**TRƯỜNG ĐẠI HỌC GIAO THÔNG VẬN TẢI**

**PHÂN HIỆU TẠI TP. HỒ CHÍ MINH**

**BỘ MÔN CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO ĐỒ ÁN TỐT NGHIỆP**

**ĐỀ TÀI: NGHIÊN CỨU TRIỂN KHAI THUẬT TOÁN YOLOV3** **VÀ XÂY DỰNG HỆ THỐNG HƯỚNG DẪN ĐẬU XE Ô TÔ PARKING VISION**

Giảng viên hướng dẫn: ThS. PHẠM THỊ MIÊN

Sinh viên thực hiện : NGUYỄN VĂN HIỆP

Mã sinh viên : 5851071024

Lớp : CÔNG NGHỆ THÔNG TIN

Khoá :58

TP. Hồ Chí Minh, tháng 06 năm 2021

**TRƯỜNG ĐẠI HỌC GIAO THÔNG VẬN TẢI**

**PHÂN HIỆU TẠI TP. HỒ CHÍ MINH**

**BỘ MÔN CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO ĐỒ ÁN TỐT NGHIỆP**

**ĐỀ TÀI:** **NGHIÊN CỨU TRIỂN KHAI THUẬT TOÁN YOLOV3 VÀ XÂY DỰNG HỆ THỐNG HƯỚNG DẪN ĐẬU XE Ô TÔ PARKING VISION**

Giảng viên hướng dẫn: ThS. PHẠM THỊ MIÊN

Sinh viên thực hiện : NGUYỄN VĂN HIỆP

Mã sinh viên : 5851071024

Lớp : CÔNG NGHỆ THÔNG TIN

Khoá :58

TP. Hồ Chí Minh, tháng 06 năm 2021

|  |  |
| --- | --- |
| TRƯỜNG ĐẠI HỌC GIAO THÔNG VẬN TẢI  PHÂN HIỆU TẠI TP. HỒ CHÍ MINH  **BỘ MÔN CÔNG NGHỆ THÔNG TIN**  --------------------------------------- | **CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM**  Độc lập – Tự do - Hạnh phúc  --------------------------------------------------- |

**NHIỆM VỤ THIẾT KẾ TỐT NGHIỆP**

Họ và tên sinh viên: NGUYỄN VĂN HIỆP

MSSV: 5851071024

Chuyên ngành: Công nghệ Thông tin

Lớp: CQ.CNTT.58

1. **Tên đề tài đồ án tốt nghiệp:**

***NGHIÊN CỨU TRIỂN KHAI THUẬT TOÁN YOLOV3 VÀ XÂY DỰNG HỆ THỐNG HƯỚNG DẪN ĐẬU XE Ô TÔ PARKING VISION.***

1. **Mục đích, yêu cầu:**
   1. **Mục đích:**

* Xây dựng website theo dõi các vị trí còn trống và đã đỗ trong bãi đỗ xe và hiển thị lên màn hình số vị trí còn trống, đã đỗ và tổng số lượng chỗ đỗ trong bãi đỗ xe ô tô theo thời gian thực.
* Xây dựng mô hình nhận diện các chỗ còn trống và đã đỗ trong bãi đỗ xe bằng hình ảnh, video, camera.
* Xây dựng website giới thiệu sản phẩm.
* Xây dựng giải pháp cho Hệ Thống Hướng Dẫn Bãi Đỗ Xe (Parking Guidance System) thông minh tiết kiệm, nhanh hơn, tiện dụng, chính xác và có thể mở rộng quy mô sau này.
  1. **Yêu cầu:**
* Tìm hiểu về Thị Giác Máy Tính (Computer Vision) và Học Sâu (Deep Learning).
* Nghiên cứu về xử lý ảnh.
* Nghiên cứu những quy trình trong xử lý ảnh.
* Nghiên cứu về Mạng Thần Kinh Tích Chập (Convolution Neural Network) và những ứng dụng của nó trong Deep Learning và Computer Vision.
* Nghiên cứu thuật toán YOLO (You Only Look Once).
* Tìm hiểu một số khái niệm liên quan đến lĩnh lực Khai Phá Dữ Liệu (Data Mining) áp dụng vào mô hình Học Sâu.
* Tìm hiểu các nguồn cơ sở dữ liệu hình ảnh cho quá trình huấn luyện mô hình Học Sâu.
* Thu thập dữ liệu hình ảnh về những chỗ trống và đã đỗ trong bãi đỗ xe. Gắn nhãn, tiền xử lý.
* Tìm hiểu về Transfer Learning, và ứng dụng vào huấn luyện mô hình.
* Ứng dụng kiến trúc Darknet huấn luyện mô hình trên tập dữ liệu lớn và tập dữ liệu mô phỏng bằng công cụ Google Colab.
* Ứng dụng thuật toán YOLOv3 (You Only Look Once, Version 3) để phát hiện những vị trí còn trống, đã đỗ trên hình ảnh, video, camera bằng Pytorch.
* Ứng dụng Nicepage thiết kế website giới thiệu sản phẩm và sử dụng Django để xử lý trang web, truyền màn hình đã xử lý lên website.

1. **Nội dung và phạm vi đề tài:** 
   1. **Nội dung đề tài:**

* Giới thiệu và phân biệt các khái niệm liên quan đến Trí Tuệ Nhân Tạo (Artificial Intelligence), Thị Giác Máy Tính (Computer Vision), Học Máy (Machine Learning), Học Sâu (Deep Learning).
* Nghiên cứu và triển khai thuật toán YOLOv3 bằng Pytorch:
* Triển khai module nhận diện qua hình ảnh.
* Triển khai module nhận diện qua video/camera.
* Kiểm thử mô hình.
* Nghiên cứu các chỉ số đánh giá mô hình.
* Xây dựng trang web hiển thị màn hình xử lý bằng Django.
* Hiển thị song song màn hình xử lý trên website và trên desktop.

1. **Phạm vi đề tài:**

* Bài toán nhận diện vật thể (Object Detection).
* Giới thiệu họ các thuật toán Mạng Thần kinh Tích chập (Convolution Neural Network) trong nhận diện vật thể.
* Ứng dụng Django để xây dựng website giới thiệu sản phẩm.
* Tích hợp module nhận diện qua camera vào Django để truyền màn hình đã xử lý lên website.

1. **Công nghệ, công cụ và ngôn ngữ lập trình:**
2. **Công nghệ:**

Python, OpenCV, Pytorch, Django, Colab Notebook, CUDA, Darknet.

1. **Công cụ:**

* Một số thư viện mã nguồn mở của Python:

Opencv-python, pandas, numpy, django, torch,...

* Visual Studio Code
* Darknet: Open Source Neural Networks
* Google Colab

1. **Ngôn ngữ lập trình:** Python
2. **Các kết quả chính dự kiến sẽ đạt được và ứng dụng:**

* Sử dụng camera tiến hành phát hiện các vị trí còn trống và đã đỗ trong thời gian thực. Hiển thị lên màn hình vị trí còn trống và đã đỗ, đếm những vị trí còn trống, đã đỗ, tổng các vị trí hiện có.
* Sử dụng Django để xây dựng website.
* Mô phỏng mô hình trực tiếp.
* Hoàn chỉnh cuốn báo cáo đề tài.
* Nắm được kiến trúc thuật toán YOLOv3 và có thể ứng dụng vào mọi đề tài liên quan.
* Nắm được các ưu, nhược điểm của thuật toán và các phương pháp tối ưu cho thuật toán.
* Nắm được những quy trình trong huấn luyện và kiểm tra mô hình trong các mô hình Deep Learning.

1. **Giáo viên và cán bộ hướng dẫn**

Họ tên: PHẠM THỊ MIÊN

Đơn vị công tác: Bộ môn Công Nghệ Thông Tin – Trường Đại học Giao thông Vận tải phân hiệu tại TP HCM

Điện thoại: 0961170638 Email:

|  |  |
| --- | --- |
| **Ngày … tháng 03 năm 2021**  **BM Công Nghệ Thông Tin** | **Đã giao nhiệm vụ TKTN**  **Giáo viên hướng dẫn** |
|  | **ThS. Phạm Thị Miên** |

Đã nhận nhiệm vụ TKTN

Sinh viên: NGUYỄN VĂN HIỆP Ký tên:

Điện thoại: 0973550322

Email: [nguyenvanhiepcmg@gmail.com](mailto:nguyenvanhiepcmg@gmail.com)

**LỜI CẢM ƠN**

Qua thời gian học tập và rèn luyện tại trường Trường Đại học Giao thông Vận tải phân hiệu tại TP HCM, đến nay chúng em đã kết thúc khoá học 4 năm và hoàn thành đồ án tốt nghiệp. Trong thời gian học tập tại trường để có được kết quả hiện tại em xin chân thành cảm ơn:

Tập thể các thầy cô giáo Bộ môn Công Nghệ Thông Tin và các thầy cô thỉnh giảng đã giảng dạy, quan tâm và tạo điều kiện thuận lợi để chúng em học tập rèn luyện trong suốt thời gian qua, giúp chúng em trang bị những kiến thức, kỹ năng cần thiết cho công việc thực tế sau này. Cảm ơn thầy cô giáo Bộ môn cũng như Ban Giám Hiệu đã cho phép em thực hiện đề tài tốt nghiệp: ***NGHIÊN CỨU TRIỂN KHAI THUẬT TOÁN YOLOV3 VÀ XÂY DỰNG HỆ THỐNG HƯỚNG DẪN ĐẬU XE Ô TÔ PARKING VISION***.

Thạc sĩ Phạm Thị Miên đã luôn quan tâm nhiệt tình hướng dẫn, giúp đỡ em trong quá trình thực hiện đồ án tốt nghiệp. Cô cũng luôn nhắc nhở, động viên mỗi khi em gặp khó khăn, nhờ vậy mà em đã hoàn thành tốt đồ án tốt nghiệp của mình đúng thời hạn được giao. Nếu không có những lời hướng dẫn, dạy bảo của cô thì em nghĩ bài báo cáo này của em sẽ rất khó có thể hoàn thiện được.

Em cũng xin gửi lời cảm ơn tới gia đình, bạn bè, những người đã động viên, giúp đỡ em rất nhiều trong thời gian học tập và làm đồ án tốt nghiệp.

Mặc dù đã cố gắng nỗ lực học hỏi không ngừng để hoàn thành đề tài, vậy nhưng thời gian thực hiện đồ án có hạn, kiến thức của em còn hạn chế. Do vậy, không tránh khỏi những thiếu sót, em rất mong nhận được những ý kiến đóng góp quý báu của thầy cô trong hội đồng bảo vệ đồ án tốt nghiệp để kiến thức của em được hoàn thiện hơn.

Đồng thời em xin cam đoan rằng nội dung đồ án của chính em nghiên cứu xây dựng nên, nếu có nội dung tham khảo đều được trích dẫn cụ thể, rõ ràng.

***TP. Hồ Chí Minh, ngày 11 tháng 06 năm 2021***

**Sinh viên thực hiện**

**Nguyễn Văn Hiệp**

**NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN**

|  |
| --- |
| ***Tp. Hồ Chí Minh, ngày … tháng 06 năm 2021***  **Giáo viên hướng dẫn**  **ThS. Phạm Thị Miên** |

**MỤC LỤC**

**DANH MỤC THUẬT NGỮ**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **STT** | **THUẬT NGỮ** | **Ý NGHĨA TIẾNG VIỆT** | **TỪ VIẾT TẮT** | **GHI CHÚ** |
| 1 | You Only Look Once | Bạn Chỉ Nhìn Một Lần | YOLO |  |
| 2 | You Only Look Once Version 2 | Bạn Chỉ Nhìn Một Lần Phiên bản 2 | YOLOv2 |  |
| 3 | You Only Look Once Version 3 | Bạn Chỉ Nhìn Một Lần Phiên bản 3 | YOLOv3 |  |

**DANH MỤC CÁC BẢNG**

**DANH MỤC HÌNH ẢNH**

# **TỔNG QUAN**

## **Sơ lược về Computer Vision và sự hỗ trợ của Deep Learning**

### ***Thị Giác Máy Tính (Computer Vision)***

Thị Giác Máy Tính (Computer Vision) đề cập đến toàn bộ quá trình mô phỏng tầm nhìn của con người trong một bộ máy phi sinh học. Điều này bao gồm việc chụp ảnh ban đầu, phát hiện và nhận dạng đối tượng, nhận biết bối cảnh tạm thời giữa các cảnh và phát triển sự hiểu biết ở mức độ cao về những gì đang xảy ra trong khoảng thời gian thích hợp. Công nghệ này từ lâu đã trở nên phổ biến trong khoa học viễn tưởng, và vì thế, chúng thường được thừa nhận như một điều điều hiển nhiên. Trong thực tế, một hệ thống cung cấp khả năng Thị Giác Máy Tính đáng tin cậy, chính xác và trong thời gian thực là một vấn đề đầy thách thức vẫn chưa được phát triển một cách đầy đủ.

Khi các hệ thống này được phát triển hoàn chỉnh, sẽ có vô số ứng dụng dựa vào Thị Giác Máy Tính như một thành phần chính. Những ví dụ điển hình đó là xe hơi tự lái, robot tự động, máy bay không người lái, thiết bị hình ảnh y tế thông minh hỗ trợ phẫu thuật và cấy ghép phẫu thuật phục hồi thị lực của con người.

Có thể Thị Giác Máy Tính mang nhiều hứa hẹn rất lớn trong tương lai, nhưng nó mang theo một sự phức tạp vốn có và luôn là thách thức đối với các hệ thống máy tính. Một phần của sự phức tạp là do thực tế Thị Giác Máy Tính không phải là một nhiệm vụ duy nhất. Thay vào đó, nó là một chuỗi các nhiệm vụ không đơn giản mà mỗi yêu cầu sử dụng các thuật toán phức tạp và đủ sức mạnh tính toán để hoạt động trong thời gian thực. Ở cấp độ cao, các tác vụ phụ của Thị Giác Máy Tính là phát hiện và phân đoạn đối tượng, phân loại hình ảnh, theo dõi đối tượng, gắn nhãn hình ảnh với các mô tả có ý nghĩa (ví dụ như chú thích hình ảnh) và cuối cùng, hiểu ý nghĩa của toàn bộ bối cảnh.

### ***Đóng góp từ Deep Learning***

Mặc dù vẫn còn những trở ngại đáng kể trong con đường phát triển của Thị Giác Máy Tính đến “cấp độ con người”, các hệ thống Deep Learning đã đạt được tiến bộ đáng kể trong việc xử lý một số nhiệm vụ phụ có liên quan. Lý do cho sự thành công này một phần dựa trên trách nhiệm bổ sung được giao cho các hệ thống Deep Learning. Điều hợp lý để nói rằng sự khác biệt lớn nhất với các hệ thống Deep Learning là chúng không còn cần phải được lập trình để tìm kiếm các đặc điểm cụ thể. Thay vì tìm kiếm các đặc điểm cụ thể bằng thuật toán được lập trình cẩn thận, các mạng lưới thần kinh bên trong các hệ thống Deep Learning được đào tạo. Ví dụ: nếu ô tô trong hình ảnh bị phân loại sai thành xe máy thì ta không tinh chỉnh các tham số hoặc viết lại thuật toán. Thay vào đó, ta tiếp tục đào tạo cho đến khi hệ thống làm cho đúng.

Với sức mạnh tính toán tăng lên được cung cấp bởi các hệ thống Deep Learning hiện đại, có sự tiến bộ ổn định và đáng chú ý hướng tới điểm mà một máy tính sẽ có thể nhận ra và phản ứng với mọi thứ mà nó nhìn thấy.

### ***Một số ứng dụng chính***

* Phân loại hình ảnh
* Phân loại hình ảnh cùng với việc bản địa hóa
* Phát hiện đối tượng
* Tái thiết hình ảnh
* Theo dõi đối tượng

Những tiến bộ trong hệ thống Deep Learning và sức mạnh tính toán đã giúp cải thiện tốc độ, độ chính xác và độ tin cậy tổng thể của hệ thống Thị Giác Máy Tính. Khi các mô hình Deep Learning cải thiện và sức mạnh tính toán trở nên dễ dàng hơn, chúng ta sẽ tiếp tục đạt được những tiến bộ và sự ổn định đối với các hệ thống tự vận hành có thể thực sự nắm bắt và phản ứng với những gì chúng cảm nhận.

**Đặt vấn đề**

***Thực trạng***

Cùng với sự phát triển vượt bậc về kinh tế xã hội, nhu cầu của con người ngày càng được nâng cao, do đó nhu cầu lưu thông hàng hoá và những đòi hỏi về đi lại ngày càng tăng. Có thể nói trong giai đoạn hiện nay, khi xu thế toàn cầu hoá diễn ra ngày càng sôi động, người ta càng nhận thức rõ hơn tầm quan trọng của giao thông vận tải.

Nếu như trên không trung, máy bay là phương tiện chiếm ưu thế, thì trên mặt đất, ô tô và vận tải ô tô lại chiếm ưu thế về năng lực vận chuyển và khả năng cơ động. Ô tô có thể hoạt động trên nhiều dạng địa hình, từ đồng bằng, miền núi đến miền biển, vận chuyển một khối lượng hàng hoá và vận chuyển người nhiều hơn bất cứ loại phương tiện vận tải nào khác.

Theo Sở Giao thông vận tải Hà Nội, trên địa bàn Hà Nội hiện có khoảng 6,9 triệu phương tiện giao thông (ô-tô, xe máy), chưa kể lượng xe ngoại tỉnh ra, vào thành phố hằng ngày. Dự tính, với tốc độ tăng trưởng xe máy 7,66%/năm; ô-tô 16,15%/năm thì đến năm 2025 Hà Nội có 1,3 triệu ô-tô và 7,3 triệu xe máy; năm 2030 có 1,7 triệu ô-tô và 7,7 triệu xe máy. Thế nhưng, điều đáng nói là hệ thống giao thông tĩnh (bãi, điểm đỗ xe công cộng) hiện nay còn quá thiếu và yếu, khiến nhu cầu đỗ xe ngày càng cấp thiết.[[1]](#footnote-1) (*theo báo nhandan.vn số ra ngày 09/03/2021*)

Thực trạng cho ta thấy cùng với sự tăng lên của các phương tiện giao thông chỉ ở Hà Nội, với các thành phố lớn thì cũng đang gặp vấn đề tương tự, ô tô đang có tốc độ tăng trưởng về số lượng khá cao về số lượng, do đó nhu cầu đỗ xe cũng ngày càng cao. Việc tìm được một chỗ đỗ xe trong thời buổi hiện nay cũng đang còn khá khó khăn. Đôi khi, có thể có chỗ trống trong bãi đỗ xe nhưng người sử dụng sẽ không biết nó nằm ở đâu, hoặc là không biết trong bãi có còn chỗ trống hay không. Rất nhiều vụ đỗ xe trước nhà dân, vỉa hè, lòng đường, công viên các nơi không phép… dẫn đến nhiều bất cập, ví dụ tình trạng tắc nghẽn giao thông và mất mỹ quan đô thị gây cản trở đường đi của chủ nhà, thậm chí nguy hiểm cho các phương tiện giao thông khác.



Hình 0.1. Chiếc bán tải bị tạt sơn khi đỗ trên vỉa hè. Ảnh Internet

### ***Các loại hình bãi đỗ xe phổ biến hiện nay***

* **Bãi đỗ xe truyền thống:** Hệ thống đỗ xe ô tô truyền thống có ở hầu hết các gia đình, văn phòng và trung tâm giải trí. Không gian đỗ xe riêng biệt được tạo ra để vào và ra trơn tru để tăng tổ chức và giảm lưu lượng. Chúng đòi hỏi ít năng lượng hơn với yêu cầu năng lượng duy nhất là chiếu sáng. Nhà để xe truyền thống cũng cực kỳ đáng tin cậy ở chỗ chức năng của họ không bị cản trở bởi sự cố mất điện. Nhà để xe truyền thống đang hoạt động suốt ngày đêm, một số bãi đỗ xe truyền thống có không gian hạn chế nên người người sử dụng phải tìm chỗ trống trước khi đỗ xe. Nếu chúng có thiết kế đơn giản, số vốn bỏ ra không lớn. Chiến lược bảo vệ xe hơi cũng có thể được cung cấp bằng cách giao nhiều lô cho các cá nhân cụ thể quản lý.

Hình 0.2. Bãi đỗ xe truyền thống tại siêu thị Emart Gò Vấp TP. HCM. Ảnh Internet

* **Hệ thống bãi đỗ xe tự động:** đòi hỏi ít diện tích hơn, loại bỏ trình điều khiển khỏi quá trình đỗ xe. Nó có thể là hoàn toàn tự động hoặc bán tự động. Chiếc xe phải được lái đến một điểm nhập cảnh nơi người lái và hành khách thoát khỏi xe. Sau đó, nó được di chuyển tự động hoặc bán tự động (với một số hỗ trợ cần thiết) đến

không gian được phân bổ của nó. Hệ thống bãi đỗ xe tự động tối đa hóa không gian hạn chế, một lợi thế trong khu vực không gian hạn chế có sẵn.



Hình 0.3. Bãi đỗ xe tự động tại Đà Nẵng. Ảnh Internet

### ***Hệ Thống Hướng Dẫn Bãi Đỗ Xe (Parking Guidance System - PGS)***

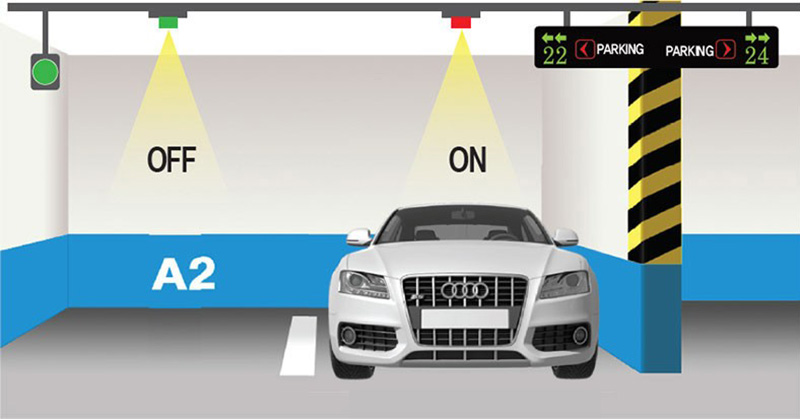
Cho dù là bãi đỗ xe truyền thống hay tự động thì trong một bãi đỗ xe hoàn chỉnh không thể thiếu hệ thống hướng dẫn đỗ xe (Viết tắt là PGS), các loại hình phổ biến tại Việt Nam hiện nay:

* **Chỉ dẫn thủ công:** Đây là phương pháp chỉ dẫn truyền thống nhất, mọi công việc đều do con người trực tiếp tham gia vào làm, khi đó khi vào chỗ để xe thì người sử dụng sẽ được một nhân viên trong bãi xe hướng dẫn đỗ xe, hoặc nhiều khi đó cũng chính là người ghi vé. Phương pháp này tốn rất nhiều nhân công và thiếu chính xác, ví dụ như một bãi đỗ xe lớn thì nhiều khi họ sẽ không nắm được số lượng chỗ trống còn lại và vị trí nào còn trống trong bãi để chỉ dẫn cho người sử dụng được. Hơn nữa còn mang lại phiền toái khá nhiều cho người sử dụng khi phải mất thời gian tự đi tìm chỗ trống, dễ gây ra va chạm với các xe khác, tốn nhiên liệu tiêu thụ cho xe và tăng mức độ ô nhiễm môi trường, làm tâm trạng người sử dụng trở nên chán nản và bực bội, gây ùn ứ các phương tiện đi vào bãi, có thể làm giảm doanh thu của bãi xe.



Hình 0.4. Minh họa hướng dẫn đỗ xe thủ công. Ảnh Internet

* **Hệ thống hướng dẫn thông minh bằng cảm biến:** Hệ thống hướng dẫn đỗ xe thông minh cung cấp sự tiện lợi trong việc hướng dẫn đỗ xe một cách tự động cho bãi đỗ xe ô tô. Người sử dụng dễ dàng nhận biết vị trí còn trống tại từng khu vực trong bãi đỗ xe khi họ đi vào bãi. Tại mỗi vị trí xe đỗ được gắn một thiết bị siêu âm để phát hiện vị trí đỗ là có xe hay không có xe. Tình trạng của các vị trí báo cáo về trung tâm theo thời gian thực. Máy tính chủ sẽ thu thập tất cả trạng thái của từng vị trí từ bộ siêu âm và điều khiển các bảng quang báo hiển thị để hướng dẫn người sử dụng dễ dàng tìm đến vị trí còn trống để đỗ xe.



Hình 0.5. Minh họa cho hệ thống hướng dẫn bằng cảm biến

Tuy nhiên giá mỗi cảm biến rất đắt đỏ, tốn khá nhiều chi phí triển khai, việc bảo trì cũng diễn ra khá khó khăn và tốn kém khi phải thuê nhân viên bên ngoài về vì yêu cầu tính chuyên môn cao, và giá linh kiện đắt đỏ, hơn nữa không phải lúc nào thì người sử dụng cũng biết là trong bãi còn chỗ trống hay không, đến khi vào nhưng hết chỗ thì cũng gây ra nhiều phiền toái. Hệ thống này chủ yếu được triển khai trong nhà có mái che, hoặc hầm.

### ***Giải pháp mới sử dụng công nghệ Computer Vision và Deep Learning: Hệ thống hướng dẫn đậu xe thông minh Parking Vision***

Để giải quyết những nhược điểm của các phương pháp trên thì giải pháp được đề xuất để giải quyết cũng như là đề tài của đồ án này đó là *nghiên cứu triển khai thuật toán YOLOv3 và xây dựng hệ thống hướng dẫn đậu xe ô tô Parking Vision*. Ta sẽ sử dụng bài toán **nhận diện vật thể** nằm trong lĩnh vực Computer Vision để nhận diện ra các vị trí đã đỗ và còn trống trong bãi, cùng với đó sử dụng công nghệ Deep Learning để huấn luyện và hỗ trợ cho Computer Vision trong việc kiểm tra dự đoán của mô hình. Một trong những sự kết hợp đó là thuật toán You Only Look Once (YOLO), là một mô hình CNN để nhận diện vật thể mà một ưu điểm nổi trội là nhanh hơn nhiều so với những mô hình cũ. Thậm chí có thể chạy tốt trên những thiết bị IOT như Raspberry Pi. Khi sử dụng phương pháp này thì có những ưu điểm sau:

* **Đối với chủ đầu tư:** Giúp chủ đầu tư quản lý bãi giữ xe một cách chuyên nghiệp, khoa học với các công nghệ hiện đại, đồng thời giảm chi phí chủ đầu tư (bao gồm chi phí đầu tư máy chủ, tên miền, màn hình hiển thị và camera đối với giải pháp này), thích hợp triển khai ở những nơi ngoài trời như bãi đỗ xe của siêu thị, trường học, nhà hàng... có thể bao quát tầm nhìn của toàn bộ xe.
* **Đối với người sử dụng:** Thuận tiện cho người sử dụng, có thể nắm bắt số lượng chỗ trống và vị trí chính xác trong bãi, giúp chủ xe giảm thời gian tìm kiếm, tiết kiệm nhiên liệu, giúp tâm trạng thoải mái hơn thay vì tự mình tìm xem vị trí nào còn trống. Nhờ đó có thể biết trước bãi còn chỗ đỗ xe hay không, có thể tự tin đến nơi đỗ xe. Hơn nữa người sử dụng có thể theo dõi tình trạng xe mình ở mọi lúc mọi nơi, sử dụng như một camera an ninh.
* **Đối với người quản lý:** Thuận tiện cho người quản lý bãi xe, giúp cho họ nắm bắt thông tin số lượng chỗ trống hiện tại, chính xác vị trí còn trống để chỉ dẫn cho các tài xế vào bãi, tạo điều kiện thuận lợi cho việc điều hành các phương tiện di chuyển.

**Tình hình nguyên cứu**

Hiện nay ở Việt Nam phương pháp hướng dẫn ô tô dựa vào Computer Vision còn chưa phát triển nhiều, có lẽ vì ngành Computer Vision ở Việt Nam chưa được ứng dụng sâu và rộng, đây là một ngành mới nổi ở Việt Nam và có nhiều triển vọng phát triển, hiện tại chỉ được áp dụng nhiều vào các giải pháp nhận diện khuôn mặt, nhận diện ký tự - OCR. Những sản phẩm của các giải pháp đó có thể kể đến như: Định danh khách hàng - eKYC, nhận dạng giấy tờ tùy thân, nhận dạng biển số xe. Nhưng công nghệ luôn phát triển không ngừng, và các bài toán đó đã được làm đi làm lại, cải tiến rất nhiều lần và cũng đã dần hoàn thiện, điều đó thúc đẩy mở ra các giải pháp khác cho chúng ta tìm hiểu. Ta nên dự đoán trước xu thế phát triển của tương lai và nghĩ ra giải pháp cho những vấn đề mới.

Hiện tại dù YOLO đã ra tới phiên bản YOLOv5, nhưng phiên bản này đang còn nhiều tranh cãi về tốc độ cũng như độ chính xác, không phải do chính chủ tác giả phát hành. Phiên bản YOLOv4 cũng đã có rất nhiều cải tiến so với YOLOv3 nhưng hiện giờ YOLOv3 vẫn còn rất mạnh về cả tốc độ lẫn độ chính xác, hơn nữa với một người mới tìm hiểu về lĩnh vực này thì tôi chọn phiên bản YOLOv3 để tìm hiểu và triển khai, bởi vì phiên bản càng mới thì kiến trúc của thuật toán càng phức tạp, nếu quá phức tạp sẽ rất khó để tìm hiểu trong thời gian ngắn.

PyTorch là một package được xây dựng dựa trên Python để thay thế Numpy và cung cấp tính linh hoạt như một nền tảng phát triển Deep Learning. Pytorch là framework được phát triển bởi Facebook, Inc[[2]](#footnote-2). Đây là một ông lớn về công nghệ đầu tư rất nhiều nguồn lực cho việc phát triển Trí tuệ nhân tạo. Pytorch được phát triển với giấy phép mã nguồn mở do đó nó tạo được cho mình một cộng đồng rất lớn. Một cộng đồng lớn đồng nghĩa với nhiều tài nguyên để học và các vấn đề của ta có thể đã có ai đó giải quyết và chia sẻ với cộng đồng. Pytorch cùng với Tensorflow[[3]](#footnote-3) và Keras[[4]](#footnote-4) là một trong những framework phổ biến được sử dụng trong các bài toán về Deep Learning hiện nay. Đặc biệt, trong các lĩnh vực nghiên cứu, hầu như các tác giả đều sử dụng pytorch để triển khai bài toán của mình.

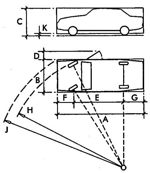
Django là một trong số những web framework bậc cao miễn phí, là mã nguồn mở được tạo ra bởi ngôn ngữ Python dựa trên mô hình mô hình MTV (gồm Model-Template-Views). Hiện framework này được phát triển, quản lý bởi Django Software Foundation. Django ra đời với mục tiêu hỗ trợ thiết kế các website phức tạp dựa trên những CSDL có sẵn. Nó hoạt động dựa theo nguyên lý ‘cắm’ các thành phần và và tái sử dụng để tạo nên các website với ít code, ít khớp nối, có khả năng phát triển và không bị trùng lặp. Lợi thế hàng đầu của Django là khả năng thiết kế, tạo lập website và các ứng dụng nhanh chóng đến bất ngờ. Vì vậy Django được sử dụng để thực hiện đồ án này.

## **Quá trình nguyên cứu**

Để hiểu rõ hơn về nghiệp vụ liên quan tới bãi đỗ xe, tôi sử dụng các mẫu bãi đỗ xe truyền thống và phổ biến nhất trên internet để tìm hiểu và triển khai giúp ích cho quá trình thu thập hình ảnh và đánh nhãn cũng như nhận diện sau này được dễ dàng hơn.

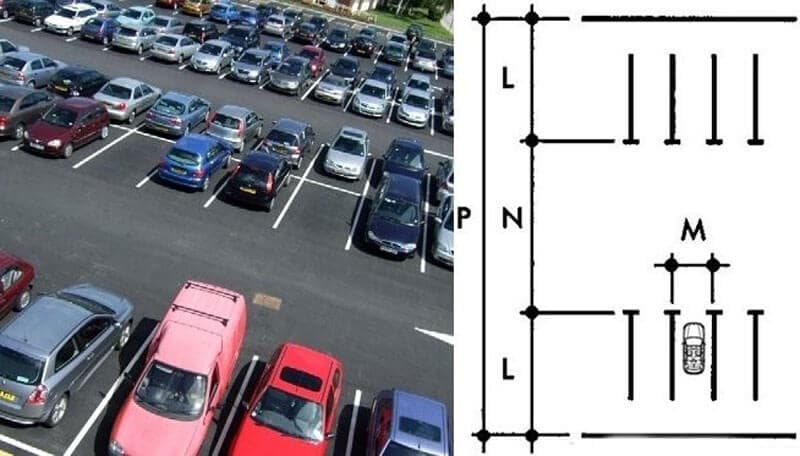
### ***Tiêu chuẩn kích thước của bãi đỗ xe ô tô***

* Kích thước ô tô tiêu chuẩn:



Hình 0.6. Minh họa kích thước ô tô tiêu chuẩn

* Chiều dài xe (A): 4.75 m
* Chiều rộng xe (B): 1.80 m
* Chiều cao xe (C): 1.70 m
* Khoảng mở tối thiểu (D): 0.50 m
* Khoảng cách giữa 2 trục bánh xe (E): 2.90 m
* Khoảng cách từ đầu xe đến trục bánh trước (F): 0.90 m
* Khoảng cách từ đuôi xe đến trục bánh sau (G): 1.10 m
* Đường kính vòng quay xe (H): 13 m
* Khoảng cách từ tâm đường kính quay đến tường (J): 14 m
* Khoảng sáng gầm xe (K): 0.10 m
* Tiêu chuẩn kích thước bãi đỗ xe thông dụng:



Hình 0.7. Tiêu chuẩn kích thước bãi đỗ xe thông dụng

Diễn giải các ký hiệu kích thước chỗ đỗ xe ô tô:

+ L: Chiều dài chỗ đỗ xe ô tô: Từ 4.75m đến 5.5m

+ M: Chiều rộng chỗ đậu xe ô tô; chiều rộng chỗ đỗ xe được quy định như sau:

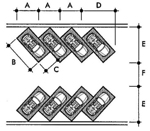
* Chiều rộng chỗ đỗ xe ô tô tiêu chuẩn : Từ 2.4m đến 2.75m
* Chiều rộng chỗ đỗ xe ô tô thời gian ngắn: Từ 2.3m đến 2.6m
* Chiều rộng chỗ đỗ xe ô tô thời gian dài: Từ 2.5m đến 2.75m
* Chiều rộng chỗ đỗ xe ô tô dành cho người khuyết tật: Từ 3m đến 3.5m

+ N: Chiều rộng lối đi ôtô :

* Chiều rộng lối đi ô tô tại làn đường có 1 chiều xe chạy : Từ 6m đến 9.15m
* Chiều rộng lối đi ô tô tại làn đường có 2 chiều xe chạy : Từ 6.95m đến 10.7m

+ P: Chiều rộng tiêu chuẩn bãi đậu xe : Từ 15.5m đến 20.1m

* Tiêu chuẩn kích thước bãi đỗ xe ô tô chéo góc 45 độ:



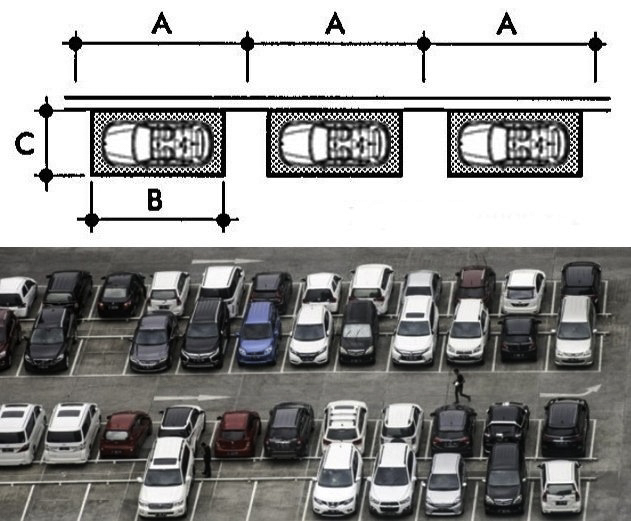
Hình 0.8. Tiêu chuẩn thiết kế bãi đỗ xe 45 độ

+ Tiêu chuẩn:

Bảng 0.1. Tiêu chuẩn kích thước bãi đỗ xe 45 độ (tính theo mét)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Vị trí** | **A** | **B** | **C** | **D** | **E** | **F** |
| Trong nhà | 3.0 | 4.6 | 2.3 | 3.26 | 5.0 | 2.8 |
| Ngoài trời | 3.39 | 5.50 | 2.40 | 3.89 | 5.50 | 2.8 - 3.0 |

* Tiêu chuẩn kích thước bãi đỗ xe ô tô song song:



Hình 0.9. Tiêu chuẩn thiết kế bãi đỗ xe song song

+ Tiêu chuẩn:

Bảng 0.2. Tiêu chuẩn kích thước bãi đỗ xe song song (tính theo mét)

|  |  |  |  |
| --- | --- | --- | --- |
| **Vị trí** | **A** | **B** | **C** |
| Trong nhà | 5.8 | 4.6 | 2.20 - 2.30 |
| Ngoài trời | 6.10 - 6.70 | 5.50 | 2.40 |

### ***Các trạng thái có thể có trong một chỗ đỗ xe ô tô***

Chỉ có 2 trạng thái chính, đó là:

* Trống: Chỗ đó chưa có ô tô nào được đỗ
* Đã đỗ: Chỗ đó đã có ô tô đỗ

### ***Các điều kiện bên ngoài ảnh hưởng đến chất lượng hình ảnh***

* Thời tiết: Mưa, nắng
* Ánh sáng: Nắng đỗ bóng, nắng gắt, nắng nhẹ, râm mát, trời tối bật đèn, sương mù,...
* Các yếu tố gây nhiễu: Cây cối, bụi bặm bám trên camera,…

### ***Một số dạng camera giám sát phổ biến hiện nay***

* **Camera quan sát có dây (Camera Analog): Camera quan sát có dây là loại camera với cảm biến CCD với hình ảnh được số hóa để xử lí. Camera ghi hình ổn định và chuyên nghiệp cùng sự hỗ trợ của nhiều thiết bị như dây cáp, đầu ghi,...**

**Camera có dây hiện nay đã được cải thiện khá nhiều về cách thức hoạt động và hỗ trợ công nghệ quan sát từ xa qua mạng Internet giúp người dùng dễ dàng quan sát và quản lí thông qua các ứng dụng được cài đặt trên điện thoại thông minh, laptop, máy tính bảng,... bạn có thể quan sát nơi làm việc, nhà ở, bãi đỗ xe,... ở mọi nơi nhờ những tính năng nổi bật của camera có dây.**

**Camera có dây đem lại hình ảnh chất lượng tốt, truyền tải ổn định, tính bảo mật cao.**



Hình 0.10. Minh họa Camera Analog

* **Camera quan sát không dây (Camera Wifi): Camera Wifi có ưu điểm lắp đặt nhanh, sử dụng dễ dàng, thuận tiện vì vậy sản phẩm camera không dây trở nên thông dụng ở thời điểm hiện tại. Camera được kết nối với Wifi ổn định và chỉ cần một thiết bị như smartphone, máy tính bảng, laptop,... có kết nối Internet là bạn có thể giám sát mọi thứ ở bất kỳ đâu. Ngoài ra, một số camera không dây hỗ trợ thẻ nhớ giúp bạn thuận tiện xem lại hình ảnh khi có cần.**

**Camera quan sát không dây tích hợp các tính năng thông minh như cảnh báo chuyển động, đàm thoại 2 chiều, xoay 360 độ,... giúp chống trộm và đảm bảo an ninh an toàn, hiệu quả.**



Hình 0.11. Minh họa Camera Wifi

### ***Quá trình giải quyết bài toán nhận diện vật thể bằng thuật toán YOLOv3***

Có 2 phương án để giải quyết bài toán nhận diện chỗ còn trống và đã đỗ với thuật toán YOLOv3, đó là:

* **Thứ nhất, sử dụng tệp weights được tác giả thuật toán YOLOv3 huấn luyện sẵn với tập dữ liệu COCO (tệp weights và tập dữ liệu COCO sẽ được giới thiệu sau), sau đó nhận diện các vị trí của ô tô, và lấy tổng các vị trí hiện có trong bãi trừ cho số lượng vị trí ô tô đó. Tuy phương án này rút ngắn thời gian huấn luyện mô hình nhưng độ chính xác không cao, dễ bị nhiễu bởi các đối tượng khác bởi vì tập dữ liệu COCO được huấn luyện để nhận dạng lên đến 80 đối tượng, và trong lúc nhận diện có thể có ô tô đang di chuyển sẽ khiến thuật toán nhận diện nhầm. Hơn nữa các vị trí còn trống sẽ không được nhận diện, cho nên thuật toán sẽ không biết khi nào ô tô được đậu đúng vị trí, vì khi ô tô đậu bên ngoài vạch cũng được nhận diện. Và cuối cùng phương án này chúng ta cần phải biết trước số lượng tổng các vị trí trong bãi. Đây là phương án tiếp cận quá nhiều bước.**
* **Thứ hai, ta sẽ huấn luyện mô hình để nhận diện được cả 2 trạng thái trống và đã đỗ trong bãi, sau đó tính tổng bằng cách lấy tổng các vật thể ta nhận diện được. Ta sẽ chọn phương án này để triển khai.**

## **Cấu trúc báo cáo đồ án tốt nghiệp**

Cấu trúc đồ án được chia thành các chương như sau:

Tổng quan – Giới thiệu tổng quan về đề tài đồ án tốt nghiệp.

Chương 1: Kiến thức nền tảng

* Phân biệt các khái niệm liên quan đến Trí Tuệ Nhân Tạo, Thị Giác Máy Tính, Học Máy, Học Sâu.
* Giới thiệu các công nghệ sử dụng.
* Các kiến thức liên quan đến Data Mining sử dụng trong bài làm.
* Tổng quan về xử lý ảnh.
* Tổng quan về Thị Giác Máy Tính và Học Sâu.

Chương 2: Mạng Thần Kinh Tích Chập (CNN)

* Giới thiệu Mạng Thần Kinh và Mạng Thần Kinh lan truyền ngược.
* Giới thiệu Mạng Thần Kinh Tích Chập.
* Các thuật toán dựa trên Mạng Thần Kinh Tích Chập trong nhận diện đối tượng.
* Giới thiệu về Transfer Learning.
* Giới thiệu các phiên bản của thuật toán YOLO.
* Giới thiệu các chỉ số đánh giá thuật toán.

Chương 3: Thiết kế, triển khai thuật toán và xây dựng website

* Giới thiệu một vài tập dữ liệu phổ biến.
* Thu thập dữ liệu, tiền xử lý, gắn nhãn và huấn luyện thuật toán.
* Triển khai thuật toán YOLOv3 nhận diện qua hình ảnh, video, camera xem trực tiếp màn hình xử lý trên máy tính.
* Kiểm tra thuật toán.
* Xây dựng website giới thiệu sản phẩm và xem màn hình xử lý.

Kết luận và kiến nghị

* Đưa ra kết quả đạt được, những thứ còn tồn tại và hướng phát triển về thuật toán lẫn website trong tương lai.

Tài liệu tham khảo.

# **KIẾN THỨC NỀN TẢNG**

## **1.1 Phân biệt các khái niệm liên quan đến Trí Tuệ Nhân Tạo, Thị Giác Máy Tính, Học Máy, Học Sâu**

***1.1.1 Trí Tuệ Nhân Tạo***

Trí Tuệ Nhân Tạo (Artificial Intelligence – viết tắt là AI) có thể được định nghĩa như một ngành của khoa học máy tính liên quan đến việc tự động hóa các hành vi thông minh. AI là một bộ phận của khoa học máy tính và do đó nó phải được đặt trên những nguyên lý lý thuyết vững chắc, có khả năng ứng dụng được của lĩnh vực này. Có thể hiểu đó là trí tuệ của máy móc được tạo ra bởi con người. Trí tuệ này có thể tư duy, suy nghĩ, học hỏi,... như trí tuệ con người. Xử lý dữ liệu ở mức rộng lớn hơn, quy mô hơn, hệ thống, khoa học và nhanh hơn so với con người.

 Trong lịch sử phát triển AI, các nhà nghiên cứu phân thành 4 hướng tiếp cận chính:

* Hành động như người (acting humanly)
* Suy nghĩ như người (thinking humanly)
* Suy nghĩ hợp lý (thinking rationally)
* Hành động hợp lý (acting rationally)
* Bài kiểm tra Turing (Turing test), đây là hướng phát triển AI hành động như người. Bài kiểm tra này được Alan Turing[[5]](#footnote-5) đề xuất vào năm 1950, mục đích để kiểm tra hệ thống máy tính đã đạt đến khả năng thông minh hay chưa. Bài test gồm một người đặt câu hỏi, một người trả lời câu hỏi và một máy tính phản hồi câu hỏi. Nếu trong quá trình trao đổi mà người đặt câu hỏi không thể phân biệt được người hay máy trả lời các câu hỏi này thì máy được xem là thông minh. Đến nay thì hướng tiếp cận này đã đạt được một số thành quả nhất định như:
* **Natural Language Processing**: Máy có khả năng đọc hiểu và giao tiếp bằng ngôn ngữ tự nhiên với người.
* **Knowledge Representation:** Máy có khả năng lưu trữ tri thức thông qua thị giác, thính giác, hay văn bản.
* **Automated Reasoning**: Máy có khả năng sử dụng tri thức đã lưu trữ để trả lời câu hỏi hay đưa ra kết luận hữu ích.
* **Machine Learning:** Máy có khả năng thích nghi với các điều kiện môi trường xung quanh để rút trích ra các nguyên lý từ tri thức thu nhận được phục vụ cho việc ra quyết định.
* **Computer Vision**: Máy có khả năng quan sát và xác định được các đối tượng xung quanh.
* **Robotics**: Máy có khả năng tương tác với đối tượng và di chuyển trong môi trường xung quanh.
* Suy nghĩ như người là một hướng tiếp cận khác thiên về khoa học nhận thức (cognitive science). Ngành này cần hẳn một thư viện để bạn đọc nghiên cứu. Vì để máy có thể suy nghĩ được như người thì bản thân ta phải hiểu được con người suy nghĩ như thế nào. Ta có ba cách để kiểm tra điều này:
* Quan sát quá trình suy nghĩ của mình.
* Quan sát hành động của một người.
* Quan sát hoạt động của não bộ.

Nếu ta có thể mô phỏng lại input (đầu vào) - output (đầu ra) tương tự như người thì đây là bằng chứng cho thấy máy có khả năng suy nghĩ như người ở góc độ nào đó. Các nhà nghiên cứu tin tưởng rằng khi kết hợp hai ngành này lại sẽ giúp chúng ta tiến đến mục tiêu AI nhanh chóng hơn.

* Suy nghĩ như người thì khó quá nên ta có thể giới hạn bài toán lại chỉ cần suy nghĩ hợp lý là được. Lúc này, logic học đã có cơ hội để sử dụng. Vào thế kỷ 19, các nhà logic học đã phát triển được các kí hiệu toán học để mô tả một phát biểu hay quan hệ giữa các đối tượng bất kỳ trong tự nhiên giúp cho máy có khả năng suy luận. Hướng tiếp cận này gặp vài trắc trở. Thứ nhất, máy khó diễn đạt được tri thức từ thế giới không chính tắc (informal) sang ngôn ngữ chính tắc (formal) đặc biệt khi tri thức đó không chắc chắn 100% là đúng. Thứ hai, có sự khác biệt lớn giữa giải quyết bài toán trên lý thuyết và bài toán ngoài thực tế. Do trong thực tế cần rất nhiều dữ kiện để giải quyết vấn đề nên nhiêu đây cũng đủ khiến cho các hệ thống tính toán bị quá tải.
* Cuối cùng là những con robot, là hướng tiếp cận hành động hợp lý. Robot hay còn gọi là agent, trong tiếng La Tinh có nghĩa làm gì đó. Các chương trình máy tính đều có khả năng làm gì đó nhưng agent được đòi hỏi nhiều hơn. Chúng phải có khả năng tự vận hành, thu nhận thông tin từ môi trường xung quanh, tồn tại được trong một thời gian dài, thích nghi với những thay đổi, đề xuất các mục tiêu và đạt được mục tiêu đó. Một con agent như vậy thường được áp dụng các bài test của Turing hay suy luận logic để đạt được các mục tiêu đề ra. Ví dụ điển hình là những con robot của Nasa được phóng lên các hành tinh có điều kiện khắc nghiệt và ít nhận được các chỉ thị của con người, làm sao chúng có thể tự thu thập các mẫu đất đá, tồn tại trong điều kiện gió bão, tự định hướng lịch trình khai thác và sử dụng năng lượng mặt trời để duy trì hoạt động của mình.

Tuy nhiên hiện nay, công nghệ AI vẫn đang còn rất nhiều hạn chế. Đơn cử như Alexa[[6]](#footnote-6) - một quản gia tuyệt vời, một trong những biểu tượng phổ biến nhất về ứng dụng của trí thông minh nhân tạo nhưng vẫn không thể vượt qua bài kiểm tra Turing.

Tóm lại, những gì chúng ta đang thực hiện với AI hiện nay nằm trong khái niệm “AI hẹp” (Narrow AI). Công nghệ này có khả năng thực hiện các nhiệm vụ cụ thể một cách tương tự, hoặc tốt hơn con người. Ví dụ về “AI hẹp” trong thực tế như công nghệ phân loại hình ảnh của Pinterest[[7]](#footnote-7) hay nhận diện khuôn mặt để gắn thẻ bạn bè trên Facebook[[8]](#footnote-8).

### ***1.1.2 Học Máy***

Học Máy (Machine Learning – viết tắt là ML) là một thuật ngữ rộng để chỉ hành động dạy máy tính cải thiện một nhiệm vụ mà nó đang thực hiện, ta có thể thấy ML chỉ là một thành phần trong hướng tiếp cận hành động như người. Cụ thể hơn, ML đề cập tới bất kỳ hệ thống mà hiệu suất của máy tính khi thực hiện một nhiệm vụ sẽ trở nên tốt hơn sau khi hoàn thành nhiệm vụ đó nhiều lần. Hay nói cách khác, khả năng cơ bản nhất của ML là sử dụng thuật toán để phân tích những thông tin có sẵn, học hỏi từ nó rồi đưa ra quyết định hoặc dự đoán về một thứ gì đó có liên quan. Thay vì tạo ra một phần mềm với những hành động, hướng dẫn chi tiết để thực hiện một nhiệm vụ cụ thể, máy tính được “huấn luyện” bằng cách sử dụng lượng dữ liệu và các thuật toán để học cách thực hiện nhiệm vụ.

Các thuật toán ML thường được chia làm 4 nhóm:

* Supervised learning (Học có giảm sát) là thuật toán dự đoán đầu ra (outcome) của một dữ liệu mới dựa trên các cặp (input, outcome) đã biết từ trước. Cặp dữ liệu này còn được gọi là (data, label), tức (dữ liệu, nhãn). Supervised learning là nhóm phổ biến nhất trong các thuật toán Machine Learning. Sử dụng các giải thuật và hướng giải quyết như: Decision tree (cây quyết định), k-NN, naive bayes, SVM, neural network (mạng thần kinh), deep learning (học sâu)…
* Unsupervised Learning (Học không giám sát): Trong thuật toán này, chúng ta không biết được outcome hay nhãn mà chỉ có dữ liệu đầu vào. Thuật toán unsupervised learning sẽ dựa vào cấu trúc của dữ liệu để thực hiện một công việc nào đó, ví dụ như phân nhóm hoặc giảm số chiều của dữ liệu để thuận tiện trong việc lưu trữ và tính toán. Sử dụng các giải thuật và hướng giải quyết như: k-means, hierachical clustering (phân cụm phân cấp).
* Semi-supervised lerning (Học bán giám sát): Các bài toán khi chúng ta có một lượng lớn dữ liệu nhưng chỉ một phần trong chúng được gán nhãn được gọi là Semi-Supervised Learning. Những bài toán thuộc nhóm này nằm giữa hai nhóm được nêu bên trên.
* Reinforcement learning (Học củng cố) là các bài toán giúp cho một hệ thống tự động xác định hành vi dựa trên hoàn cảnh để đạt được lợi ích cao. Hiện tại, Reinforcement Learning chủ yếu được áp dụng vào Lý Thuyết Trò Chơi, các thuật toán cần xác định nước đi tiếp theo để đạt được điểm số cao nhất. Bao gồm các phương pháp: Active (chủ động), passive (thụ động), generalization (tổng quát hóa).

Nếu không có ML, AI hiện tại sẽ bị hạn chế khá nhiều bởi nó mang lại cho máy tính sức mạnh để tìm ra mọi thứ mà không được lập trình rõ ràng. Ví dụ về một loại ML, giả sử ta muốn một chương trình có thể xác định được mèo trong các bức ảnh:

* Đầu tiên, ta cung cấp cho AI một tập hợp các đặc điểm của loài mèo để máy nhận dạng, ví dụ như màu sắc lông, hình dáng cơ thể, kích thước…
* Tiếp theo, ta cung cấp một số hình ảnh cho AI, trong đó một số hoặc tất cả các hình ảnh có thể được dán nhãn "mèo" để máy có thể chọn hiệu quả hơn các chi tiết, đặc điểm có liên quan đến mèo.
* Sau khi máy đã nhận được đủ dữ liệu cần thiết về mèo, nó phải biết cách tìm một con mèo trong một bức tranh - “Nếu trong hình ảnh có chứa các chi tiết X, Y, hoặc Z nào đó, thì 95% khả năng đó là một con mèo”.

Nhìn chung, ứng dụng của ML ngày nay là vô cùng phổ biến và độ hữu ích thì không phải bàn cãi nhiều nữa.

***1.1.3 Học Sâu***

Học Sâu (Deep Learning – viết tắt là DL) chỉ là một phương pháp nằm trong hướng giải quyết học có giám sát của ML. Nó như là một loại ML với "neural network" sâu có thể xử lý dữ liệu theo cách tương tự như một bộ não con người có thể thực hiện. Điểm khác biệt chính ở đây là con người không sẽ phải dạy một chương trình Deep Learning biết một con mèo trông như thế nào, mà chỉ cần cung cấp cho nó đủ hình ảnh cần thiết về loài mèo, và nó sẽ tự mình hình dung, tự học. Các bước cần làm như sau:

* Cung cấp cho máy rất nhiều ảnh về mèo.
* Thuật toán sẽ kiểm tra ảnh để xem các đặc điểm, chi tiết chung giữa các bức ảnh.
* Mỗi bức ảnh sẽ được giải mã chi tiết dưới nhiều cấp độ, từ các hình dạng lớn, chung đến các ô nhỏ và nhỏ hơn nữa. Nếu một hình dạng hoặc các đường được lặp lại nhiều lần, thuật toán sẽ gắn nhãn nó như là một đặc tính quan trọng.
* Sau khi phân tích đủ hình ảnh cần thiết, thuật toán giờ đây sẽ biết được các mẫu nào cung cấp bằng chứng rõ ràng nhất về mèo và tất cả những gì con người phải làm chỉ là cung cấp các dữ liệu thô.

Tóm lại Deep learning là loại ML mà trong đó máy tự đào tạo chính nó. Deep Learning đòi hỏi rất nhiều dữ liệu đầu vào và sức mạnh tính toán hơn là ML, nhưng nó đã bắt đầu được triển khai bởi các công ty công nghệ lớn như Facebook, Amazon[[9]](#footnote-9). Trong đó, một trong những cái tên nổi tiếng nhất về ML là AlphaGo[[10]](#footnote-10), một máy tính có thể chơi cờ vây với chính bản thân nó cho đến khi nó có thể dự đoán những đường đi nước bước chính xác nhất đủ để đánh bại nhiều nhà vô địch trên thế giới.

***1.1.4 Thị Giác Máy Tính***

Thị Giác Máy Tính (Computer Vision – viết tắt là CV) là một lĩnh vực liên ngành của Trí Tuệ Nhân Tạo nhằm mục đích cung cấp cho máy tính và các thiết bị khác có khả năng tính toán hiểu biết ở mức độ cao từ cả hình ảnh và video kỹ thuật số, bao gồm chức năng thu thập, xử lý và phân tích hình ảnh kỹ thuật số. Đây là lý do tại sao Thị Giác Máy Tính, một phần, là một lĩnh vực phụ khác của Trí Tuệ Nhân Tạo, chủ yếu dựa vào Học Máy và các thuật toán Học Sâu để xây dựng các ứng dụng Thị Giác Máy Tính. Ngoài ra, Thị Giác Máy Tính bao gồm một số công nghệ hoạt động cùng nhau: Đồ họa máy tính (Computer graphics), Xử lý hình ảnh (Image processing), Xử lý tín hiệu (Signal processing), Công nghệ cảm biến (Sensor technology), Toán học (Mathematics) hoặc thậm chí Vật lý (Physics).

**Từ các khái niệm trên ta có thể khái quát mối quan hệ giữa AI, ML, DL, CV thông qua sơ đồ sau:**

Graphical user interface

Description automatically generated

Hình 1.1. Sơ đồ mối quan hệ giữa AI, ML, DL, CV

**1.2 Xử lý ảnh**

***1.2.1 Khái niệm***

Xử lý ảnh (Image Processing), viết tắt là XLA, là đối tượng nghiên cứu của lĩnh vực Thị Giác Máy Tính, là quá trình biến đổi từ một ảnh ban đầu sang một ảnh mới với các đặc tính và tuân theo ý muốn của người sử dụng. Xử lý ảnh có thể gồm quá trình phân tích, phân lớp các đối tượng, làm tăng chất lượng, phân đoạn và tách cạnh, gán nhãn cho vùng hay quá trình biên dịch các thông tin hình ảnh của ảnh.

Cũng như xử lý dữ liệu bằng đồ hoạ, xử lý ảnh số là một lĩnh vực của tin học ứng dụng. Xử lý dữ liệu bằng đồ họa đề cập đến những ảnh nhân tạo, các ảnh này được xem xét như là một cấu trúc dữ liệu và được tạo bởi các chương trình. Xử lý ảnh số bao gồm các phương pháp và kỹ thuật biến đổi, để truyền tải hoặc mã hoá các ảnh tự nhiên. Mục đích của xử lý ảnh gồm:

* Biến đổi ảnh làm tăng chất lượng ảnh.
* Tự động nhận dạng ảnh, đoán nhận ảnh, đánh giá các nội dung của ảnh.

Nhận biết và đánh giá các nội dung của ảnh là sự phân tích một hình ảnh thành những phần có ý nghĩa để phân biệt đối tượng này với đối tượng khác, dựa vào đó ta có thể mô tả cấu trúc của hình ảnh ban đầu. Có thể liệt kê một số phương pháp nhận dạng cơ bản như nhận dạng ảnh của các đối tượng trên ảnh, tách cạnh, phân đoạn hình ảnh,… Kỹ thuật này được dùng nhiều trong y học (xử lý tế bào, nhiễm sắc thể), nhận dạng chữ trong văn bản.

***1.2.2 Các quá trình xử lý ảnh***

Diagram

Description automatically generated

Hình 1.2. Các giai đoạn chính trong xử lý ảnh

* Thu nhận ảnh (Image acquisition): Đây là công đoạn đầu tiên mang tính quyết định đối với quá trình XLA. Ảnh đầu vào sẽ được thu nhận qua các thiết bị như camera, sensor, máy scanner,v.v… và sau đó các tín hiệu này sẽ được số hóa. Việc lựa chọn các thiết bị thu nhận ảnh sẽ phụ thuộc vào đặc tính của các đối tượng cần xử lý. Các thông số quan trọng ở bước này là độ phân giải, chất lượng màu, dung lượng bộ nhớ và tốc độ thu nhận ảnh của các thiết bị.
* Tiền xử lý (image processing):Ở bước này, ảnh sẽ được cải thiện về độ tương phản, khử nhiễu, khử bóng, khử độ lệch,… với mục đích làm cho chất lượng ảnh trở lên tốt hơn nữa, chuẩn bị cho các bước xử lý phức tạp hơn về sau trong quá trình XLA. Quá trình này thường được thực hiện bởi các bộ lọc.
* Phân đoạn ảnh (segmentation):phân đoạn ảnh là bước then chốt trong XLA. Giai đoạn này phân tích ảnh thành những thành phần có cùng tính chất nào đó dựa theo biên hay các vùng liên thông. Tiêu chuẩn để xác định các vùng liên thông có thể là cùng màu, cùng mức xám… Mục đích của phân đoạn ảnh là để có một miêu tả tổng hợp về nhiều phần tử khác nhau cấu tạo lên ảnh thô. Vì lượng thông tin chứa trong ảnh rất lớn, trong khi đa số các ứng dụng chúng ta chỉ cần trích một vài đặc trưng nào đó, do vậy cần có một quá trình để giảm lượng thông tin khổng lồ đó. Quá trình này bao gồm phân vùng ảnh và trích chọn đặc tính chủ yếu.
* Trích xuất đặc trưng (feature extraction): Kết quả của bước phân đoạn ảnh thường được cho dưới dạng dữ liệu điểm ảnh thô, trong đó hàm chứa biên của một vùng ảnh, hoặc tập hợp tất cả các điểm ảnh thuộc về chính vùng ảnh đó. Trong cả hai trường hợp, sự chuyển đổi dữ liệu thô này thành một dạng thích hợp hơn cho việc xử lý trong máy tính là rất cần thiết. Để chuyển đổi chúng, câu hỏi đầu tiên cần phải trả lời là nên biểu diễn một vùng ảnh dưới dạng biên hay dưới dạng một vùng hoàn chỉnh gồm tất cả những điểm ảnh thuộc về nó. Biểu diễn dạng biên cho một vùng phù hợp với những ứng dụng chỉ quan tâm chủ yếu đến các đặc trưng hình dạng bên ngoài của đối tượng, ví dụ như các góc cạnh và điểm uốn trên biên chẳng hạn. Biểu diễn dạng vùng lại thích hợp cho những ứng dụng khai thác các tính chất bên trong của đối tượng, ví dụ như vân ảnh hoặc cấu trúc xương của nó. Sự chọn lựa cách biểu diễn thích hợp cho một vùng ảnh chỉ mới là một phần trong việc chuyển đổi dữ liệu ảnh thô sang một dạng thích hợp hơn cho các xử lý về sau. Chúng ta còn phải đưa ra một phương pháp mô tả dữ liệu đã được chuyển đổi đó sao cho những tính chất cần quan tâm đến sẽ được làm nổi bật lên, thuận tiện cho việc xử lý chúng.
* Phân tích dữ liệu (data analysis): Đây là bước cuối cùng trong quá trình XLA. Trong đó có quá trình nhận dạng ảnh có thể được nhìn nhận một cách đơn giản là việc gán nhãn cho các đối tượng trong ảnh. Ví dụ đối với nhận dạng chữ viết, các đối tượng trong ảnh cần nhận dạng là các mẫu chữ, ta cần tách riêng các mẫu chữ đó ra và tìm cách gán đúng các ký tự của bảng chữ cái tương ứng cho các mẫu chữ thu được trong ảnh. Giải thích là công đoạn gán nghĩa cho một tập các đối tượng đã được nhận biết.

Chúng ta cũng có thể thấy rằng, không phải bất kỳ một ứng dụng XLA nào cũng bắt buộc phải tuân theo tất cả các bước xử lý đã nêu ở trên, ví dụ như các ứng dụng chỉnh sửa ảnh nghệ thuật chỉ dừng lại ở bước tiền xử lý. Một cách tổng quát thì những chức năng xử lý bao gồm cả nhận dạng và giải thích thường chỉ có mặt trong hệ thống phân tích ảnh tự động hoặc bán tự động, được dùng để rút trích ra những thông tin quan trọng từ ảnh, ví dụ như các ứng dụng nhận dạng ký tự quang học, nhận dạng chữ viết tay… Một điều kiện tiên quyết và quan trọng cho Thị Giác Máy Tính khác với xử lý ảnh là việc thao tác trên nhiều hình ảnh. Trong khi xử lý ảnh hoàn toàn hoạt động với duy nhất một tấm hình ở dạng số hóa, Computer Vision hoạt động phù hợp hơn trên một luồng hình ảnh có mối quan hệ tạm thời đã biết trước.

***1.2.3 Một số hệ màu cơ bản***

Không gian màu (Colour Space) được hiểu là các mô hình toán để miêu tả màu sắc. Mỗi không gian màu đều có một tác dụng và ứng dụng trong các bài toán khác nhau.

* **Hệ màu RGB**: là không gian màu phổ biến dùng trong máy tính, máy ảnh, điện thoại và nhiều thiết bị kĩ thuật số khác nhau. Không gian màu này khá gần với cách mắt người tổng hợp màu sắc. Nguyên lý cơ bản là sử dụng 3 màu sắc cơ bản R (red - đỏ), G (green - xanh lục) và B (blue - xanh lam) để biểu diễn tất cả các màu sắc. Thông thường, trong mô hình 24 bit mỗi kênh màu sẽ sử dụng 8bit để biểu diễn, tức là giá trị R, G, B nằm trong khoảng 0 - 255. Bộ 3 số này biểu diễn cho từng điểm ảnh, mỗi số biểu diễn cho cường độ của một màu. Với mô hình màu 24bit thì số màu tối đa có thể tạo ra là 255 x 255 x 255 = 16581375 màu. Một điểm cân lưu ý là với các thư viện đọc ảnh và hiển thị ảnh như matplotlib[[11]](#footnote-11), Pillow[[12]](#footnote-12) thì các ảnh được đọc theo RGB tuy nhiên Opencv[[13]](#footnote-13) đọc ảnh theo các kênh BGR.

Diagram, shape, venn diagram, circle

Description automatically generated

Hình 1.3. Hệ màu RGB

* Không gian màu HSV (còn gọi là HSB) là một cách tự nhiên hơn để mô tả màu sắc, dựa trên 3 số liệu:
* H: (Hue) là phần màu của mô hình màu và được biểu thị dưới dạng một số từ 0 đến 360 độ.
* S: (Saturation) Độ bão hòa là lượng màu xám trong màu, từ 0 đến 100 phần trăm. Một hiệu ứng mờ nhạt có thể có được từ việc giảm độ bão hòa về không để giới thiệu nhiều màu xám hơn.
* B (hay V): (Bright hay Value) Độ sáng hay giá trị hoạt động kết hợp với độ bão hòa và mô tả độ sáng hoặc cường độ của màu sắc, từ 0-100 phần trăm, trong đó 0 là hoàn toàn đen và 100 là sáng nhất và cho thấy màu sắc nhất.

A picture containing text, light

Description automatically generated

Hình 1.4. Hệ màu HSV

### **1.2.4 Biễu diễn ảnh**

Ảnh trong thực tế là một ảnh liên tục cả về không gian và giá trị độ sáng. Để có thể xử lý ảnh bằng máy tính thì cần thiết phải tiến hành số hóa ảnh. Quá trình số hóa biến đổi các tín hiệu liên tục sang tín hiệu rời rạc thông qua quá trình lấy mẫu (rời rạc hóa về không gian) và lượng tử hóa các thành phần giá trị mà về nguyên tắc bằng mắt thường không thể phân biệt được hai điểm liền kề nhau. Các điểm như vậy được gọi là các pixel (Picture Element) hay các phần tử ảnh hoặc điểm ảnh. Ở đây cần phân biệt khái niệm pixel hay đề cập đến trong các hệ thống đồ họa máy tính. Để tránh nhầm lẫn ta gọi khái niệm pixel này là pixel thiết bị. Khái niệm pixel thiết bị có thể xém xét như sau: khi ta quan sát màn hình (trong chế độ đồ họa), màn hình không liên tục mà gồm các điểm nhỏ, gọi là pixel. Mỗi pixel gồm một tập tọa độ (x, y) và màu.

Như vậy mỗi ảnh là tập hợp các điểm ảnh. Khi được số hóa nó thường được biểu diễn bởi mảng 2 chiều I(n,p): n là dòng và p là cột.

Về mặt toán học có thể xem ảnh là một hàm hai biến f(x,y) với x, y là các biến tọa độ. Giá trị số ở điểm (x,y) tương ứng với giá trị xám hoặc độ sáng của ảnh (x là các cột còn y là các hàng). Giá trị của hàm ảnh f(x,y) được hạn chế trong phạm vi của các số nguyên dương.

0 ≤ f(x,y) ≤ fmax.

Với ảnh đen trắng mức xám của ảnh có thể được biểu diễn bởi một số như sau:



0

)

(

)

(



*d*

*S*

*c*

*k*

*BW*

*f*

Trong đó SBW() là đặc tính phổ của cảm biến được sử dụng và k là hệ số tỷ lệ xích. Vì sự cảm nhận độ sáng có tầm quan trọng hàng đầu đối với ảnh đen trắng nên *SBW(**)* được chọn giống như là hiệu suất sáng tương đối. Vì *f* biểu diễn công suất trên đơn vị diện tích, nên nó bao giờ cũng không âm và hữu hạn.

*0≤ f ≤ fmax*

Trong đó *fmax* là giá trị lớn nhất mà *f* đạt được. Trong xử lý ảnh, *f* được chia thang sao cho nó nằm trong một phạm vi thuận lợi nào đó.

Thông thường đối với ảnh xám, giá trị *fmax*là 255 ( 28=256) bởi vì mỗi phần tử ảnh được mã hóa bởi một byte. Khi quan tâm đến ảnh màu ta có thể mô tả màu qua ba hàm số: thành phần màu đỏ qua *R(x,y)*, thành phần màu lục qua *G(x,y)* và thành phần màu lam qua *B(x,y)*. Bộ ba giá trị *R*, *G*, và *B* nhận được từ:



0

)

(

)

(



*d*

*S*

*c*

*k*

*R*

*R*



0

)

(

)

(



*d*

*S*

*c*

*k*

*G*

*G*



0

)

(

)

(



*d*

*S*

*c*

*k*

*B*

*B*

Ở đó *SR(* *), SG(* *) và SB(* *)* theo thứ tự là những đặc tính phổ của các cảm biến (bộ lọc) đỏ, lục và lam. *R*, *G*, *B* cũng không âm và hữu hạn.

Ảnh có thể được biểu diễn theo một trong hai mô hình: mô hình Vector hoặc mô hình Raster.

**Mô hình Vector:** Ngoài mục đích tiết kiệm không gian lưu trữ, dễ dàng hiển thị và in ấn, các ảnh biểu diễn theo mô hình vector còn có ưu điểm cho phép dễ dàng lựa chọn, sao chép, di chuyển, tìm kiếm…Theo những yêu cầu này thì kỹ thuật biểu diễn vector tỏ ra ưu việt hơn. Trong mô hình này, người ta sử dụng hướng vector của các điểm ảnh lân cận để mã hóa và tái tạo lại hình ảnh ban đầu. Các ảnh vector được thu nhận trực tiếp từ các thiết bị số hóa như Digitalize hoặc được chuyển đổi từ các ảnh Raster thông qua các chương trình vector hóa.

**Mô hình Raster:** là mô hình biểu diễn ảnh thông dụng nhất hiện nay. Ảnh được biểu diễn dưới dạng ma trận các điểm ảnh. Tùy theo nhu cầu thực tế mà mỗi điểm ảnh có thể được biểu diễn bởi một hay nhiều bit. Mô hình Raster thuận lợi cho việc thu nhận, hiển thị và in ấn. Các ảnh được sử dụng trong phạm vi của đề tài này cũng là các ảnh được biểu diễn theo mô hình Raster.

Khi xử lý các ảnh Raster chúng ta có thể quan tâm đến mối quan hệ trong vùng lân cận của các điểm ảnh. Các điểm ảnh có thể xếp hàng trên một lưới (raster) hình vuông, lưới hình lục giác hoặc theo một cách hoàn toàn ngẫu nhiên với nhau.

Diagram

Description automatically generated with medium confidence

Hình 1.5. Quan hệ trong vùng lân cận giữa các điểm ảnh

Cách sắp xếp theo hình vuông là được quan tâm đến nhiều nhất và có hai loại:

điểm 4 láng giềng (4 liền kề) hoặc 8 láng giềng (8 liền kề). Với điểm 4 láng giềng, một điểm ảnh I(i, j) sẽ có điểm kế cận theo 2 hướng i và j; trong khi đó với điểm 8 láng giềng, điểm ảnh I(i, j) sẽ có 4 điểm kế cận theo 2 hướng i, j và 4 điểm kế cận theo hướng chéo 45 o (Xem hình 1.3).

***1.2.5 Độ phân giải ảnh***

Độ phân giải( Resolution) của ảnh là mật độ điểm ảnh được ấn định trên một ảnh số được hiển thị. Khoảng cách giữa các điểm ảnh phải được chọn sao cho mắt người vẫn thấy được sự liên tục của ảnh. Việc lựa chọn khoảng cách thích hợp tạo nên một mật độ phân giải và được phân bố theo trục x và y trong không gian hai chiều.

***1.2.6 Phạm vi ứng dụng của xử lý ảnh***

Xử lý ảnh đã đem lại nhiều ứng dụng trong nhiều lĩnh vực khác nhau: y học, khoa học hình hình sự, khí tượng thuỷ văn, quản lý, sản xuất...

**1.3** **Một số kiến thức về Data Mining sử dụng trong bài làm**

***1.3.1 Khái niệm về Data Mining***

Data mining, hay được dịch ra là khai thác dữ liệu. Đó là một quá trình khám phá các mẫu trong các tệp dữ liệu rất lớn, liên quan đến các phương thức tại giao điểm của máy học, thống kê học và hệ thống cơ sở dữ liệu. Data mining là một lĩnh vực liên ngành của Khoa học máy tính và thống kê máy tính với mục tiêu tổng thể là trích xuất thông tin (bằng phương pháp thông minh) từ bộ dữ liệu và chuyển đổi thông tin thành cấu trúc dễ hiểu để sử dụng tiếp.

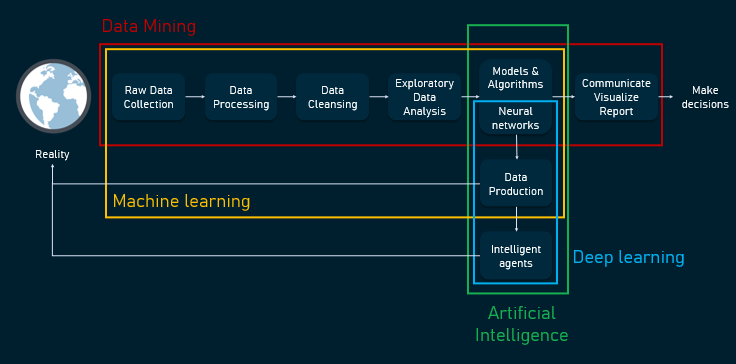
Data mining là bước phân tích của quy trình “khám phá tri thức trong cơ sở dữ liệu”. Ngoài bước phân tích thô, nó còn liên quan đến các khía cạnh quản lý dữ liệu và cơ sở dữ liệu, xử lý trước dữ liệu, cân nhắc mô hình và suy luận, cân nhắc phức tạp, xử lý hậu của các cầu trúc được phát hiện, trực quan hóa và cập nhật trực tuyến.

Trên thực tế, thuật ngữ “Data mining” là một cách gọi sai, vì mục tiêu là khai thác các mẫu và kiến thức từ một lượng lớn dữ liệu, chứ không phải khai thác (mining) dữ liệu. Tuy nhiên, đây là một tên gọi thông dụng và thường được áp dụng cho bất kỳ dạng dữ liệu hoặc xử lý thông tin quy mô lớn nào (thu thập, trích xuất, lưu trữ, phân tích và thống kê) cũng như mọi ứng dụng của hệ thống hỗ trợ quyết định máy tính, bao gồm cả trí tuệ nhân tạo và trí tuệ kinh doanh.

Thống kê học liên quan các con số và định lượng dữ liệu. Có nhiều công cụ để tìm các thuộc tính có liên quan của dữ liệu nhưng điều này khá gần với toán học thuần túy.

Machine Learning sử dụng các kỹ thuật Data Mining và các thuật toán học tập khác để xây dựng các mô hình về những gì đang xảy ra đằng sau một số dữ liệu để nó có thể dự đoán kết quả trong tương lai. Toán học là cơ sở cho nhiều thuật toán, nhưng điều này hướng tới lập trình nhiều hơn.

Ta có thể nhìn thấy mối quan hệ giữa các lĩnh vực AI, ML, DL, Data Mining trong ngành Khoa học máy tính qua sơ đồ sau:



Hình 1.6. Mối liên hệ giữa AI, ML, DL, Data Mining trong Khoa học máy tính

***1.3.2 Các phép toán sử dụng trong bài làm***

**a) Biến đổi dữ liệu**

Biến đổi dữ liệu (data transformation) là ứng dụng các hàm toán học xác định cho từng điểm trong tập dữ liệu, nghĩa là, mỗi điểm dữ liệu zᵢ được thay thế bằng giá trị biến đổi yᵢ = f, trong đó f là một hàm.

**Phân phối chuẩn (normal distribution):**

Là dạng phân phối xác suất quan trọng nhất trong thống kê.

Một biến có dạng phân phối chuẩn khi phân phối xác suất của nó có dạng đường cong hình chuông “bell-shaped curve”. Mean nằm chính giữa và các giá trị của biến xoay quanh mean. Trong đó các giá trị mean (giá trị trung bình), median (trung vị), mode (yếu vị) có giá trị như nhau và được xác định bởi 2 tham số: Giá trị trung bình (mean) và phương sai (variance).

Tuy vậy không phải lúc nào dữ liệu trong thực tế cũng theo phân phối chuẩn, chúng thường bị lệch và làm cho kết quả phân tích không hợp lệ, để giải quyết vấn đề này ta sử dụng phép biến đổi logarit.

Diagram

Description automatically generated

Hình 1.7. Minh họa phân phối chuẩn

**Logarithm transformation (Phép biến đổi logarit)**:

Phép biến đổi logarit là một phương pháp biến đổi dữ liệu, trong đó nó thay thế mỗi biến x bằng một Log(x). Việc lựa chọn cơ số logarit thường do nhà phân tích phụ thuộc vào mục đích của mô hình thống kê. Trong bài viết này, ta sẽ tập trung vào sự biến đổi logarit tự nhiên (cơ số e), ký hiệu là ln.

Khi dữ liệu liên tục ban đầu không tuân theo đường cong hình chuông, ta có thể dùng phép biến đổi logarit lên dữ liệu này để làm cho dữ liệu này “chuẩn” nhất có thể để kết quả phân tích thống kê từ dữ liệu này trở nên hợp lệ hơn. Nói cách khác, việc chuyển đổi logarit làm giảm hoặc loại bỏ độ lệch của dữ liệu ban đầu. Lưu ý quan trọng ở đây là dữ liệu gốc phải tuân theo hoặc gần đúng theo phân phối log-normal. Nếu không, phép biến đổi logarit sẽ không hoạt động.

Chart, histogram

Description automatically generated

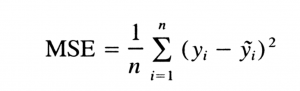
Hình 1.8. Minh họa tác động của phép biến đổi logarit

**b) Sai số toàn phương trung bình (Mean squared error):**

Mean squared error, viết tắt là MSE, là trung bình của bình phương các sai số, tức là sự khác biệt giữa các ước lượng và những gì được đánh giá. MSE là một hàm rủi ro, tương ứng với giá trị kỳ vọng của sự mất mát sai số bình phương hoặc mất mát bậc hai. Sự khác biệt xảy ra do ngẫu nhiên, hoặc vì các ước lượng không tính đến thông tin có thể cho ra một ước tính chính xác hơn.

MSE đánh giá chất lượng của một ước lượng (ví dụ, một hàm toán học lập bản đồ mẫu dữ liệu của một tham số của dân số từ đó các dữ liệu được lấy mẫu) hoặc một yếu tố dự báo (ví dụ, một bản đồ chức năng có số liệu vào tùy ý để một mẫu của các giá trị của một số biến ngẫu nhiên). Định nghĩa của một MSE khác với những gì tương ứng cho dù là một trong những mô tả một ước lượng, hay một yếu tố dự báo.

**Công thức tính MSE:**



Với: yi  là biến độc lập, ỹi là giá trị ước lượng, n là số các giá trị dự báo

**1.4 Phần mềm Nicepage**

Nicepage là một trình xây dựng trang web mới mẻ, khác biệt với các trình xây dựng phổ biến khác trên thị trường. Nó nhằm mục đích trả lại quyền tự do sáng tạo cho các nhà thiết kế web, với trình chỉnh sửa kéo và thả nâng cao cho phép nhà thiết kế web tự do sáng tạo tối đa.

**Các tính năng chính:**

* Tự do kéo thả các thành phần của trang web
* Hỗ trợ Responsive
* Hỗ trợ các hiệu ứng animation
* Hỗ trợ các khối tạo sẵn
* Hỗ trợ nhiều mẫu trang web tạo sẵn

Cung cấp miễn phí giới hạn tính năng và bản trả phí sử dụng nhiều tính năng hơn. Tuy nhiên với nhu cầu thiết kế trang web đơn giản với thời gian nhanh chóng và miễn phí nên tôi chọn phần mềm này để thiết kế trang web của mình.

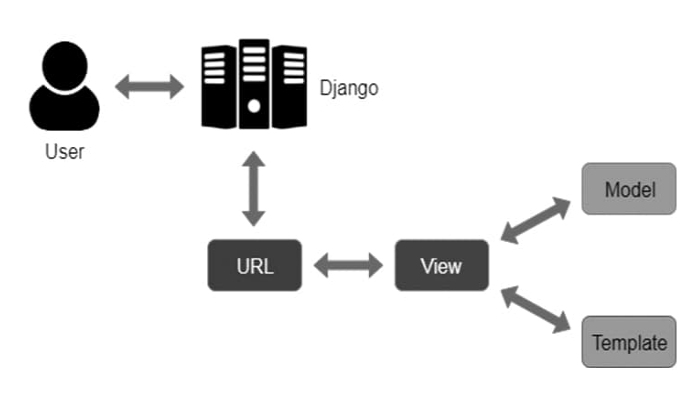
**1.5 Django Framework**

### ***1.5.1 Ưu điểm của Django:***

1. Sử dụng ngôn ngữ lập trình Python
2. Django đang là giải pháp cốt lỗi trong các lĩnh vực: IOT, các Công ty Blue chip, IT giants ([NASA](https://vinasupport.com/tag/nasa/), [Google](https://vinasupport.com/tag/google-inc/),…), các công ty tài chính
3. Sử dụng ORM, [Mô hình MVC](https://vinasupport.com/tag/mo-hinh-mvc/) (MTV)
4. Hỗ trợ Multi-Site, Đa ngôn ngữ
5. Nhiều tài liệu và cộng đồng hỗ trợ
6. Hỗ trợ giao diện quản lý Admin
7. Dễ dàng mở rộng

### ***1.5.2 Mô hình MTV (Model-Template-Views) trong Django***

Django sử dụng mô hình MTV tương tự như mô hình MVC (Model-View-Controller) trong các framework khác.



Hình 1.9. Mô hình MTV của Django

Bảng 1.1. Mô tả chức năng các thành phần trong mô hình MTV.

| **Thành phần** | **Mô tả** |
| --- | --- |
| **Model** | Nơi thiết kế ra những table cho database, từ đó Django ORM đã cung cấp những phương thức xử lý, nghiệp vụ lên database |
| **Template** | Là những mẫu được thiết kế ra và xử lý output ra mã html/css cho trang web |
| **Views** | Các hàm để xử lý khi có request từ người dùng |

### ***1.5.3 Cấu trúc Project Django***

* **\_\_init\_\_.py**: đây là 1 file cơ bản trong Python dùng để biến folder chứa nó thành package, giúp tao có thể import
* **setttings.py**: đây là file cấu hình project. (VD: cấu hình database, đặt múi giờ, cài thêm thư viện, ...)
* **urls.py**: đây là file giúp chúng ta tạo các đường dẫn urls của trang web để liên kết các webpage lại với nhau
* **wsgi.py**: đây là file giúp chúng ta deploy project lên server

1. https://nhandan.vn/baothoinay-dothi/tim-co-che-cho-cac-diem-do-xe-637883/ [↑](#footnote-ref-1)
2. Facebook, Inc. là một công ty truyền thông xã hội và công nghệ Mỹ có trụ sở tại Menlo Park, California. [↑](#footnote-ref-2)
3. TensorFlow là một thư viện phần mềm mã nguồn mở dành cho máy học trong nhiều loại hình tác vụ nhận thức và hiểu ngôn ngữ. [↑](#footnote-ref-3)
4. Keras là một thư viện phần mềm mã nguồn mở cung cấp giao diện Python cho các mạng nơ-ron nhân tạo. [↑](#footnote-ref-4)
5. Alan Mathison Turing OBE FRS là một nhà toán học, logic học và mật mã học người Anh, được xem là cha đẻ của ngành khoa học máy tính. [↑](#footnote-ref-5)
6. Amazon Alexa, hay được gọi tắt là Alexa, là trợ lý ảo được phát triển bởi Amazon [↑](#footnote-ref-6)
7. Pinterest là website chia sẻ ảnh theo dạng mạng xã hội [↑](#footnote-ref-7)
8. Facebook là một mạng xã hội phổ biến nhất hiện nay [↑](#footnote-ref-8)
9. Amazon.com, Inc. là một công ty công nghệ đa quốc gia của Mỹ có trụ sở tại Seattle, Washington [↑](#footnote-ref-9)
10. AlphaGo là chương trình máy tính cờ vây do Google DeepMind phát triển tại London [↑](#footnote-ref-10)
11. Matplotlib là một thư viện vẽ đồ thị cho ngôn ngữ lập trình Python và phần mở rộng toán học số NumPy của nó [↑](#footnote-ref-11)
12. Pillow là một phần của PIL- Python Image Library, thư viện xử lý ảnh của Python [↑](#footnote-ref-12)
13. OpenCV là một thư viện các chức năng lập trình chủ yếu nhắm vào thị giác máy tính thời gian thực [↑](#footnote-ref-13)