so với ảnh gốc một góc bất kỳ. Do đó, để cho các bước xử lý tiếp theo (phân tích và nhận dạng) được chính xác cẩn phải có thao tác hiệu chỉnh độ nghiêng của ảnh thu nhận được.

1. **Phân đoạn hay phân vùng ảnh (Segmentation)**

Phân vùng ảnh là tách một ảnh đầu vào thành các vùng thành phần để biểu diễn phân tích, nhận dạng ảnh. Mỗi vùng ảnh phải có một đặc tính đồng nhất, hình dáng của một đối tượng có thể đươc miêu tả hoặc bởi các tham số của đường biên hoăc các tham số của vùng mà nó chiếm dữ.

Có thể thấy kỹ thuật phát hiện biên và phân vùng ảnh là hai bài toán đối ngẫu của nhau. Dò biên để phân vùng được ảnh và ngược lại phân vùng được ảnh ta có thể phát hiện được biên.

[2] Có rất nhiều kỹ thuật phân đoạn ảnh, chúng ta có thể chia thành ba lớp khác nhau:

* **Các kỹ thuật cục bộ:** dựa vào các thuộc tính cục bộ của điểm ảnh và lán giềng của nó.
* **Các kỹ thuật toàn thể:** phân ảnh dựa trên thông tin chung của toàn bộ ảnh (vd sử dụng lược đồ xám của ảnh).
* **Các kỹ thuật tách (split), hợp (merge) và growing:** sử dụng các khái niệm đồng nhất và gắn về hình học.

1. **Biểu diễn ảnh (Image Representation) [1]**

Đầu ra ảnh sau phân đoạn chữa các điểm ảnh của vùng ảnh (ảnh đã phân đoạn) cộng với mã liên kết với các vùng lân cận. Việc biến đổi các số liệu này thành dạng thích hợp là cần thiết cho xử lý tiếp theo bằng máy tính. Việc chọn các tính chất để thể hiện ảnh gọi là trích chọn đặc trưng (*Feature Selection*) gắn với việc tách các đăc tính của ảnh dưới dạng các thông tin định lượng hoặc làm cơ sở để phân biệt lớp đối tượng này với đối tượng khác trong phạm vi ảnh nhận được.

1. **Nhận dạng và nội suy ảnh (Image Recognition and Interpretation) [1]**

Nhận dạng ảnh là quá trình xác định ảnh. Quá trình này thường thu được bằng cách so sánh với mẫu chuẩn đã được học (hoặc lưu) từ trước. Nội suy là phán đoán theo ý nghĩa trên cơ sở nhận dạng.

Một số đối tượng nhận dạng khá phổ biến hiện nay đang được áp dụng trong khoa học và công nghệ là: nhận dạng ký tự (chữ in, chữ viết tay, chữ ký điện tử), nhận dạng văn bản (Text), nhận dạng vân tay, nhận dạng mã vạch, nhận dạng mặt người…

1. **Cơ sở chi thức (Knowledge Base) [1]**

Anh là một đối tượng khá phức tạp về đường nét, độ sáng tối, dung lượng điểm ảnh, môi trường để thu ảnh phong phú kéo theo nhiễu. Trong nhiều khâu xử lý và phân tích ảnh ngoài việc đơn giản hóa các phương pháp toán học đảm bảo tiện lợi cho xử lý, người ta mong muốn bắt chước quy trình tiếp nhận và xử lý ảnh theo cách của con người. Trong các bước xử lý đó, nhiều khâu hiện nay đã xử lý theo các phương pháp trí tuệ con người. Vì vậy, ở đây các cơ sở tri thức được phát huy.

### 2.1.2. Ứng dụng của thị giác máy tính [3][4]

Thị giác máy tính cho phép các máy tính cũng như robot, các phương tiện điều khiển từ máy tính và mọi thứ từ nhà máy, thiết bị nông trại đến xe hơi và máy bay có thể thực hiện một số hoạt động tự động, nó hoạt động một cách hiệu quả, thậm chí an toàn hơn.

* **Phát hiện khuyết điểm:**  Cho đến bây giờ thì việc phát hiện ra các yếu tố bị lỗi thường được tiến hành bởi những người giám sát chỉ định nhưng nếu mở rộng hơn họ không thể nào kiểm soát được toàn bộ cả một quy trình hệ thống được. Với thị giác máy tính, chúng ta có thể kiểm tra tất cả các lỗi nhỏ nhất với tốc độ nhanh và tốt hơn mắt thường con người gấp nhiều lần. Thuật toán này được thiết kế và đào tạo đặc biệt cho từng ứng dụng cụ thể thông qua hình ảnh có khiếm khuyết và không có khuyết tật.
* **Trình đọc tự động:** Chắc hẳn ứng dụng này không hề xa lạ với bạn khi mà nó xuất hiện thường xuyên khi bạn sử dụng Google translate, nó có thể nhận được văn bản với bất cứ ngôn ngữ nào thông qua camera trên điện thoại thông mình.
* **Vận hành tự động:** Những chiếc xe tự lái đang làm mưa làm gió trên thị trường công nghệ hiện nay, lĩnh vực này phụ thuộc rất nhiều vào thị giác máy tính và học sâu (*Deep learning*). Công nghệ này phân tích dữ liệu thu thập được từ hàng triệu người lái xe, học hỏi hành vi lái xe để tự động tìm làn đường, ước tính độ cong, phát hiện các mối nguy hiểm và nhận diện các tín hiệu giao thông….

Các ứng dụng của Thị giác máy tính trong đời sống của chúng ta còn nhiều và đa dạng hơn nữa, nó mang tới một làn sóng mới cho nhân loại, nơi máy mọc có thể nhìn được thế giới bên ngoài như cách mà con người nhìn trong hàng nghìn năm nay…

## 2.2. THỊ GIÁC MÁY TÍNH VỚI DEEP LEARNING

### 2.2.1. Học sâu (Deep learning)

Học sâu (*Deep learning*) là một tập hợp con của Học máy (*Machine learning*) được xây dựng dựa trên thuật toán Neural network. Cùng với khả năng tính toán vượt trội của máy tính và lượng dữ liệu khổng lồ mà con người tạo ra, Deep learning đang có những bước đột phá thực sự.

1. **Mạng thần kinh nhân tạo (Artifical Neural Network)**

Theo Wikipedia: Mạng thần kinh nhân tạo hay thường gọi ngắn gọn là **mạng neural** (*tiếng Anh là Artifical Neural Network – ANN hay Neural Network*) là một [mô hình toán học](https://vi.wikipedia.org/wiki/M%C3%B4_h%C3%ACnh_to%C3%A1n_h%E1%BB%8Dc) hay [mô hình tính toán](https://vi.wikipedia.org/wiki/M%C3%B4_h%C3%ACnh_t%C3%ADnh_to%C3%A1n) được xây dựng dựa trên các [mạng neural sinh học](https://vi.wikipedia.org/w/index.php?title=M%E1%BA%A1ng_n%C6%A1-ron_sinh_h%E1%BB%8Dc&action=edit&redlink=1). Nó gồm có một nhóm các [neural nhân tạo](https://vi.wikipedia.org/w/index.php?title=N%C6%A1-ron_nh%C3%A2n_t%E1%BA%A1o&action=edit&redlink=1) (nút) nối với nhau, và xử lý thông tin bằng cách truyền theo các kết nối và tính giá trị mới tại các nút (cách tiếp cận connectionism đối với [tính toán](https://vi.wikipedia.org/wiki/T%C3%ADnh_to%C3%A1n)). Trong nhiều trường hợp, mạng neural nhân tạo là một [hệ thống thích ứng](https://vi.wikipedia.org/w/index.php?title=H%E1%BB%87_th%E1%BB%91ng_th%C3%ADch_%E1%BB%A9ng&action=edit&redlink=1) (adaptive system) tự thay đổi cấu trúc của mình dựa trên các thông tin bên ngoài hay bên trong chảy qua mạng trong quá trình học.

[6] Mạng neural ra đời xuất phát từ ý tưởng mô phỏng bộ não con người. Giống như con người, Mạng neural được học bởi kinh nghiệm, lưu những kinh nghiệm đó và sử dụng trong tình huống phù hợp.

* **Học có giám sát (*supervised learning*)**

Học có giám sát là nhóm thuật toán dự đoán đầu ra (output) của dữ liệu mới (new input) dựa trên các cặp dữ liệu đã biết trước. Cặp dữ liệu này còn được gọi là dữ liệu - nhãn (data - label). Đây là nhóm phổ biến nhất trong các thuật toán học máy.

Nhóm thuật toán học có giám sát gồm các bài toán chính sau:

Phân loại (classification): Các nhãn của dữ liệu đầu vào được chia thành các nhóm hữu hạn.

Hồi quy (regression): Nhãn là một giá trị thực cụ thể. Ở nghiên cứu này, nhóm tác giả đã áp dụng bài toán hồi quy để dự báo phân bố độ rỗng của vỉa.

* **Học không giám sát (*unsupervised learning*)**

Trong thuật toán này không biết trước được đầu ra hay nhãn của tập dữ liệu đầu vào, chỉ dựa vào cấu trúc của dữ liệu để thực hiện công việc như: phân nhóm (clustering) hoặc giảm số chiều của dữ liệu (dimension reduction) để thuận tiện trong việc lưu trữ và tính toán. Học không giám sát là khi chỉ có dữ liệu đầu vào X mà không biết nhãn Y tương ứng.

* **Học bán giám sát (*semi – supervised learning*)**

Các bài toán khi có một lượng lớn dữ liệu X nhưng chỉ có một phần được gán nhãn được gọi là học bán giám sát. Những bài toán thuộc nhóm này nằm giữa 2 nhóm trên

* **Học củng cố (*reinforcement learning*)**

Học củng cố giúp hệ thống tự động xác định hành vi dựa trên hoàn cảnh để đạt được lợi ích cao nhất (maximising the performance).

<https://vi.wikipedia.org/wiki/M%E1%BA%A1ng_th%E1%BA%A7n_kinh_nh%C3%A2n_t%E1%BA%A1o>

<https://vi.wikipedia.org/wiki/M%E1%BA%A1ng_th%E1%BA%A7n_kinh_nh%C3%A2n_t%E1%BA%A1o>

<https://en.wikipedia.org/wiki/Deep_learning>

[1]

<https://www.slideshare.net/elLeonNo1/gio-trnh-x-l-nh>

[2]

<http://luanvan.net.vn/luan-van/xu-ly-anh-tien-xu-ly-va-phan-doan-anh-71860/>

[4]

<https://longvan.net/thi-giac-may-tinh-la-gi-ung-dung-thi-giac-may-tinh.html>

[5]

<http://tailieuhoctap.vn/chi-tiet-sach/412-luan-van-de-tai-tham-khao/luan-van-de-tai-cao-hoc/769878-ung-dung-cua-thi-giac-may-tinh-trong-san-xuat-nong-nghiep>

<https://users.soict.hust.edu.vn/ductq/XLA%20Lecture.pdf>

[6]

<http://www.pvn.vn/DataStore/Documents/2019/TapchiDK/so%207.%202019/mang%20neuron%20Ta%20Quoc%20Dung.pdf>