**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG THPT TRẤN BIÊN**



**BẢN THUYẾT MINH SẢN PHẨM DỰ THI**

**CUỘC THI LẬP TRÌNH DÀNH CHO HỌC SINH TRUNG HỌC PHỔ THÔNG VÀ SẢN PHẨM SÁNG TẠO CÔNG NGHỆ THÔNG TIN DÀNH CHO SINH VIÊN CAO ĐẲNG, ĐẠI HỌC NĂM 2024 (HUE-ICT CHALLENGE-2024)**

**Tên sản phẩm:**

**PHẦN MỀM NHẬN DIỆN NGÔN NGỮ KÝ HIỆU TRỰC TUYẾN HỖ TRỢ NGƯỜI KHUYẾT TẬT**

**Lĩnh vực:**

**Phần mềm**

**Học sinh thực hiện: PHAN LÊ QUANG THẢO**

MỤC LỤC

[**PHẦN MỘT. NỘI DUNG** 2](#_Toc161089706)

[**I. TÊN SẢN PHẨM** 2](#_Toc161089707)

[**II. Ý TƯỞNG** 3](#_Toc161089708)

[**III. CHI TIẾT SẢN PHẨM** 3](#_Toc161089709)

[**1. Tính mới của sản phẩm** 3](#_Toc161089710)

[**3. Cài đặt thuật toán nhận diện ngôn ngữ ký hiệu cho phần mềm** 3](#_Toc161089711)

[**4. Nguyên tắc hoạt động của phần mềm** 3](#_Toc161089712)

[**4.2. Quy trình huấn luyện mô hình học máy nhận diện ngôn ngữ ký hiệu** 4](#_Toc161089713)

[**4.3. Thiết kế giao diện phần mềm** 6](#_Toc161089714)

[**4.4. Quy trình nhận diện ngôn ngữ ký hiệu thành văn bản** 6](#_Toc161089715)

[**4.5. Quy trình nhận diện giọng nói thành ngôn ngữ ký hiệu** 7](#_Toc161089716)

[**4.6. Đánh giá độ hiệu quả mô hình nhận diện ngôn ngữ ký hiệu** 8](#_Toc161089717)

[**PHẦN HAI. HIỆU QUẢ ĐẠT ĐƯỢC CỦA SẢN PHẨM** 9](#_Toc161089718)

[**1. Kết quả nghiên cứu** 9](#_Toc161089719)

[**2. Kết luận nghiên cứu** 9](#_Toc161089720)

[**PHẦN BA. TÀI LIỆU NGHIÊN CỨU** 9](#_Toc161089721)

[**PHẦN BỐN. CAM KẾT VỀ BẢN QUYỀN SẢN PHẨM** 10](#_Toc161089722)

# **PHẦN MỘT. NỘI DUNG**

# **I. TÊN SẢN PHẨM**

- Phần mềm nhận diện ngôn ngữ ký hiệu trực tuyến hỗ trợ người khuyết tật

# **II. Ý TƯỞNG**

 - Nhận thức được sự khó khăn trong giao tiếp trực tuyến của người khuyết tật nên em quyết định phát triển 1 phần mềm nhận diện ngôn ngữ trực tuyến giúp người khuyết tật giao tiếp dễ dàng hơn.

# **III. CHI TIẾT SẢN PHẨM**

## **1. Tính mới của sản phẩm**

- Phần mềm sử dụng song song 2 thuật toán là LSTM và YOLOv8 để nhận diện từ đó làm cho phần mềm hoạt động chính xác hơn

- Phần mềm giống như là 1 ứng dụng video call được tích hợp them phần nhận diện. Tính năng này chưa có trên các phần mềm hay ứng dụng nhận diện ngôn ngữ ký hiệu trên thị trường và các cuộc thi
 **2. Xây dựng cơ sở dữ liệu ngôn ngữ ký hiệu.**

- Cơ sở dữ liệu ngôn ngữ ký hiệu là một tập hợp các dữ liệu về các dấu hiệu ngôn ngữ ký hiệu. Các dữ liệu này bao gồm:

* + Hình ảnh và cử chỉ tay của các dấu hiệu ngôn ngữ ký hiệu.
	+ Nhãn của các hình ảnh, nhãn có thể hình dung đơn giản là 1 tập dữ liệu mà mô hình học máy dựa vào đó và đưa ra kết quả của hình ảnh nhận diện được, việc dán nhãn cho hình ảnh ngôn ngữ ký hiệu đã thu thập là một quá trình tương đối dài, cần sự tỉ mỉ và cẩn thận vì quá trình này quyết định 1 phần kết quả của mô hình học máy trong quá trình huấn luyện.
* Cơ sở dữ liệu ngôn ngữ ký hiệu là phương tiện chính được sử dụng để huấn luyện thuật toán nhận diện ngôn ngữ ký hiệu.

## **3. Cài đặt thuật toán nhận diện ngôn ngữ ký hiệu cho phần mềm**

* Thuật toán nhận diện ngôn ngữ ký hiệu là một chương trình máy tính có khả năng nhận dạng các dấu hiệu ngôn ngữ ký hiệu từ các dữ liệu đầu vào (Tập hình ảnh và tập nhãn đã được thu thập, chuẩn bị từ trước ). Thuật toán được sử dụng để huấn luyện mô hình cụ thể là dùng thư viện Ultralytics phiên bản YOLOv8 để ghi nhận hình ảnh sau đó bắt đầu áp dụng các thuật toán trong thư viện để huấn luyện mô hình. Ngoài ra phần mềm còn được tích hợp thêm thuật toán LSTM để nhận diện tốt hơn. LSTM là mô hình được huấn luyện từ dữ liệu các đốt ngón tay của tay người dùng.

## **4. Nguyên tắc hoạt động của phần mềm**

- Phương pháp nghiên cứu chủ yếu là nghiên cứu tài liệu của Ultralytics. Các tài liệu nghiên cứu bao gồm tài liệu của Ultralytics, các bài báo khoa học, sách chuyên khảo, luận văn, trang web và các kênh Youtube về Machine Learning hay cụ thể là Object Detection, Computer Vision.

**4.1. Quy trình thu thập và xây dựng tập dữ liệu cho dự án**

* Em áp dụng mô hình nhận diện tay người của Mediapipe để nhận diện các ngôn ngữ ký hiệu thay vì gán nhãn 1 cách thủ công chính vì thế mà em đã tự động hóa được quá trình khó khăn của 1 dự án nhận diện.
* Cách tự động hóa được thực hiện như sau. Đầu tiên em sử dụng model của Mediapipe để chạy, khi chạy thì sẽ đặt cho lần chạy đó 1 lớp ngôn ngữ ký hiệu, sau đó bắt đầu trích xuất kết quả trong quá trình nhận diện từ đó ghi nhận lại các tọa độ của bàn tay và đưa ra các chỉ số của hộp giới hạn là x1 x2 y1 y2, tiếp theo dùng công thức toán chuyển đổi từ hộp giới hạn sang YOLO format là x y w h, sau khi có tọa độ với YOLO format thì chương trình sẽ viết xuống tệp text 4 tham số đó cùng với chỉ số class ban đầu đã đặt cho lần chạy đó, sau quá trình này thì dự án đã có được 1 tập dữ liệu lớn, chất lượng cao và tiết kiệm được rất nhiều thời gian, từ đó có phát triển được nhiều thứ tốt hơn trong dự án ( Tổng cộng có hơn 30 lớp đối tượng và hơn 30.000 ảnh ). Phần mã cho quá trình tự động hóa thu thập dữ liệu này để ở phần phụ lục của bản báo cáo.
* Sau khi có tập dữ liệu thì em bắt tay vào sắp xếp phù hợp và đưa dữ liệu lên các Online Notebook như Google Colab hay Kaggle Notebook để tận dụng được GPU miễn phí từ các nhà cung cấp này, việc này giúp dự án giảm được số tiền siêu lớn và tăng tốc độ huấn luyện. Các nền tảng này đã cung cấp các GPU rất tốt cho việc huấn luyện mô hình liên quan tới hình ảnh như Tesla T4 hay Tesla P100.

## **4.2. Quy trình huấn luyện mô hình học máy nhận diện ngôn ngữ ký hiệu**

* Quá trình huấn luyện sử dụng thuật toán nhận diện YOLOv8, thuật toán sẽ hoạt động rất tốt khi được huấn luyện trên một tập dữ liệu lớn gồm các hình ảnh chất lượng và nhãn đối tượng chính xác, chi tiết. Quá trình huấn luyện sử dụng thuật toán tối ưu hóa stochastic gradient descent (SGD).
* Trong YOLOv8, stochastic gradient descent (SGD) được sử dụng để cập nhật trọng số của mô hình. Trọng số là những giá trị quyết định cách mô hình hoạt động. Bằng cách cập nhật trọng số, mô hình có thể trở nên chính xác hơn.
* Sử dụng YOLOv8 để phát hiện các đối tượng trong một bức ảnh, hay các ngôn ngữ ký hiệu trong dự án nói riêng. Nếu mô hình không phát hiện ra tất cả các đối tượng nào hết, thì hàm mất mát sẽ cao hơn ngưỡng cho phép. Stochastic gradient descent sẽ sử dụng hàm mất mát này để cập nhật trọng số của mô hình.
* Ví dụ: Nếu mô hình không phát hiện ra chữ ‘H’ trong bức ảnh, thì hàm mất mát sẽ cao. Stochastic gradient descent sẽ cập nhật trọng số của mô hình để giúp mô hình phát hiện ra chữ ‘H’ trong lần dự đoán tiếp theo. Việc cập nhật trọng số này diễn ra trong quá trình huấn luyện mô hình, cữ mỗi lần dự đoán trên tập test thì SGD sẽ có được giá trị mà hàm mất mát trả về cao thì SGD sẽ thay đổi trọng số, sau nhiều lần thử trên tập test thì sẽ có 1 trọng số hoàn hảo cho mô hình, từ đó góp thêm phần chính xác cho mô hình.
* Đầu vào của mô hình YOLOv8 là một nguồn có thể là ảnh, một video hoặc có thể là webcam, mô hình sẽ nhận dạng nguồn đó có đối tượng nào hay không, sau đó sẽ xác định tọa độ của đối tượng trong bức ảnh, từ đó đưa ra được độ tin cậy của đối tượng ( Confidence ) và lớp đối tượng ( Class ).
* YOLOv8 sẽ chia hình ảnh thành một số khu vực nhỏ, sau đó dự đoán các đối tượng có thể có mặt trong mỗi khu vực đó. Mô hình sử dụng mạng convolutional neural network (CNN) để trích xuất các đặc trưng từ hình ảnh, sau đó sử dụng các thuật toán học máy để dự đoán xác suất xuất hiện của từng đối tượng trong mỗi khu vực.
* Có thể hiểu nôm na rằng CNN hoạt động bằng cách phân tích các ô nhỏ trong ảnh, gọi là các pixel. Mỗi pixel có một giá trị, đại diện cho mức độ sáng hoặc tối của pixel đó. CNN sử dụng các phép toán gọi là tích chập để phân tích các giá trị của các pixel này.
* Các phép tích chập này giúp CNN phát hiện các đặc trưng trong ảnh. Ví dụ, CNN có thể phát hiện các đường thẳng, góc, hoặc các hình dạng khác.
* Sau khi phát hiện các đặc trưng, CNN sử dụng các phép toán khác để xác định vị trí và loại của các ngôn ngữ ký hiệu trong ảnh.
* Để dễ hiểu, có thể ví CNN như một con người đang nhìn vào một bức ảnh. Con người sẽ nhìn vào các pixel của bức ảnh và cố gắng phát hiện các đặc trưng, chẳng hạn như các đường thẳng, góc, hoặc các hình dạng khác. Sau khi phát hiện các đặc trưng, con người sẽ sử dụng các kiến thức của mình để xác định vị trí và loại của các đối tượng trong bức ảnh.
* Ví dụ, khi nhìn thấy một bức ảnh có một người đang đứng, con người sẽ phát hiện các đặc trưng như sau :
	+ Các đường thẳng tạo thành khuôn mặt, cơ thể, và cánh tay của người đó.
	+ Các góc tạo thành mắt, mũi, miệng, và khuỷu tay của người đó.
	+ Các hình dạng tạo thành tóc, quần áo, và giày của người đó.

-  Sau khi phát hiện các đặc trưng này, con người sẽ sử dụng kiến thức của mình để xác định rằng đó là một người.

- CNN hoạt động theo cách tương tự như con người. CNN sử dụng các phép toán tích chập để phát hiện các đặc trưng trong ảnh, sau đó sử dụng các kiến thức của mình để xác định vị trí và loại của các đối tượng trong ảnh.

- Trong mô hình YOLOv8 thì các đặc trưng được CNN phát hiện bao gồm các đặc trưng của ngôn ngữ ký hiệu như các đường thẳng tạo thành các nét của ngôn ngữ ký hiệu, các hình dạng tạo thành các vùng của ngôn ngữ ký hiệu

- Sau khi phát hiện các đặc trưng này, CNN sử dụng các kiến thức của mình để xác định loại ngôn ngữ ký hiệu.

- Nhờ sử dụng CNN, YOLOv8 có thể phát hiện ngôn ngữ ký hiệu trong ảnh một cách chính xác và hiệu quả.

- Tuy nhiên, khi sử dụng YOLOv8 để phát hiện ngôn ngữ ký hiệu, cần lưu ý một số điểm như ngôn ngữ ký hiệu là một loại đối tượng đặc biệt, có các đặc điểm khác biệt so với các đối tượng thông thường. Ví dụ, ngôn ngữ ký hiệu thường có kích thước nhỏ, các cạnh và góc sắc nét, và có thể xuất hiện trong nhiều tư thế và vị trí khác nhau.

- Vì vậy để mô hình YOLOv8 có thể phát hiện ngôn ngữ ký hiệu hiệu quả, cần có một tập dữ liệu huấn luyện phù hợp. Tập dữ liệu này cần bao gồm nhiều hình ảnh ngôn ngữ ký hiệu với nhiều tư thế và vị trí khác nhau kết hợp với 1 điều kiện sáng tốt.

- Quá trình hoạt động YOLOv8 với dataset là các hình ảnh ngôn ngữ ký hiệu:

 - Bước 1: Đầu tiên, hình ảnh được chia thành một lưới 3x3x13, với mỗi ô lưới có kích thước 32x32x32 pixel.

 - Bước 2: Sau đó, mạng CNN được sử dụng để trích xuất các đặc trưng từ mỗi ô lưới. Các đặc trưng này được sử dụng để dự đoán các đối tượng có thể có mặt trong mỗi ô lưới.

 - Bước 3: Dự đoán của mô hình bao gồm:

 - Xác suất của việc có một đối tượng trong ô lưới đó.

 - Loại đối tượng có thể có mặt trong ô lưới đó.

 - Tọa độ của tâm đối tượng trong ô lưới đó.

 - Kích thước của đối tượng trong ô lưới đó.

 - Bước 4: Mô hình sẽ sử dụng các thuật toán học máy để kết hợp các dự đoán từ các ô lưới khác nhau để xác định vị trí và kích thước của các đối tượng trong hình ảnh.

 - Bước 5: Cuối cùng, mô hình sẽ trả về kết quả của quá trình phát hiện đối tượng.

## **4.3. Thiết kế giao diện phần mềm**

* Tiếp nối quá trình huấn luyện là phần thiết kế giao diện phần mềm. Ở đây dự án dùng thư viện PyQt5 và PyQt5-Tools của ngôn ngữ Python để thiết kế, phần mềm PyQt rất dễ sử dụng, em đã sử dụng thêm kiến thức CSS của mình để giúp cho phần giao diện được đẹp hơn và sinh động hơn. CSS là 1 ngôn ngữ dùng để định dạng , dùng cho web và ứng dụng máy tính.
* Hơn nữa, vì muốn phát triển 1 phần mềm miễn phí nên em đã quyết định không làm phần đăng nhập hay đăng ký mà làm thẳng vào giao diện chính. Tuy vậy nhưng em sẽ vẫn đảm bảo được sự bảo mật thông tin về hình ảnh khi sử dụng và dữ liệu khi sử dụng của người dùng.

## **4.4. Quy trình nhận diện ngôn ngữ ký hiệu thành văn bản**

* Đầu tiên, em viết hàm chức năng cho các nút, các nhãn trong giao diện ( Back End ), các hàm chức năng này là các hàm dùng để thêm các việc cần làm cho các nút, các nhãn trong phần giao diện.
* Phương thức hoạt động của mô hình nhận diện ngôn ngữ ký hiệu trong phần mềm**:**
	+ Khi người dùng nhấn vào nút start camera trên giao diện thì hệ thống sẽ gọi webcam, hoặc camera tùy vào thiết bị.
	+ Sau đó hệ thống áp dụng mô hình đã huấn luyện chạy song song với webcam để có thể nhận diện thời gian thực.
	+ Khi nhận diện được tên ký hiệu thì hệ thống sẽ dựa vào độ tin cậy của ký hiệu đang được nhận diện đó ( Confidence Score of Class ).
	+ Nếu như độ tin cậy của ký hiệu lớn hơn hoặc bằng 0.85 thì em sẽ đưa tên của ký hiệu ngôn ngữ đó lên label text.
	+ Phần code cho chức năng chính này được để ở phần phụ lục của bản báo cáo.
* Phát triển chức năng chức năng nhận diện giọng nói và chuyển thành ngôn ngữ ký hiệu.

## **4.5. Quy trình nhận diện giọng nói thành ngôn ngữ ký hiệu**

* Vì nhận thấy rằng nếu phần mềm chỉ có phần nhận diện ngôn ngữ ký hiệu để giúp người khuyết tật giao tiếp với người bình thường là chưa đủ hoàn thiện nên em đã quyết định phát triển phần chức năng thứ 2 nữa là chức năng nhận diện giọng nói sau đó chuyển đổi giọng nói thành ngôn ngữ ký hiệu dưới dạng video.
* Cách hoạt động của chức năng này như sau:
	+ Khi nhấn vào nút bắt đầu ghi âm thì chương trình sẽ sử dụng module speech-recognition của google kết hợp với python. Module này hoạt động theo các bước sau:
		- Đầu tiên, hệ thống thu thập âm thanh trực tiếp từ micro khi người dùng nhấn vào nút ghi âm
		- Kế tiếp, module xử lý âm thanh bằng thuật toán có sẵn được Google huấn luyện trước
		- Cuối cùng thì module xác định cụm từ và từ có trong âm thanh, dựa vào độ tin cậy của cụm từ hay từ mà module sẽ đưa ra kết quả cuối cùng là output\_text
	+ Sau khi đã có output\_text thì hệ thống sẽ duyệt biến char qua từng ký tự của output\_text sau đó thêm các đường dẫn vào img\_list.
	+ Cuối cùng hệ thống sử dụng hàm ImageSequenceClip của module moviepy để xử lý video. Việc xử lý video hoạt động theo các bước như sau:
		- Tạo 1 frame list là 1 tập chứa các ảnh, sử dụng cv2.imread để đọc ảnh thông qua các đường dẫn ở img\_list.
		- Duyệt qua từng ảnh trong frame list, nếu ảnh đó hợp lệ thì sử dụng cv2.cvtColor để đổi màu ảnh từ dạng BGR qua RGB. Chuyển xong thì hệ thống thêm ảnh vào result.
		- Kế đến hệ thống sử dụng ImageSequenceClip để xử lý video.
		- Cuối cùng phần mềm sẽ tự động sao lưu video với code là ‘libx264’ vào đường dẫn đã cho trước.
	+ Sau khi đã dựng được video thì em xây dựng hàm chuyển video lên label:
		- Phần mềm sử dụng pyqtSignal để làm tín hiệu, kế tiếp dùng hàm cv2.VideoCapture để đọc từng khung hình trong video sau đó hiển thị lên img\_label từng khung hình bằng lệnh emit trong PyQt.
		- Phần mềm liên kết phần đưa video lên label với đối tượng nút play\_video để khi nhấn Play Video thì trên giao diện thì video sẽ được phát.
		- Phần mềm có 1 hàm stop\_video để dừng lại video, được hoạt động bằng cách dùng hàm stop của phương thức QTheard.
		- Ngoài ra hệ thống có làm thêm 1 video về chữ cái với các chữ đen nền trắng nhằm làm cho phần hoạt động của tính năng này thêm sinh động và bắt mắt hơn. Quá trình tạo ra video này cũng giống với quá trình tạo ra video các ký hiệu ngôn ngữ tay người.
		- Phần code cho chức năng này được để ở phần phụ lục của bản báo cáo.

## **4.6. Đánh giá độ hiệu quả mô hình nhận diện ngôn ngữ ký hiệu**

* Độ hiệu quả của hệ thống nhận diện ngôn ngữ ký hiệu thông qua các chỉ số sau khi huấn luyện.
	+ Loss: Là hàm mất mát của mô hình. Giá trị của hàm mất mát càng nhỏ thì mô hình càng tốt.
	+ Precision: Là độ chính xác của mô hình. Giá trị của độ chính xác càng cao thì mô hình càng chính xác.
	+ Recall: Là độ hoàn thành của mô hình. Giá trị của độ hoàn thành càng cao thì mô hình càng hoàn thành.
	+ mAP (mean average precision): Là chỉ số trung bình của độ chính xác và độ hoàn thành. Giá trị của mAP càng cao thì mô hình càng tốt.
* Ý nghĩa của các chỉ số:
	+ Loss: Loss là một hàm toán học được sử dụng để đo lường sự khác biệt giữa kết quả dự đoán của mô hình và kết quả thực tế. Giá trị của loss càng nhỏ thì mô hình càng tốt.
	+ Precision: Precision là tỷ lệ giữa số đối tượng được mô hình phát hiện chính xác và tổng số đối tượng được mô hình phát hiện. Giá trị của precision càng cao thì mô hình càng chính xác.
	+ Recall: Recall là tỷ lệ giữa số đối tượng được mô hình phát hiện chính xác và tổng số đối tượng thực tế có trong ảnh. Giá trị của recall càng cao thì mô hình càng hoàn thành.
	+ mAP: mAP là chỉ số trung bình của precision và recall. Giá trị của mAP càng cao thì mô hình càng tốt.
* Ngoài các chỉ số này ra thì YOLOv8 sau quá trình huấn luyện còn trả về IOU:
	+ IOU là viết tắt của Intersection over Union, hay còn gọi là tỷ lệ giao thoa trên tổng diện tích. Trong nhận dạng ngôn ngữ ký hiệu, IOU là một thước đo mức độ khớp giữa vùng dự đoán của mô hình và vùng thực tế của đối tượng trong ảnh.
	+ IOU cao cho thấy mô hình đã dự đoán chính xác vị trí và kích thước của đối tượng.
	+ Ví dụ: nếu một mô hình nhận dạng ngôn ngữ ký hiệu dự đoán rằng một bàn tay ở vị trí (100, 100) đến (200, 200), và vùng thực tế của bàn tay đó ở vị trí (110, 110) đến (190, 190), thì IOU sẽ là 0,9.
* Tóm lại, IOU là một thước đo quan trọng trong nhận dạng ngôn ngữ ký hiệu, giúp mô hình xác định chính xác hơn các đối tượng trong ảnh.

# **PHẦN HAI. HIỆU QUẢ ĐẠT ĐƯỢC CỦA SẢN PHẨM**

## **1. Kết quả nghiên cứu**

* Cho ra được dữ liệu, mô hình huấn luyện.
* Nhận diện trên thời gian thực ( Real-time detect object ).
* Tốc độ xử lý dữ liệu nhanh và có độ chính xác cao.
* Phần mềm hoạt động mượt mà, chất lượng video cao, dễ sử dụng, thân thiện với người dùng và miễn phí.
* Biểu diễn được kết quả rõ ràng.

## **2. Kết luận nghiên cứu**

* Theo em thì sản phẩm nhận diện ngôn ngữ ký hiệu có tiềm năng to lớn trong việc cải thiện khả năng giao tiếp của người khiếm thính và mở ra những cơ hội mới trong các lĩnh vực như giáo dục, y tế, và xã hội. Sản phẩm đã đạt được những thành tựu đáng kể trong việc cải thiện khả năng nhận diện và hiểu ngôn ngữ ký hiệu, từ đó mang lại trải nghiệm người dùng tốt hơn và góp phần vào sự đa dạng và tích cực trong xã hội.
* Tuy nhiên, để làm cho sản phẩm nhận diện ngôn ngữ ký hiệu trở thành một phần không thể thiếu trong cuộc sống hàng ngày của người khiếm thính, bản thân em cần tiếp tục nỗ lực trong việc phát triển và cập nhật các tính năng, đồng thời tăng cường giáo dục và sự nhận thức trong cộng đồng. Đặc biệt, việc hợp tác với các chuyên gia, tổ chức, và cộng đồng người khiếm thính là quan trọng để đảm bảo rằng sản phẩm đáp ứng đúng nhu cầu và mang lại giá trị thực sự cho người sử dụng.
* Sự phát triển của sản phẩm nhận diện ngôn ngữ ký hiệu không chỉ là một bước quan trọng trong hành trình hỗ trợ người khiếm thính mà còn là một minh chứng cho sức mạnh của công nghệ trong việc tạo ra những thay đổi tích cực trong xã hội. Sản phẩm này có thể giúp tạo ra một môi trường bao dung, nơi mọi người có thể tương tác và giao tiếp một cách tự nhiên và hiệu quả.

# **PHẦN BA. TÀI LIỆU NGHIÊN CỨU**

<https://viblo.asia/p/tim-hieu-ve-yolo-trong-bai-toan-real-time-object-detection-yMnKMdvr57P>

<https://vinbigdata.com/cong-nghe-hinh-anh/yolov8-co-gi-nang-cap-so-voi-cac-phien-ban-truoc.html>

<https://pypi.org/project/ultralytics/>

<https://docs.ultralytics.com/>

<https://github.com/ultralytics>

<https://colab.research.google.com/github/ultralytics/ultralytics/blob/main/examples/tutorial.ipynb>

<https://docs.ultralytics.com/reference/engine/results/>

<https://docs.ultralytics.com/modes/predict/>

<https://universe.roboflow.com/>

<https://www.youtube.com/watch?v=wuZtUMEiKWY>

# **PHẦN BỐN. CAM KẾT VỀ BẢN QUYỀN SẢN PHẨM**

* Sản phẩm chưa từng được công bố hoặc tham gia trong bất kỳ cuộc thi nào.
* Sản phẩm đúng bản quyền của sinh viên dự thi, trường hợp có sử dụng mã nguồn mở phải tuân thủ các yêu cầu của giấy phép mã nguồn mở của các tổ chức, cá nhân phát hành mã nguồn mở.