**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 Django để xây dựng website giới thiệu sản phẩm, 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 và các bạn học cùng lớp 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 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 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 3. Bãi đỗ xe tự động tại Đà Nẵng. Ảnh Internet

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

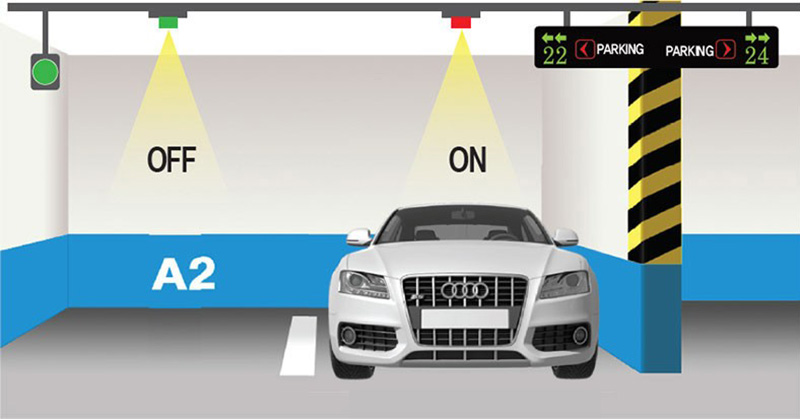
Cho dù là bãi đỗ xe truyền thống hay tự động thì trong một bãi đỗ xe hoàn chỉnh không thể thiếu hệ thống hướng dẫn đỗ xe (Viết tắt là PGS), các loại hình phổ biến tại Việt Nam hiện nay:

* **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 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 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 làm nên. 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. 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 và Keras 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.

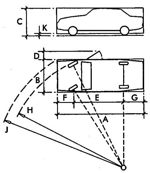
Django là một trong số những web framework bậc cao miễn phí, là mã nguồn mở được tạo ra bởi ngôn ngữ Python dựa trên mô hình mô hình MTV (gồm Model-Template-Views). Hiện framework này được phát triển, quản lý bởi Django Software Foundation. Django ra đời với mục tiêu hỗ trợ thiết kế các website phức 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.

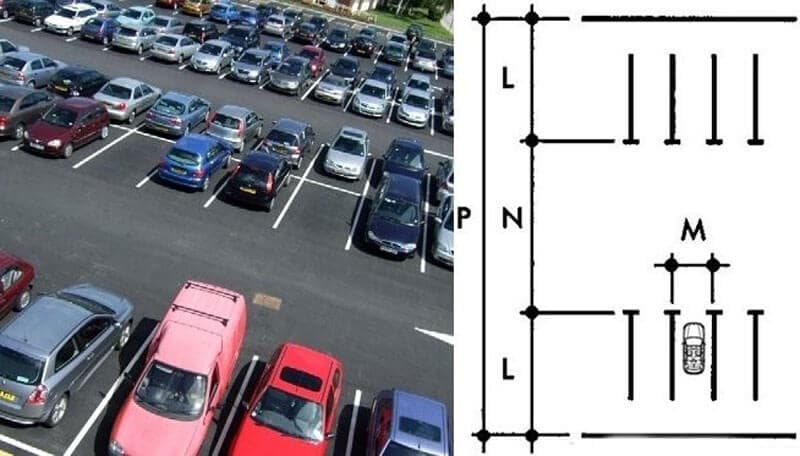
### ***Tiêu chuẩn kích thước của bãi đỗ xe ô tô***

* Kích thước ô tô tiêu chuẩn:



Hình 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 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:

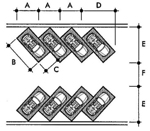
* Chiều rộng chỗ đỗ xe ô tô tiêu chuẩn : Từ 2.4m đến 2.75m
* Chiều rộng chỗ đỗ xe ô tô thời gian ngắn: Từ 2.3m đến 2.6m
* Chiều rộng chỗ đỗ xe ô tô thời gian dài: Từ 2.5m đến 2.75m
* Chiều rộng chỗ đỗ xe ô tô dành cho người khuyết tật: Từ 3m đến 3.5m

+ 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:



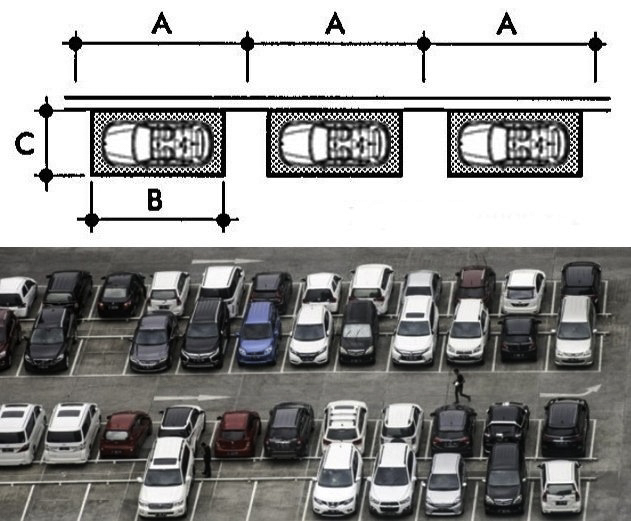
Hình 8. Tiêu chuẩn thiết kế bãi đỗ xe 45 độ

+ Tiêu chuẩn:

Bảng 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 9. Tiêu chuẩn thiết kế bãi đỗ xe song song

+ Tiêu chuẩn:

Bảng 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 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 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 lĩnh vực Khoa Học Dữ Liệu, Thị Giác Máy Tính.
* Giới thiệu về Transfer Learning.
* 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 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 và website trong tương lai.

Tài liệu tham khảo.

1. https://nhandan.vn/baothoinay-dothi/tim-co-che-cho-cac-diem-do-xe-637883/ [↑](#footnote-ref-1)