**ĐẠI HỌC QUỐC GIA TP.HCM**

**TRƯỜNG ĐẠI HỌC BÁCH KHOA**

**KHOA KHOA HỌC VÀ KỸ THUẬT MÁY TÍNH**

****

**BÁO CÁO**

**ĐỒ ÁN CHUYÊN NGÀNH**

**MÔ HÌNH ROBOT PHÂN LOẠI HÀNG HOÁ**

Ngành: Kỹ thuật máy tính

|  |  |
| --- | --- |
| **HỘI ĐỒNG:** |  |
| **GVHD:** | Ts. Phạm Hoàng Anh  Phạm Công Thái |
| **GVPB:** |  |
| **SVTH:** | Nguyễn Tuấn Vinh (1915944) |

TP. HỒ CHÍ MINH, THÁNG 5 NĂM 2024

# TỔNG QUAN

## Phân loại hàng hoá và những lợi ích từ việc phân loại hàng hoá hiệu quả.

Trong quá trình Sản xuất hàng hoá và Phân phối hàng hoá, phân loại sản phẩm là một công việc quan trọng trong công đoạn xử lý. Đối với một doanh nghiệp, việc phân loại sản phẩm nhanh chống và hiệu quả giúp:

* Nâng cao hiệu quả hoạt động:
  + Tăng tốc độ xử lý sản phẩm, xử lý nguyên liệu, giảm thời gian lưu kho, giúp doanh nghiệp tiết kiệm chi phí và nâng cao năng suất. Thông thường, một doanh nghiệp sẽ phải đầu tư 3 – 10% cho chi phí kho hàng, 3 – 5% cho chi phí về nhân lực cho các hoạt động giám sát và quản lý kho, 6 – 24% cho chi phí đầu tư vào hàng tồn kho và 2 – 5% cho chi phí về các thiệt hàng hàng hoá.[[1]](#footnote-1)
  + Giảm thiểu sai sót trong quá trình phân loại, đảm bảo chất lượng sản phẩm và hạn chế rủi ro cho doanh nghiệp.
  + Giảm thiểu tổn thất do hư hỏng, lỗi thời hoặc thất lạc.
  + Cải thiện khả năng truy xuất nguồn gốc sản phẩm, giúp doanh nghiệp quản lý hàng hóa hiệu quả hơn.
* Tăng cường khả năng cạnh tranh:
  + Giúp doanh nghiệp đáp ứng nhu cầu thị trường một cách nhanh chóng và hiệu quả.
  + Nâng cao uy tín và thương hiệu của doanh nghiệp, thu hút khách hàng tiềm năng.
  + Đáp ứng đơn hàng của khách hàng nhanh chóng và chính xác hơn. Từ đó, doanh nghiệp có thể cung cấp dịch vụ khách hàng tốt hơn bằng cách giải quyết các vấn đề về đơn hàng và giao hàng một cách nhanh chóng.
  + Tạo lợi thế cạnh tranh so với các đối thủ trong cùng ngành.
* Thúc đẩy phát triển kinh tế:
  + Góp phần vào sự phát triển của ngành công nghiệp logistics và chuỗi cung ứng.
  + Tạo ra nhiều việc làm, thúc đẩy sự phát triển của kinh tế địa phương.
  + Nâng cao năng lực cạnh tranh của quốc gia trên thị trường quốc tế.

Bên cạnh đó, công việc phân loại sản phẩm là một công việc nhàm chán, lặp đi lặp lại và đòi hỏi tính chính xác, cẩn thận và tỉ mĩ cao. Để giảm thiểu chi phí đầu tư cho việc phân loại hàng hoá, thoả mãn các yêu cầu và tính chất công việc kể trên nhưng vẫn đáp ứng được tốc độ yêu cầu trong các quy trình sản xuất tự động thì việc ra đời của các robot phân loại sản phẩm là điều tất yếu trong thời đại công nghiệp hoá - hiện đại hoá hiện nay.

## Tổng quan về thị trường Robot phân loại sản phẩm hiện nay

* Mức độ phát triển:

Ngành công nghiệp robot phân loại sản phẩm ra đời vào đầu những năm 1980, với tốc độ phát triển nhanh chóng trong những thập kỷ qua. Tuy nhiên, ngành công nghiệp robot phân loại sản phẩm đã chứng kiến sự phát triển đáng kể trong thập kỷ gần đây, đặc biệt từ năm 2010 trở đi. Từ việc sử dụng công nghệ máy học đến các tiến bộ trong robot hình thái và cảm biến, ngành này đã phát triển với tốc độ nhanh chóng. Theo số liệu từ The International Federation of Robotics (IFR) trong Báo cáo thế giới về robot công nghiệp năm 2020 (IFR presents World Robotics Report 2020), Báo cáo cho thấy kỷ lục 2,7 triệu robot công nghiệp đang hoạt động trong các nhà máy trên khắp thế giới - tăng 12%. Doanh số bán robot công nghiệp mới vẫn ở mức cao với 373.000 chiếc được xuất xưởng trên toàn cầu vào năm 2019. Con số này ít hơn 12% so với năm 2018, nhưng vẫn là doanh số bán hàng cao thứ 3 từng được ghi nhận[[2]](#footnote-2). Trong thời gian này, thị trường robot phân loại sản phẩm đã trở thành một phần không thể tách rời trong chuỗi cung ứng sản phẩm toàn cầu. Theo Mordor Intelligence, thị trường robot phân loại sản phẩm toàn cầu dự kiến đạt 114,67 tỷ USD vào năm 2023 và tăng trưởng với tốc độ CAGR 17,64% từ 2023 đến 2028.[[3]](#footnote-3)

* Các yếu tố ảnh hưởng:

Các yếu tố thúc đẩy sự phát triển của ngành bao gồm sự tăng cao trong nhu cầu tự động hóa cao trong các ngành công nghiệp, đặc biệt là logistics, thực phẩm và đồ uống, dược phẩm, điện tử đi kèm với nhu cầu nâng cao hiệu quả, năng suất, giảm chi phí sản xuất đảm bảo an toàn lao động và giảm thiểu sai sót do con người. Bên cạnh đó còn có sự phát triển của các công nghệ như trí tuệ nhân tạo, thị giác máy tính, học máy. Tuy nhiên, ngành sản xuất và phát triển Robot phân loại sản phẩm còn gặp nhiều hạn chế như chi phí đầu tư ban đầu cao, nhu cầu về nhân lực có trình độ kỹ thuật cao để vận hành - bảo trì và khả năng ứng dụng hạn chế cho một số sản phẩm đặc biệt

* Đặc điểm robot phân loại sản phẩm hiện nay:
* Cấu tạo: Các robot phân loại sản phẩm hiện nay thường được thiết kế để có khả năng xử lý sản phẩm ở nhiều kích thước, hình dạng và trọng lượng khác nhau. Chúng thường được trang bị cánh tay robot, hệ thống camera, các bộ cảm biến, hệ thống điều khiển, hệ thống băng tải.
* Nguyên lý hoạt động: Sử dụng camera để nhận diện sản phẩm, sau đó sử dụng cánh tay robot để gắp và phân loại sản phẩm theo các tiêu chí định trước.
* Số bậc tự do: Các robot phân loại sản phẩm hiện nay có từ 4 đến 6 bậc tự do, cho phép thực hiện các thao tác phức tạp như gắp, xoay, di chuyển sản phẩm.
* Hạn chế: Khả năng nhận diện sản phẩm phức tạp và khả năng thích ứng với môi trường thay đổi còn hạn chế.
* Công nghệ áp dụng: Trí tuệ nhân tạo, thị giác máy tính, học máy, robot cộng tác.
* Nhà cung cấp robot phân loại sản phẩm:

Trong nước Việt Nam, một số công ty hàng đầu trong lĩnh vực này bao gồm Công ty TNHH Công Nghệ Robots Việt Nam và Công ty TNHH Sản Xuất và Thương Mại Robot Vina. Ngoài ra, các nhà sản xuất robot hàng đầu trên thế giới như Fanuc, ABB và Yaskawa cũng có mặt trên thị trường Việt Nam thông qua các đại lý và đối tác địa phương.

# LỜI CAM ĐOAN

Tôi xin cam đoan rằng đề tài *Mô hình Robot phân loại sản phẩm* được tiến hành một cách minh bạch, công khai. Mọi thứ được dựa trên sự cố gắng cũng như sự nỗ lực của bản thân cùng với sự giúp đỡ không nhỏ từ Ts. Phạm Hoàng Anh và Thầy Phạm Công Thái

Các số liệu và kết quả nghiên cứu được đưa ra trong đồ án là trung thực và không sao chép hay sử dụng kết quả của bất kỳ đề tài nghiên cứu nào tương tự. Nếu như phát hiện rằng có sự sao chép kết quả nghiên cứu đề những đề tài khác bản thân tôi xin chịu hoàn toàn trách nhiệm.

# LỜI CẢM ƠN

# MỤC LỤC

# THUẬT NGỮ VÀ TỪ VIẾT TẮT

# MỞ ĐẦU

## Giới thiệu

Trong bối cảnh của ngành công nghiệp sản xuất hiện đại, việc tự động hóa quy trình phân loại sản phẩm trở nên ngày càng quan trọng. Dự án nhằm phát triển một mô hình robot có khả năng phân loại sản phẩm một cách tự động để cải thiện hiệu suất và giảm chi phí cho các nhà sản xuất.

Robot được thiết kế để tự động phát hiện và phân loại các loại sản phẩm khác nhau bằng cách sử dụng công nghệ thị giác máy tính (computer vision). Điều này cho phép robot nhận biết và phân loại sản phẩm dựa trên các đặc điểm hình dáng và màu sắc của chúng. Để thực hiện các nhiệm vụ này, robot sử dụng các vi điều khiển thuộc dòng ESP32 series, đảm bảo tính linh hoạt và hiệu suất trong việc xử lý dữ liệu và điều khiển các hành động của robot.

Không chỉ có khả năng phân loại sản phẩm, robot cũng được trang bị khả năng tự hành. Điều này cho phép robot tự điều hướng và thực hiện các nhiệm vụ phân loại sản phẩm một cách hiệu quả và đáng tin cậy. Bằng cách này, robot có thể hoạt động một cách độc lập và hiệu quả mà không cần sự can thiệp của con người.

Bên cạnh đó, với khả năng vận chuyển hàng hóa trên khoảng cách xa, robot trở thành một giải pháp lý tưởng cho các nhà kho hoặc nhà máy lớn. Sự kết hợp giữa khả năng phân loại sản phẩm và khả năng tự hành của robot giúp nâng cao hiệu suất và tăng cường tính linh hoạt trong quy trình vận chuyển hàng hóa, đặc biệt là trong môi trường công nghiệp đòi hỏi sự chính xác và hiệu quả.

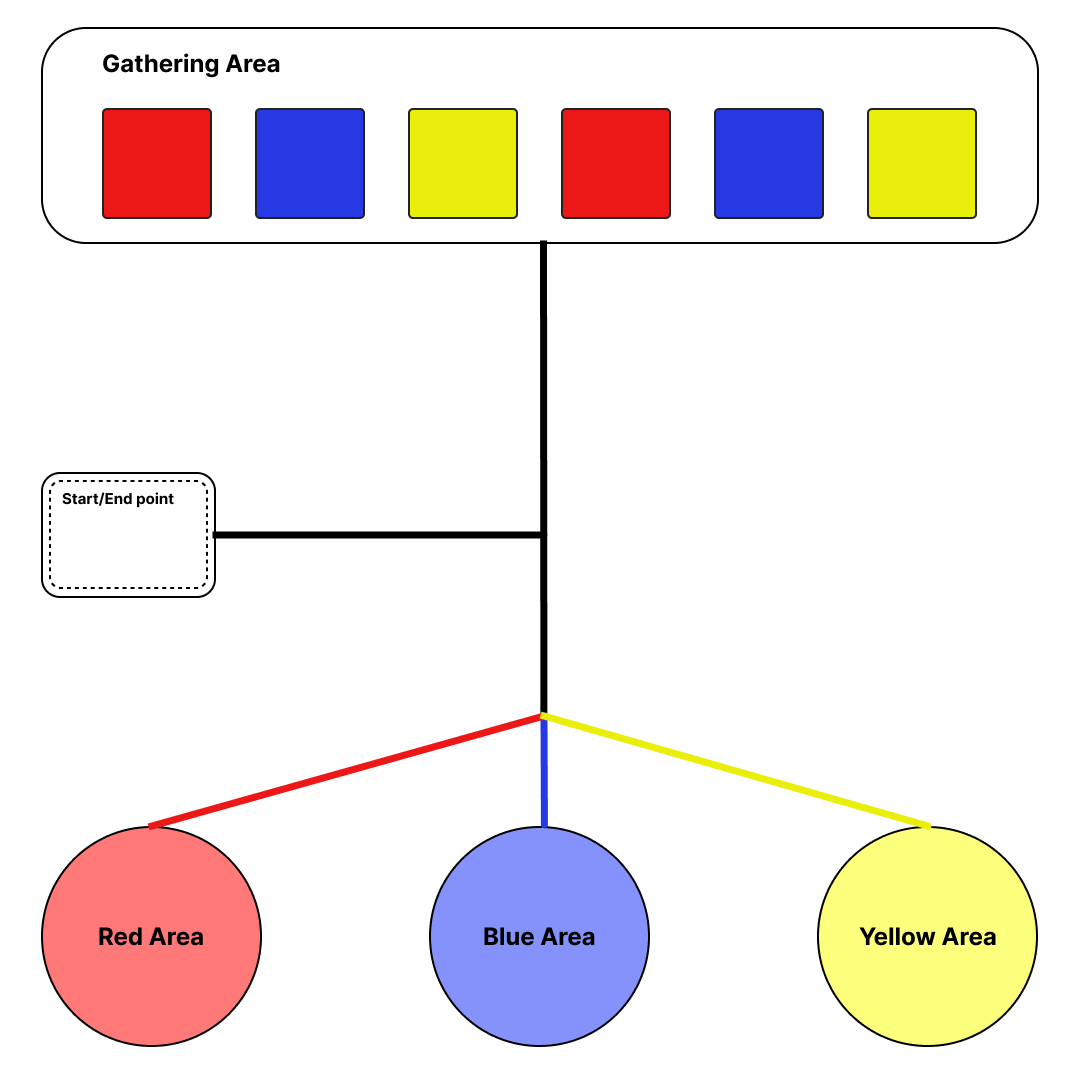
Robot phân loại sản phẩm mà em phát triển có những điểm khác biệt so với các loại robot phân loại phổ biến hiện nay. Thay vì sử dụng cánh tay robot, em đã áp dụng một đầu kẹp để cầm nắm hàng hóa, tạo điểm nhấn độc đáo và giảm độ phức tập trong quy trình thiết kế. Đồng thời, robot của em không được thiết lập cố định vị trí mà có thể di chuyển trên những tuyến đường được vẽ sẵn, giúp tối ưu hóa quy trình làm việc và tăng khả năng sử dụng trong các môi trường công nghiệp đa dạng.

Robot có một số ưu điểm nổi bậc như có khả năng phân loại được nhiều loại hàng hóa nếu được training đầy đủ, đồng thời việc thiết lập và cài đặt robot cũng rất dễ dàng. Chỉ cần training cho robot các loại hàng hóa cần phân loại, các vị trí của các khu vực phân loại và vẽ đường đi cho robot là có thể sử dụng được. Hơn nữa, khả năng vận chuyển ở khoảng cách xa của robot là một điểm mạnh lớn, làm cho nó trở thành lựa chọn phù hợp cho các nhà kho hoặc nhà máy lớn.

Tuy nhiên, nhược điểm của robot cũng không thể bỏ qua. Độ chính xác không cao, dễ phân loại sai hoặc di chuyển nhầm đường là một trong những thách thức lớn đối với robot của em. Vấn đề về năng lượng cũng là một hạn chế đáng lưu ý, khi robot sử dụng pin để hoạt động và không thể hoạt động liên tục. Cuối cùng, cần có thời gian training lại khi có hàng hóa mới, tạo ra sự bất tiện cho quy trình sản xuất và vận hành của nhà máy hay nhà kho.

## Mục tiêu của đề tài

### Mục tiêu chung

Mục tiêu chung của đề tài là nghiên cứu, thiết kế và hiện thực một Robot phân loại sản phẩm và vận chuyển sản phẩm đến khu vực được chỉ định thỏa mãn yêu cầu của Test case sau:

Hình . Mô hình của test case

Trong đó:

|  |  |  |
| --- | --- | --- |
| Hình ảnh | Tên gọi | Mô tả |
|  | Khối vuông (Cube) | Những khối vuông tượng trưng cho sản phẩm cần được phân loại  Sẽ chỉ có 3 loại khối vuông chính là xanh, đỏ và vàng |
|  | Khu vực phân loại (Classified area) | Đây là các khu vực mà Robot cần phải đưa các khối vuông phía trên vào hay nói cách khác là các khu vực chứa các sản phẩm đã được phân loại.  Trong test case này, có 3 khu vực phân loại lần lượt là Blue area, Red area và Yellow area. |
|  | Khu vực tập kết (Gathering area) | Đây là khu vực tập kết các khối vuông.  Các khối vuông được đặt trong khu vực này để sẵn sàng cho việc phân loại. |
|  | Điểm bắt đầu/ kết thúc (Start/End point) | Đây là nơi robot bắt đầu cũng như kết thúc chu trình làm việc của mình.  Khu vực này có thể được tích hợp cổng sạc để sạc cho Robot khi Robot gần hết năng lượng. |
|  | Đường kẻ (Line) | Những đường kẻ được vẽ sẵn để kết nối các khu vực với nhau.  Robot dựa vào những đường vẽ này để xác định vị trí của các khu vực. |

Bảng Miêu các thành phần trong test case

**Yêu cầu:**

Ban đầu, các khối vuông được đặt trong *Gathering area*. Từ *Start point,* Robot sẽ xuất phát và tiến đến *Gethering area.* Sau đó, Robot tiến hành nhận diện và phân loại các khối vuông rồi vận chuyển các khối vuông đến các khu vực phân loại *(Classified Area)* tương ứng với màu sắc của từng khối vuông. Khi đã phân loại hết tất cả các khối vuông trong *Gethering area*, Robot phải quay lại điểm *End point*.

**Thông tin đầu vào:**

1. Các sản phẩm được sắp xếp thẳng hàng với nhau trong khu vực Gathering. Các sản phẩm được tượng trưng bằng các khối vuông khác màu nhau.
2. Ba khu vực dành riêng cho các loại sản phẩm khác nhau. Các khu vực này là các khu vực để chứa các khối vuông theo màu sắc.
3. Các đường kẻ được vẽ sẳn để Robot có thể lựa chọn tuyến đường đến các khu vực.
4. Khu vực khởi đầu và kết thúc. Nơi đây sẽ là nơi bắt đầu cũng như kết thúc chu trình làm việc của Robot. Trong thực tế thì đây có thể là trạm sạc pin cho Robot.

**Kết quả mong muốn:**

1. Các khối vuông trong khu vực Gathering được vận chuyển và đặt tại các khu vực tương ứng với màu sắc của từng khối vuông.
2. Robot quay trở lại khu vực xuất phát.

**Các giới hạn:**

**Nhiệm vụ của Robot:**

1. Tự động phát hiện được các sản phẩm,
2. Phân loại được loại sản phẩm,
3. Di chuyển đến vị trí sản phẩm để lấy sản phẩm
4. Vận chuyện sản phẩm đó đến khu vực phân loại tương ứng,
5. Lặp lại các bước trên cho đến khi hết tất cả sản phẩm. Sau đó quay lại điểm bắt đầu.  
   Như vậy, ta có thể thấy Robot có những chức năng chung như sau:

* Phát hiện các sản phẩm khác nhau trong khu vực tập kết (Gathering Area) và phân loại theo thành các nhóm theo tiêu chí được định nghĩa trước.
* Xác định vị trí tương đối của sản phẩm trên khung hình của camera của từng sản phẩm.
* Sử dụng cánh tay robot để gấp và thả sản phẩm.
* Di chuyển theo những tuyến đường được thiết kế sẵn.

### Mục tiêu cụ thể

#### Về phần cứng

* Thiết kế và chế tạo khung cho robot chắc chắn, chịu lực tốt và di chuyển linh hoạt trên địa hình phẳng,
* Tính toán và xác định vị trí đặt camera trên robot sao cho có tầm nhìn bao quát có thể nhìn thất các vật thể và cánh tay của Robot,
* Lựa chọn và tích hợp các cảm biến, động cơ và linh kiện phù hợp cho các chức năng đã được nêu,
* Thiết kế và chế tạo cánh tay robot có độ chính xác cao. Xác định chiều dài hoạt động của cánh tay, độ rộng khi mở và đóng của bàn tay sao cho phù hợp với hầu hết các kích thước của sản phẩm.

#### Về phần mềm

* Phát triển thuật toán thị giác máy tính:
  + Xác định được sản phẩm,
  + Phân loại sản phẩm theo các tiêu chí đã được định nghĩa,
  + Vẽ được khung bao quanh sản phẩm và xác định được vị trí tương đối của vật thể trên khung hình của camera,
* Lập trình phần mềm điều khiển robot:
  + Điều khiển di chuyển robot trên địa hình, phát hiện đường kẻ và di chuyển theo nó,
  + Điều khiển di chuyển camera,
  + Điều khiển hoạt động của cánh tay robot,
  + Giao tiếp giữa ESP32 WROOM 32 và ESP32 CAM.

#### Thử nghiệm và đánh giá

* Chuẩn bị bộ dataset dành cho việc training mô hình AI theo các tiêu chí,
* Xây dưng môi trường thử nghiệm,
* Đánh giá độ chính xác của mô hình AI với các sản phẩm khác nhau,
* Chạy thử và đánh giá robot trên môi trường thử nghiệm,
* Đề xuất các giải pháp khắc phục các trường hợp lỗi.

## Ý nghĩa thực tiễn của đề tài

### Nâng cao hiệu quả và năng suất tự động hoá

Robot phân loại hàng hóa có thể làm việc liên tục 24/7 mà không cần nghỉ ngơi, thay thế sức lao động thủ công, giảm thiểu sai sót và tai nạn lao động.

Robot có thể phân loại sản phẩm với tốc độ và độ chính xác cao hơn con người, giúp tăng năng suất và hiệu quả hoạt động trong các kho hàng, nhà máy, trung tâm thương mại,...

### Giảm chi phí hoạt động

Sử dụng robot phân loại hàng hóa giúp tiết kiệm chi phí nhân công, chi phí bảo trì và chi phí sửa chữa so với hệ thống phân loại thủ công.

Robot có thể hoạt động trong môi trường khắc nghiệt, nguy hiểm, giảm thiểu chi phí bảo hộ lao động và chi phí y tế.

### Tăng cường khả năng cạnh tranh

Doanh nghiệp áp dụng robot phân loại hàng hóa sẽ có lợi thế cạnh tranh về hiệu quả, năng suất và chi phí so với các doanh nghiệp sử dụng phương pháp phân loại thủ công.

Robot có thể giúp doanh nghiệp đáp ứng nhu cầu thị trường ngày càng cao về tốc độ và độ chính xác trong việc phân loại hàng hóa.

### Mở ra tiềm năng ứng dụng rộng rãi

Robot phân loại hàng hóa có thể được ứng dụng trong nhiều lĩnh vực khác nhau như logistics, sản xuất, thương mại điện tử, nông nghiệp,...

Với sự phát triển của công nghệ trí tuệ nhân tạo và thị giác máy tính, robot phân loại hàng hóa sẽ ngày càng thông minh và linh hoạt hơn, mở ra nhiều tiềm năng ứng dụng mới trong tương lai.

Tóm lại, việc phát triển robot phân loại hàng hóa là một đề tài có ý nghĩa thực tiễn quan trọng, góp phần nâng cao hiệu quả, năng suất, giảm chi phí và tăng cường khả năng cạnh tranh cho doanh nghiệp. Với tiềm năng ứng dụng rộng rãi, robot phân loại hàng hóa hứa hẹn sẽ đóng vai trò quan trọng trong việc tối ưu hóa hoạt động logistics và sản xuất trong tương lai.

## Đối tượng nghiên cứu và phạm vi giới hạn của đề tài

Phần nội dung của đề tài tập trung vào nghiên cứu, thiết kế và hiện thực Robot có khả năng giải quyết được bài toán đã đề ra trong Test case của Mục tiêu đề tài. Cụ thể, phạm vi nghiên cứu bao gồm 5 đối tượng chính sau:

### Khung Robot:

Nghiên cứu và thiết kế một khung cơ bản cho Robot, bao gồm cấu trúc cơ học và kết nối các thành phần điện tử.

### Esp32 wroom 32 và esp32 cam:

Sử dụng vi điều khiển ESP32 WROOM 32 và ESP32 CAM để điều khiển và giao tiếp với các thiết bị và cảm biến khác nhau trên Robot.

### Computer vision:

Nghiên cứu và triển khai các thuật toán phát hiện, nhận diện và phân loại vật thể sử dụng công nghệ computer vision, nhằm hỗ trợ quá trình phân loại sản phẩm của Robot.

### Robot dò line:

Thiết kế và hiện thực chức năng dò line cho Robot, giúp nó có khả năng di chuyển trên các tuyến đường được vẽ sẵn và đến được đích một cách chính xác.

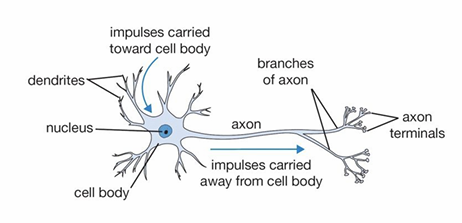
### Thiết kế và nguyên lý hoạt động của cánh tay robot:

Nghiên cứu và thiết kế cánh tay robot, bao gồm cơ chế hoạt động, khả năng cầm nắm và di chuyển sản phẩm, đảm bảo tính linh hoạt và chính xác trong quá trình phân loại.

# CƠ SỞ LÝ THUYẾT

## Deep Learning

### Giới thiệu chung

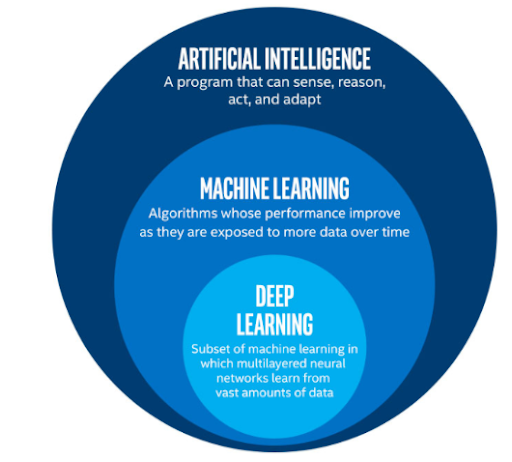
****Học sâu (Deep Learning) là một lĩnh vực con của Trí tuệ Nhân tạo (AI) lấy cảm hứng từ cấu trúc và hoạt động của não bộ con người.

Nó sử dụng các mạng nơ-ron nhân tạo (ANN) được xây dựng từ các "nơ-ron nhân tạo" đơn giản, mô phỏng hoạt động của các tế bào thần kinh.

Hình . Cấu trúc của nơ-ron

Mỗi nơ-ron nhận nhiều tín hiệu đầu vào, xử lý chúng thông qua một hàm kích hoạt, và tạo ra một tín hiệu đầu ra. Các nơ-ron được sắp xếp thành nhiều lớp, với mỗi lớp thực hiện một chức năng cụ thể.

So với các phương pháp học máy truyền thống, Deep Learning có khả năng học hỏi từ dữ liệu phức tạp và phi tuyến tính tốt hơn, mang lại hiệu quả cao trong nhiều lĩnh vực như xử lý ảnh, nhận dạng giọng nói, dịch tự động và nhiều hơn nữa. Tuy nhiên, Deep Learning cũng đòi hỏi lượng dữ liệu lớn và tài nguyên tính toán mạnh mẽ để huấn luyện mô hình hiệu quả.



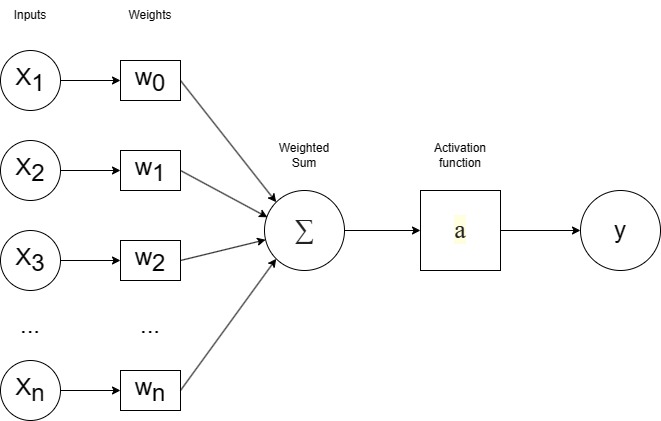
Hình . AI, Machine Learning và Deep Learning

### Artificial Neural Networks (ANNs)

#### Perceptron

Perceptron là thành phần cơ bản nhất cấu thành nên một mạng neural. Chúng đơn giản là các hàm toán học nhận đầu vào từ một hoặc nhiều số, thực hiện các phép toán và trả về kết quả đầu ra.

Có thể hình dung hoạt động của perceptron trong hình vẽ sau:



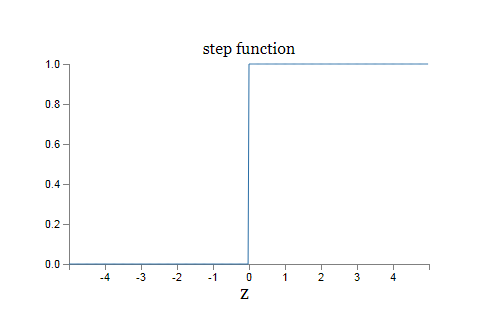
Hình . Cấu trúc của một Perceptron

Trong đó,

* Hàm này áp dụng phép biến đổi phi tuyến tính cho tổng có trọng số và độ lệch (bias ), tạo ra kết quả đầu ra.
* :Kết quả của quá trình

Các hàm kích hoạt (Activation function) phổ biến:

1. **Hàm Step:**



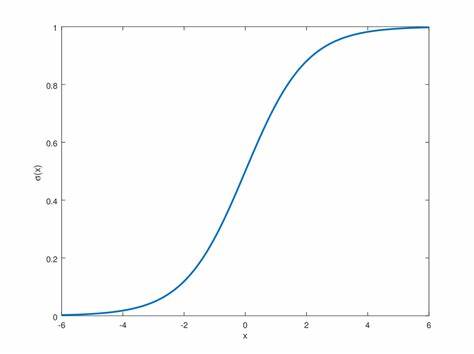
Hình . Hàm step

1. **Hàm RELU (Rectified Linear Unit):**



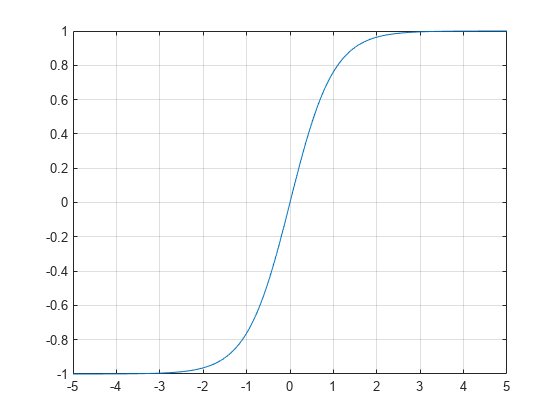
Hình . Hàm RELU

1. **Hàm Sigmoid:**



Hình . Hàm Sigmoid

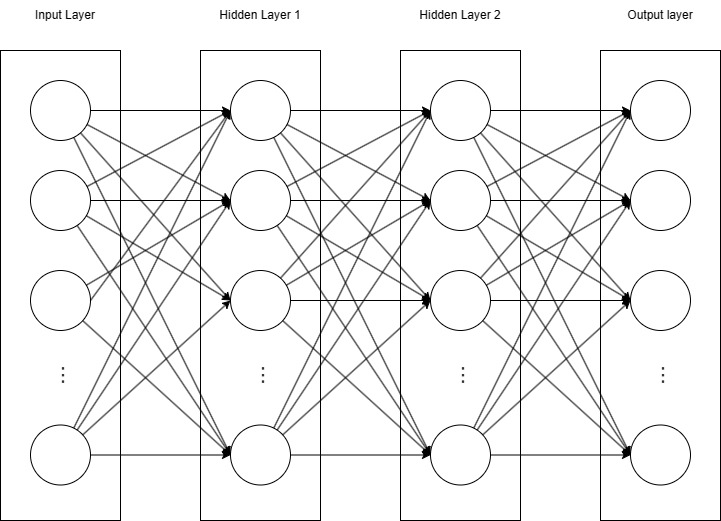
1. **Hàm Tanh:**



Hình . Hàm Tanh

#### Network Structure

Mạng neural là sự kết hợp của của các tầng perceptron hay còn được gọi là perceptron đa tầng (multilayer perceptron) như hình vẽ bên dưới:



Hình . Cấu trúc mạng neural

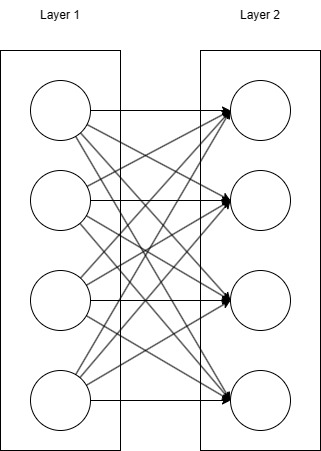
Một mạng neural sẽ có 3 kiểu tầng:

* Tầng vào (Input layer): Là tầng bên trái cùng của mạng thể hiện cho các đầu vào của mạng.
* Tầng ra (Output layer): Là tầng bên phải cùng của mạng thể hiện cho các đầu ra của mạng.
* Tầng ẩn (Hidden layer): Là tầng nằm giữa tầng vào và tầng ra thể hiện cho việc suy luận logic của mạng. Lưu ý rằng, một NN chỉ có 1 tầng vào và 1 tầng ra nhưng có thể có nhiều tầng ẩn.

Số lượng của một mạng là không cố định. Nó được ký hiệu với:

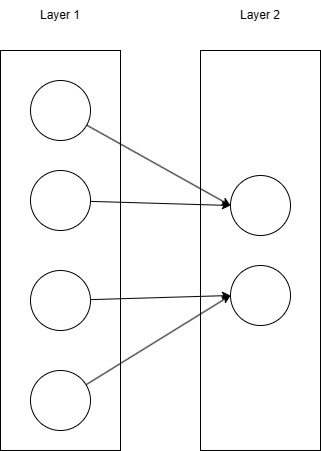
Khi tính số tầng trong một mạng neural, ta sẽ không tính tầng Input.

Các neural trong một lớp không được kết nối trực tiếp. Thông tin truyền một chiều từ lớp này sang lớp tiếp theo, với mỗi nơ-ron trong lớp nhận đầu vào từ lớp trước và gửi đầu ra của nó đến lớp tiếp theo. Có 2 kiểu kết nối chính:  
Fully-connected: mỗi neural trong cùng một tầng sẽ kết nối đôi một với các neural với tầng tiếp theo.



Hình . Fully Connected

Pooling: Khác với fully-connected, pooling chỉ cho phép một vài neural trong cùng một tầng kết nối với một neural của tầng tiếp theo. Điều này sẽ giảm số lượng neural ở layer đó.



Hình . Pooling

#### Network Architectures

Có một số kiến trúc khác nhau cho các mạng neural, mỗi kiến trúc có điểm mạnh và điểm yếu riêng. Một số kiến trúc phổ biến nhất bao gồm:

* Mạng neural tiến nguồn (Feedforward Neural Networks): Đây là loại kiến trúc ANN đơn giản nhất, trong đó thông tin chảy theo một hướng từ đầu vào đến đầu ra. Các lớp được kết nối đầy đủ, có nghĩa là mỗi neural trong một lớp được kết nối với tất cả các neural trong lớp tiếp theo.
* Mạng neural hồi quy (Recurrent Neural Networks - RNNs): Những mạng này có một thành phần "bộ nhớ", nơi thông tin có thể chảy trong chu kỳ qua mạng. Điều này cho phép mạng xử lý các chuỗi dữ liệu, như chuỗi thời gian hoặc âm thanh.
* Mạng neural tích chập (Convolutional Neural Networks - CNNs): Những mạng này được thiết kế để xử lý dữ liệu có một định dạng lưới, như hình ảnh. Các lớp bao gồm các lớp tích chập, có nhiệm vụ học để phát hiện các đặc điểm cụ thể trong dữ liệu, và các lớp gộp, giảm kích thước không gian của dữ liệu. Đây là kiến trúc chính cần tìm hiểu để để giải quyết bài toán trong đề tài lần này.
* Bộ mã hóa tự động (Autoencoders): Đây là các mạng neural được sử dụng cho việc học không giám sát. Chúng bao gồm một bộ mã hóa chuyển dữ liệu đầu vào thành một biểu diễn chiều thấp hơn và một bộ giải mã chuyển biểu diễn trở lại dữ liệu ban đầu.
* Mạng neural sinh đối nghịch (Generative Adversarial Networks - GANs): Đây là các mạng neural được sử dụng cho việc mô hình hóa sinh. Chúng bao gồm hai phần: một bộ tạo ra học để tạo ra các mẫu dữ liệu mới, và một bộ phân biệt học để phân biệt giữa dữ liệu thực và dữ liệu được tạo ra.

#### Lan truyền tuyến (Forward Propagation)

Như bạn thấy thì tất cả các nốt mạng (nơ-ron) được kết hợp đôi một với nhau theo một chiều duy nhất từ tầng vào tới tầng ra. Tức là mỗi nốt ở một tầng nào đó sẽ nhận đầu vào là tất cả các nốt ở tầng trước đó mà không suy luận ngược lại. Hay nói cách khác, việc suy luận trong mạng NN là suy luận tiến (feedforward):

Trong đó:

* : Tầng hiện tại . Quy ước sử dụng và để chỉ tầng inputs và output.
* : Số lượng node tại tầng hiện tại
* : Số lượng node tại tầng trước.
* : Node thứ của tầng hiện tại,
* : Node thứ của tầng trước đó,
* : Ví dụ huấn luyện (training example) hiện tại với là số lượng ví dụ huấn luyện.
* : Tổng trọng số của các lần kích hoạt của lớp trước đó, được dịch chuyển theo độ lệch (bias).
* : Trọng số cân bằng lần kích hoạt thứ của tầng trước.
* : Giá trị bias của tầng hiện tại
* : lần kích hoạt tại tầng hiện tại.
* : lần kích hoạt tại tầng trước.
* : hàm kích hoạt được sử dụng ở tầng hiện tại

Nói một cách ngắn gọn, một node trong lớp hiện tại phụ thuộc vào mọi node ở tầng trước và hình ảnh trực quan sau đây có thể giúp chúng ta thấy điều đó rõ ràng hơn:

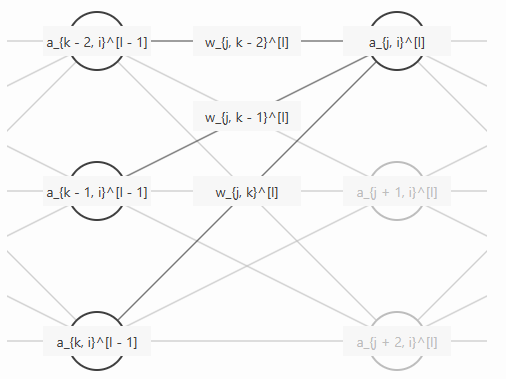


Figure Một node tại tầng hiện tại

Bên cạnh đó, một node ở lớp trước sẽ ảnh hưởng đến mọi node trong lớp hiện tại và với sự thay đổi về phần đánh dấu, chúng ta cũng có thể thấy điều đó rõ ràng hơn:

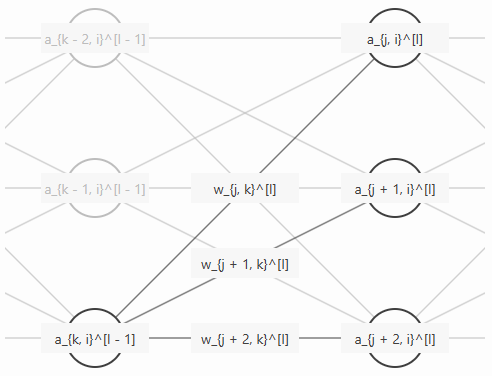


Figure Một node ở tầng trước đó

Như vây, ta có thể vector hoá các nút như sau:

,

Hoặc ta có thể viết như sau:

Trong đó , và cuối cùng, . Dấu 2 chấm ở đây dùng dể làm rõ rằng là cột thứ của . Tương tự với những cái khác.

Tiếp theo, chúng ta vector hoá các ví dụ huấn luyện:

Trong đó, , , và .

Ta có thể lập hai ký hiệu bổ sung sau:

Trong đó, là ma trận của tầng inputs và để chỉ ma trận tầng dự đoán/ outputs.

Cuối cùng, ta có thể định nghĩa hàm chi phí (cost function) như sau:

as

Trong đó, để chỉ ma trận mục tiêu (the targets) và có thể được điều chỉnh theo nhu cầu của bài toán.

#### Lan truyền ngược (Backpropagation)

Lan truyển ngược là quá trình giúp điều chỉnh các trọng số và bias của các node trong mạng neural để sai lệch giữa đầu ra dự đoán và đầu ra thực tế đạt tối thiểu.

Phương pháp phổ biến nhất để tối ưu mạng neural đa tầng chính là gradient descent (GD). Để áp dụng GD, chúng ta cần tính được gradient của hàm mất mát theo từng ma trận trọng số vector điều chỉnh (bias).

Giả sử là hàm mất mát của bài toán, trong đó , là tập hợp tất cả các ma trận trọng số và vector điều chỉnh. , là cặp dữ liệu huấn luyện với mội cột tương ứng. Để có thể áp dng các phương pháp gradient descent, chúng ta cần tính được các ; ,

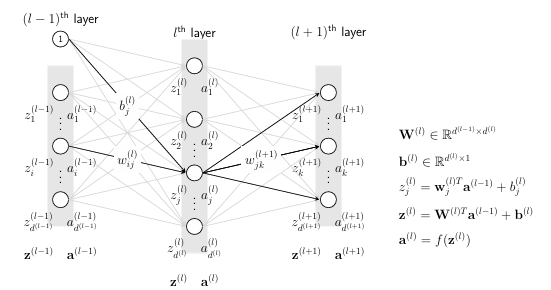
Xét ví dụ của hàm mất mát là hàm sai số trung bình bình phương (MSE):

Với là số cặp dữ liệu trong tập huấn luyện. Theo các công thức này, việc tính toán trực tiếp các giá trị gradient tương đối phức tạp vì hàm mất mát không phụ thuộc trực tiếp vào các ma trận trọng số và vector điều chỉnh. Phương pháp phổ biến nhất được sử dụng là *lan truyền ngược* (backpropagation), giúp tính gradient ngược từ tầng cuối cùng đến tầng đầu tiên. Tầng cuối cùng được tính toán trước vì nó ảnh hưởng trực tiếp tới đầu ra dự đoán và hàm mất mát. Việc tính toán gradient của các ma trận trọng số trong các tầng trước được thực hiện dựa trên nguyên tắc chuỗi quen thuộc cho gradient của hàm hợp.

Thuật toán Stochastic Gradient Descent thường được sử dụng để cập nhật các ma trận trọng số và vector điều chỉnh dựa trên một cặp điểm dữ liệu huấn luyện . Đơn giản hơn, ta coi là hàm mất mát nếu chỉ xét cặp điểm này. Ở đây là hàm mất mát bất kỳ, không chỉ hàm MSE như ở trên. Đạo hàm riêng của hàm mất mát theo chỉ một thành phần của ma trận trọng số của tầng đầu ra:

Trong đó thường là một đại lượng không quá khó để tính toán và vì . Tương tự, gradient của hàm mất mát theo hệ số tự do của tầng cuối cùng là

Với đạo hàm riêng theo trọng số ở các tầng , hãy quan sát hình minh hoạ bên dưới:



Hình Mô phỏng cách tính lan truyền ngược. Tầng cuối có thể là tầng đầu ra.

Ở đây, tại các nút, đầu vào z và đầu ra a được viết riêng để tiện theo dõi. Dựa vào đó, bằng quy nạp ngược từ cuối, ta có thể tính được:

Với:

Trong đó, và được hiểu là hàng thứ của ma trận (chú ý dấu hai chấ, khi không có dấu này, chúng ta mặc định dùng nó để ký hiệu cho vector cột). Dấu ∑ tính tổng ở dóng thứ hai trong phép tính trên xuất hiệ vì đóng góp vào việc tính tất cả các . Biểu thức đạo hàm ngoài dấu ngoặc lớn xuất hiện vì . Tới đây, ta có thể thấy rằng việc hàm kích hoạt có đạo hàm đơn giản sẽ có ích rất nhiều trong việc tính toán. Với cách làm tương tự, ta có thể suy ra

Nhận thấy rằng trong những công thức trên, việc tính các đóng một vai trò quan trọng. Hơn nữa, để tính được giá trị này, ta cần tính được các . Nói cách khác, ta cần tính ngược các giá trị này từ tầng cuối cùng. Tên gọi *lan truyền ngược* xuất phát từ đây.

Sau khi tìm được giá trị của và , ta cập nhật giá trị trọng số và bias của các node theo thuật toán Gradient Descent theo công thức sau:

Với là một số dương được gọi là tốc độ học (learning rate). Dấu trừ thể hiện việc và cần đi ngược lại đạo hàm riêng tương ứng là và .

## Convolutional Neural Networks (CNNs)

### Giới thiệu

Mạng neural Nhân tạo (ANN) đã tạo nên cuộc cách mạng trong nhiều lĩnh vực nhờ khả năng học hỏi các mẫu phức tạp từ dữ liệu. Tuy nhiên, các kiến trúc ANN truyền thống lại gặp khó khăn trong việc xử lý hiệu quả dữ liệu dạng lưới, chẳng hạn như hình ảnh, nơi mối quan hệ giữa các pixel đóng vai trò quan trọng. Hạn chế này cản trở hiệu suất của chúng trong các nhiệm vụ như phân loại đối tượng, trong đó việc xác định và phân biệt các đối tượng trong ảnh là điều cần thiết.

Mạng neural tích chập (CNN) giải quyết thách thức này bằng cách giới thiệu một kiến trúc chuyên biệt được thiết kế cho xử lý ảnh và phân loại đối tượng. Lấy cảm hứng từ cấu trúc sinh học của vỏ não thị giác, CNN tận dụng các lớp tích chập để trích xuất các đặc trưng không gian từ hình ảnh. Các lớp này sử dụng các bộ lọc (kernels) di chuyển trên ảnh đầu vào, phát hiện các cạnh, đường thẳng và các đặc trưng cấp thấp khác trong các vùng cụ thể. Khi mạng đi qua nhiều lớp tích chập hơn, nó học cách kết hợp các đặc trưng cấp thấp này thành các đặc trưng cấp cao phức tạp hơn, có liên quan đến việc xác định đối tượng. Khả năng trích xuất đặc trưng theo thứ bậc này cho phép CNN vượt trội trong các nhiệm vụ như phân loại đối tượng, trong đó việc nhận biết các đối tượng bên trong ảnh đòi hỏi phải hiểu các thành phần cấu thành và mối quan hệ không gian của chúng.

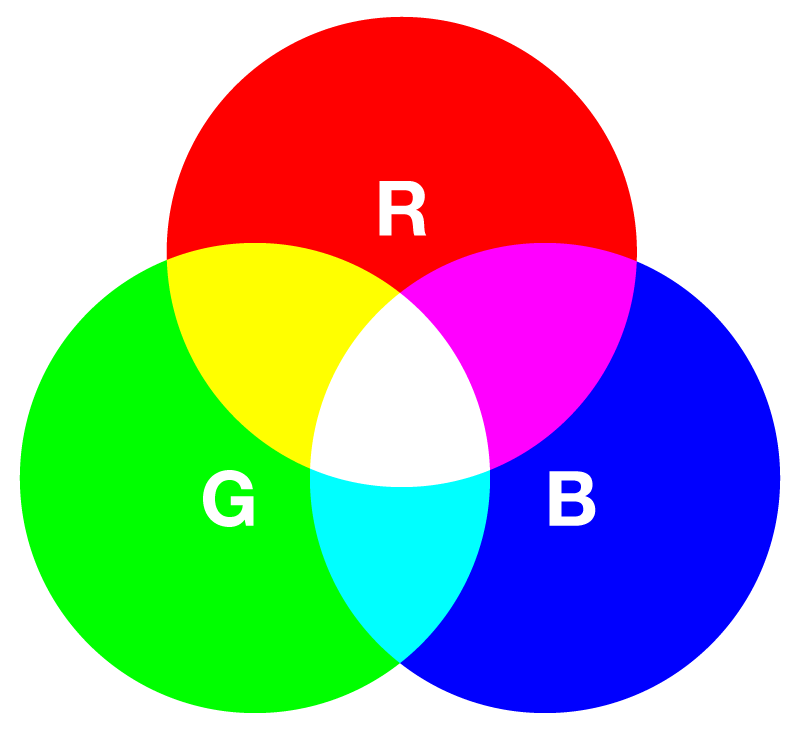
So với ANN truyền thống, CNN mang lại một số lợi thế cho việc phân loại đối tượng.

* Thứ nhất, chúng loại bỏ nhu cầu thiết kế thủ công các đặc trưng, một quá trình tốn thời gian và phụ thuộc vào lĩnh vực cụ thể. Bằng cách tự động học các đặc trưng thông qua phép tích chập, CNN có thể thích ứng với nhiều lớp đối tượng và độ phức tạp của hình ảnh khác nhau.
* Thứ hai, CNN tận dụng việc chia sẻ tham số, trong đó một bộ lọc duy nhất được áp dụng trên các phần khác nhau của ảnh. Điều này giúp giảm số lượng tham số cần học so với các mạng lưới đầy đủ cùng kích thước, cải thiện hiệu quả đào tạo và giảm nguy cơ overfitting[[4]](#footnote-4). Cuối cùng, kết nối cục bộ trong các lớp tích chập thúc đẩy tính bất biến tịnh tiến, một đặc tính quan trọng cho phân loại đối tượng. Điều này có nghĩa là mạng có thể nhận ra các đối tượng bất kể vị trí của chúng trong ảnh, giúp nó thích ứng với các biến thể trong bố cục hình ảnh.

### Các khái niệm cơ bản

#### Ảnh trên máy tính

Thông thường, máy tính sử dụng hệ màu RGB[[5]](#footnote-5). Khi trộn ba màu trên theo tỉ lệ nhất định có thể tạo thành các màu khác nhau.



Hình Hệ màu RGB

Như vậy, với một màu cho sẵn, ta có thể tìm được bộ ba giá trị của màu đó. Với mỗi bộ 3 số r, g, b nguyên trong khoảng [0, 255] sẽ cho ra một màu khác nhau. Do có 256 cách chọn r, 256 cách chọn màu g, 256 cách chọn b => tổng số màu có thể tạo ra bằng hệ màu RGB là: 256 \* 256 \* 256 = 16777216 màu.

Ảnh trên máy tính được chia làm 2 loại là ảnh màu và ảnh xám.

* **Ảnh màu**

Ảnh màu là một ma trận các pixel mà mỗi pixel biểu diễn một điểm màu. Mỗi điểm màu được biểu diễn bằng bộ 3 số (r,g,b). Như vậy, ta có thể vector hoá một bức ảnh có kích thức với là chiều rộng và chiều cao của bức ảnh:

* **Ảnh xám**

#### Tensor

#### Chuyển hệ màu của ảnh

#### Kernel và phép tính convolution

#### Padding

#### Stride

#### Flatten

### Convolutional layers

y11 = b+(x111 w111+x121 w121+x131 w131+x211 w211+x221 w221+x231 w231+x311 w311+x321 w321+x331 w331)+(x112 w112+x122 w122+x132 w132+x212 w212+x222 w222+ x232 w232+x312 w312+x322 w322+x332 w332)+(x113 w113+x123 w123+x133 w133+x213 w213 +x223 w223+x233 w233+x313 w313+x323 w323+x333 w333)= 25

### Pooling layers

### Fully connected layers

### CNN Architectures for Object Classification

### Transfer Learning

## Computer Vision

### Digital Image Fundamentals

### Image Segmentation

### Object Detection and Recognition

### Image Classification

# PHƯƠNG PHÁP NGHIÊN CỨU

*Xem 5 phương pháp nghiên cứu thực tiễn*

# KẾT QUẢ

# BÀN LUẬN VÀ KIẾN NGHỊ

# TÀI LIỆU THAM KHẢO

Logistics Việt Nam (11/07/2022). Tìm hiểu các loại chi phí lưu kho trong doanh nghiệp, truy cập 27/02/2024, từ <https://alslogistics.vn/tim-hieu-cac-loai-chi-phi-luu-kho-trong-doanh-nghiep/>

Mordor Intelligence (2024). Robotics Industry Size & Share Analysis - Growth Trends & Forecasts (2024 - 2029), truy cập 27/02/2022 từ <https://www.mordorintelligence.com/industry-reports/robotics-market>

Interational Federation of Robotics (2020). IFR presents World Robotics Report 2020, truy cập 27/02/2024, từ <https://ifr.org/ifr-press-releases/news/record-2.7-million-robots-work-in-factories-around-the-globe>

1. Tìm hiểu các loại chi phí lưu kho trong doanh nghiệp - Logistics Việt Nam – 11/07/2022 trên alslogistics.vn [↑](#footnote-ref-1)
2. IFR presents World Robotics Report 2020 - https://ifr.org/ifr-press-releases/news/record-2.7-million-robots-work-in-factories-around-the-globe [↑](#footnote-ref-2)
3. Robotics Industry Size & Share Analysis - Growth Trends & Forecasts (2024 - 2029) -   
   Source: https://www.mordorintelligence.com/industry-reports/robotics-market  
    [↑](#footnote-ref-3)
4. Overfitting (hay quá khớp) là hiện tượng khi mô hình xây dựng thể hiện được chi tiết bộ dữ liệu huấn luyện. Điều này có nghĩa là cả dữ liệu nhiễu, hoặc dữ liệu bất thường trong tập huấn luyện đều được chọn và học để đưa ra quy luật mô hình. Những quy luật này sẽ không có ý nghĩa nhiều khi áp dụng với bộ dữ liệu mới có thể có dạng dữ liệu nhiễu khác. Khi đó, nó ảnh hưởng tiêu cực tới độ chính xác của mô hình nói chung. [↑](#footnote-ref-4)
5. RGB là viết tắt của Red (đỏ), Green (xanh lục), Blue (xanh lam), là ba màu chính của ánh sáng khi tách ra từ lăng kính [↑](#footnote-ref-5)