

**TRƯỜNG ĐẠI HỌC XÂY DỰNG**

**KHOA CÔNG NGHỆ THÔNG TIN**

---------------🙦 🕮 🙤---------------



ĐỒ ÁN

**TỐT NGHIỆP ĐẠI HỌC**

NGÀNH CÔNG NGHỆ THÔNG TIN

**Computer Vision**

|  |  |
| --- | --- |
| Mã đồ án: | 1 |
| Sinh viên thực hiện: | **Nguyễn Thạc Trung Hiếu** |
| Mã sinh viên: | 0189866 |
| Lớp: | 66PM |
| Giảng viên hướng dẫn | TS. **Phạm Hữu Tùng** |

Hà Nội, 01/2026

**PHIẾU GIAO NHIỆM VỤ ĐỒ ÁN TỐT NGHIỆP**

**TÓM TẮT**

**NỘI DUNG ĐỒ ÁN TỐT NGHIỆP**

**MỤC LỤC**

**[Chương 1: Đặt vấn đề và định hướng giải pháp 1](#_heading=h.2q3ftj2iw4a0)**

**[1.1](#_heading=h.pbjtlfq9c7jj)****Lý do chọn đề tài** 1

**[1.1.1.](#_heading=h.awf6kzveq5mf)****Cơ sở thực tiễn** 1

**[1.1.2.](#_heading=h.dbm9ce9i77ku)****Đặt vấn đề** 1

**[1.2](#_heading=h.kr12hg6iqa0h)****Mục tiêu đồ án** 1

**[1.3](#_heading=h.3b22n4z3yvfe)****Phạm vi thực hiện** 2

**[1.4](#_heading=h.gg5bgpqx40hg)****Phương pháp thực hiện** 2

**[Chương 2: Các kết quả đạt được 3](#_heading=h.pttun6c484fj)**

**[2.1](#_heading=h.l28k8zfqs4zk)****Mô tả bài toán** 3

**[2.2](#_heading=h.l02qh2i0kxid)****Yêu cầu chức năng** 3

**[2.2.1.](#_heading=h.kkf8tqr14qdo)****Quản lý người dùng** 3

**[2.2.2.](#_heading=h.i7zga46sz7b)****Quản lý sân thể thao** 3

**[2.2.3.](#_heading=h.1u5t38xfq1sb)****Đặt sân** 3

**[2.2.4.](#_heading=h.uul1arsnclox)****Thanh toán** 3

**[2.2.5.](#_heading=h.ouwnxvemhryt)****Đánh giá** 3

**[2.2.6.](#_heading=h.29frxa8dazyu)****Báo cáo và thống kê** 4

**[2.3](#_heading=h.aygvjt23jts1)****Yêu cầu phi chức năng** 4

**[2.4](#_heading=h.jebuxxjtl19b)****Use Case Diagram tổng quát** 4

**[2.5](#_heading=h.g8odstljbxca)****Mô tả các Use Case chính** 5

**[Chương 3: Kết luận 1](#_heading=h.91eb4z447a5d)**

**[3.1 Lựa chọn mô hình kiến trúc](#_heading=h.2g39tuqiqgu2)** [1](#_heading=h.2g39tuqiqgu2)

**[3.1.1 Khái niệm về Microservices](#_heading=h.9eq6a8mcn63g)** [1](#_heading=h.9eq6a8mcn63g)

**[3.1.2 Khái niệm về Spring Cloud](#_heading=h.7kqgrc45hic1)** [1](#_heading=h.7kqgrc45hic1)

**[3.2 Sơ đồ kiến trúc tổng thể](#_heading=h.f9xylgssown1)** [1](#_heading=h.f9xylgssown1)

**MỤC LỤC HÌNH ẢNH**

*[Hình 1: Use case tổng quan hệ thống](#_heading=h.15eb3nv1k1og)* [4](#_heading=h.15eb3nv1k1og)

# **Chương 1: Đặt vấn đề và định hướng giải pháp**

* 1. **Lý do chọn đề tài**
     1. **Cơ sở thực tiễn**

# Trong những năm gần đây, công nghệ Computer Vision (thị giác máy tính) đã có sự phát triển vượt bậc và ngày càng được ứng dụng rộng rãi trong nhiều lĩnh vực tại Việt Nam cũng như trên thế giới. Theo báo cáo của Markets and Markets, thị trường Computer Vision toàn cầu đạt giá trị 11.32 tỷ USD năm 2021 và dự kiến tăng trưởng lên 17.4 tỷ USD vào năm 2026, với tốc độ tăng trưởng kép hàng năm (CAGR) là 7.3%. Tại Việt Nam, các lĩnh vực ứng dụng Computer Vision phổ biến nhất bao gồm bán lẻ và thương mại điện tử (35%), y tế (28%), an ninh và giám sát (42%), sản xuất công nghiệp (25%), và giao thông thông minh (18%). Công nghệ này đang được triển khai mạnh mẽ tại các thành phố lớn như Hà Nội và TP. Hồ Chí Minh, từ các hệ thống thanh toán không tiếp xúc, nhận diện khuôn mặt, đến kiểm soát chất lượng sản phẩm trong nhà máy.

Sự phát triển của Computer Vision được thúc đẩy bởi nhiều yếu tố thực tiễn quan trọng. Thứ nhất, khả năng tính toán đã tăng lên đáng kể nhờ sự phát triển của GPU (Graphics Processing Unit) và TPU (Tensor Processing Unit), cho phép xử lý hàng triệu phép tính song song trong thời gian thực. Các thiết bị edge computing như Raspberry Pi, Jetson Nano, và Intel Neural Compute Stick cũng ngày càng phổ biến với giá thành hợp lý, giúp triển khai Computer Vision dễ dàng hơn. Thứ hai, dữ liệu hình ảnh và video ngày càng dồi dào với hàng tỷ ảnh được tải lên internet mỗi ngày, cùng với các bộ dữ liệu mở như ImageNet (14 triệu ảnh), COCO (330,000 ảnh), và Open Images (9 triệu ảnh) cung cấp nguồn tài nguyên khổng lồ để huấn luyện mô hình. Thứ ba, các framework mã nguồn mở như TensorFlow, PyTorch, OpenCV, và các mô hình pre-trained như YOLO, ResNet, MobileNet đã làm giảm đáng kể rào cản kỹ thuật, cho phép các nhà phát triển không cần đào tạo mô hình từ đầu.

Tuy nhiên, việc triển khai Computer Vision trong thực tế vẫn còn nhiều thách thức cần giải quyết. Về mặt kỹ thuật, độ chính xác của mô hình phụ thuộc nhiều vào chất lượng và số lượng dữ liệu huấn luyện, đặc biệt trong các điều kiện ánh sáng khác nhau, góc chụp đa dạng, hoặc các yếu tố che khuất. Chi phí đầu tư ban đầu cho phần cứng chuyên dụng, camera chất lượng cao, và server xử lý vẫn còn cao đối với các doanh nghiệp vừa và nhỏ tại Việt Nam. Vấn đề về quyền riêng tư và bảo mật dữ liệu cũng là mối quan ngại lớn, đặc biệt khi ứng dụng liên quan đến nhận diện khuôn mặt hoặc giám sát. Hơn nữa, thiếu nguồn nhân lực có chuyên môn sâu về Computer Vision và AI khiến nhiều tổ chức gặp khó khăn trong việc phát triển và duy trì hệ thống. Trong bối cảnh chuyển đổi số đang diễn ra mạnh mẽ, việc xây dựng các giải pháp Computer Vision tối ưu, dễ triển khai và phù hợp với điều kiện thực tế là yêu cầu cấp thiết để đáp ứng nhu cầu ngày càng tăng của doanh nghiệp và người dùng.

* + 1. **Đặt vấn đề**

Trong những năm gần đây, đột quỵ đã trở thành một trong những nguyên nhân gây tử vong và tàn tật hàng đầu tại Việt Nam cũng như trên thế giới. Theo Tổ chức Y tế Thế giới (WHO), mỗi năm có khoảng 15 triệu người bị đột quỵ toàn cầu, trong đó 5 triệu người tử vong và 5 triệu người bị tàn tật vĩnh viễn. Tại Việt Nam, số ca đột quỵ ước tính khoảng 200,000 ca mỗi năm, với tỷ lệ tử vong lên đến 25-30% trong tháng đầu tiên. Thời gian vàng để cứu sống và giảm thiểu di chứng cho bệnh nhân đột quỵ chỉ trong vòng 3-4.5 giờ đầu kể từ khi xuất hiện triệu chứng. Một trong những dấu hiệu cảnh báo quan trọng của đột quỵ là ngã đột ngột, thường kèm theo liệt nửa người, méo mặt, mất khả năng cân bằng, và khó khăn trong việc đứng dậy. Tuy nhiên, việc phân biệt giữa cú ngã do đột quỵ và ngã bình thường (do trượt chân, vấp ngã, chóng mặt nhẹ) là vô cùng khó khăn, đặc biệt đối với người thân hoặc nhân viên y tế không có mặt ngay lập tức, dẫn đến việc chậm trễ trong cấp cứu và gia tăng nguy cơ tử vong hoặc di chứng nặng nề.

Trong bối cảnh công nghệ Computer Vision (thị giác máy tính) đang phát triển mạnh mẽ, việc ứng dụng vào lĩnh vực y tế, đặc biệt là phát hiện sớm đột quỵ, đang được quan tâm nghiên cứu và triển khai. Theo báo cáo của Markets and Markets, thị trường Computer Vision trong y tế đạt giá trị 1.8 tỷ USD năm 2021 và dự kiến tăng lên 4.5 tỷ USD vào năm 2026, với tốc độ tăng trưởng 20.1% mỗi năm. Các ứng dụng Computer Vision trong y tế đã chứng minh hiệu quả trong chẩn đoán hình ảnh X-quang (độ chính xác 92-95%), phát hiện khối u qua CT/MRI (89-94%), và phân tích võng mạc mắt (90-93%). Đặc biệt, công nghệ phát hiện té ngã và phân tích tư thế người đã được nghiên cứu rộng rãi với độ chính xác 85-92% trong môi trường kiểm soát. Tại các quốc gia phát triển như Mỹ, Nhật Bản, và châu Âu, hệ thống giám sát người cao tuổi sử dụng Computer Vision đã được triển khai tại các viện dưỡng lão, bệnh viện, và ngay cả tại gia đình, giúp phát hiện kịp thời các sự cố ngã và gọi cấp cứu tự động.

Sự phát triển của Computer Vision trong phát hiện đột quỵ được thúc đẩy bởi nhiều yếu tố thực tiễn quan trọng. Thứ nhất, khả năng tính toán thời gian thực đã tăng lên đáng kể nhờ sự phát triển của GPU (Graphics Processing Unit) và TPU (Tensor Processing Unit), cho phép xử lý video với tốc độ 30-60 khung hình/giây, đủ nhanh để phát hiện sự cố ngay lập tức. Các thiết bị edge computing như Jetson Nano, Raspberry Pi 4, và camera AI tích hợp sẵn khả năng xử lý cũng ngày càng phổ biến với giá thành từ 2-10 triệu đồng, phù hợp cho triển khai gia đình và bệnh viện. Thứ hai, dữ liệu y tế hình ảnh ngày càng phong phú với các bộ dữ liệu mở như UR Fall Detection Dataset, Multiple Cameras Fall Dataset, và các nghiên cứu lâm sàng về hành vi ngã của bệnh nhân đột quỵ cung cấp nguồn tài nguyên để huấn luyện mô hình chính xác. Thứ ba, các framework mã nguồn mở như MediaPipe (Google) cho phát hiện pose estimation, OpenPose cho phân tích khung xương người, cùng với TensorFlow, PyTorch, và các mô hình pre-trained như MoveNet, PoseNet đã làm giảm đáng kể rào cản kỹ thuật, cho phép các nhà phát triển xây dựng hệ thống phát hiện ngã với độ chính xác cao mà không cần đào tạo mô hình từ đầu.

Tuy nhiên, việc triển khai Computer Vision để phân biệt ngã do đột quỵ và ngã bình thường trong thực tế vẫn còn nhiều thách thức cần giải quyết. Về mặt kỹ thuật, độ chính xác của mô hình phụ thuộc nhiều vào chất lượng dữ liệu huấn luyện với các trường hợp đột quỵ thực tế, trong khi dữ liệu này rất khan hiếm và khó thu thập do tính nhạy cảm y tế. Các yếu tố môi trường như ánh sáng yếu, góc camera không thuận lợi, nhiều người trong khung hình, hoặc vật che khuất có thể làm giảm độ chính xác xuống 60-70%. Chi phí đầu tư ban đầu cho hệ thống camera giám sát chất lượng cao, server xử lý, và kết nối mạng ổn định vẫn còn là rào cản đối với nhiều gia đình và cơ sở y tế quy mô nhỏ tại Việt Nam. Vấn đề về quyền riêng tư là mối quan ngại lớn khi camera giám sát liên tục trong không gian sống cá nhân, đòi hỏi giải pháp xử lý dữ liệu cục bộ và mã hóa bảo mật cao. Hơn nữa, việc phân biệt chính xác giữa ngã do đột quỵ (với các dấu hiệu như liệt nửa người, tư thế bất thường, không tự đứng dậy được) và ngã bình thường (vấp ngã nhưng vẫn cử động bình thường) đòi hỏi mô hình AI phải học các đặc trưng y khoa phức tạp, cần sự hợp tác chặt chẽ giữa kỹ sư AI và bác sĩ thần kinh.

Thực trạng trên đặt ra yêu cầu xây dựng một hệ thống phát hiện và cảnh báo ngã thông minh dựa trên Computer Vision, đáp ứng nhu cầu chăm sóc sức khỏe cho người cao tuổi và bệnh nhân có nguy cơ đột quỵ cao tại Việt Nam. Hệ thống cần có khả năng phân tích tư thế người theo thời gian thực, phát hiện cú ngã đột ngột, phân biệt ngã do đột quỵ thông qua các dấu hiệu như liệt nửa người, tư thế bất thường, không tự cử động sau ngã, và tự động gửi cảnh báo đến người thân hoặc trung tâm cấp cứu. Hệ thống cũng cần đảm bảo bảo mật dữ liệu cá nhân, chi phí triển khai hợp lý, dễ dàng lắp đặt tại gia đình hoặc bệnh viện, và có giao diện thân thiện để người dùng không chuyên có thể sử dụng, phù hợp với đặc thù dân số già hóa và tỷ lệ đột quỵ cao tại Việt Nam.

* 1. **Mục tiêu đồ án**

Đồ án nhằm Xây dựng hệ thống phát hiện và phân loại tự động các cú ngã do đột quỵ và ngã bình thường sử dụng công nghệ Computer Vision và Deep Learning, nhằm hỗ trợ cảnh báo sớm và giảm thiểu thời gian cấp cứu cho bệnh nhân đột quỵ.

* 1. **Phạm vi thực hiện**
     + Phạm vi nghiên cứu:

· **Ngôn ngữ lập trình chính**: Python 3.8+ cho phát triển mô hình Computer Vision và xử lý video

· **Framework Deep Learning**: TensorFlow 2.x hoặc PyTorch 1.x

· **Thư viện Computer Vision**:

* OpenCV (cv2) cho xử lý video và hình ảnh
* MediaPipe hoặc OpenPose cho Pose Estimation

· **Mô hình AI sử dụng**:

* Pose Estimation: MediaPipe Pose, MoveNet Lightning/Thunder, hoặc PoseNet
* Phân loại: CNN, LSTM, hoặc mô hình kết hợp (CNN-LSTM, ConvLSTM)

· **API Backend**: Flask hoặc FastAPI để xây dựng RESTful API phục vụ mô hình AI

* 1. **Phương pháp thực hiện**
* Ngôn ngữ lập trình: Python, NodeJs

# **Chương 2: Các kết quả đạt được**

* 1. **Mô tả bài toán**

Bài toán đặt ra là xây dựng một hệ thống giám sát thông minh dựa trên công nghệ Computer Vision nhằm cung cấp một giải pháp tự động:

· Hệ thống cần có khả năng **tự động phát hiện** khi một người ngã xuống đất từ tư thế đứng hoặc ngồi thông qua phân tích video từ camera giám sát.

· Phát hiện phải diễn ra **theo thời gian thực** (real-time) với độ trễ tối thiểu (< 3 giây) để đảm bảo cảnh báo kịp thời.

· Giảm thiểu **báo động giả** (false positive) từ các hành động bình thường như ngồi xuống, nằm nghỉ, cúi xuống nhặt đồ.

* 1. **Yêu cầu chức năng**

### Phát hiện ngã tự động

* **Mô tả**: Hệ thống theo dõi video từ camera và tự động phát hiện khi có người ngã xuống
* **Input**: Video stream từ camera
* **Output**: Cảnh báo "Phát hiện ngã" kèm timestamp
* **Yêu cầu**:
  + Xử lý real-time (20-30 FPS)
  + Độ chính xác ≥ 85%
  + Phản hồi trong vòng 2-3 giây

### 1.2. Phân loại loại ngã

* **Mô tả**: Phân biệt ngã bình thường và ngã nghi ngờ đột quỵ
* **Input**: Chuỗi frame video của sự kiện ngã
* **Output**:
  + "Ngã bình thường" - Người có thể tự đứng dậy
  + "Ngã nghi ngờ đột quỵ" - Có dấu hiệu bất thường (liệt, không cử động)
* **Yêu cầu**:
  + Độ chính xác phân loại ≥ 80%
  + Phân tích dựa trên tư thế và chuyển động

### 1.3. Gửi cảnh báo tự động

* **Mô tả**: Tự động thông báo đến người thân/y tế khi phát hiện ngã nghi ngờ đột quỵ
* **Kênh thông báo**:
  + Email
  + SMS (tùy chọn)
  + Thông báo trên web (real-time)
* **Nội dung cảnh báo**:
  + Thời gian xảy ra
  + Loại ngã
  + Vị trí camera
  + Link xem video

## 2. **Chức năng giám sát và hiển thị**

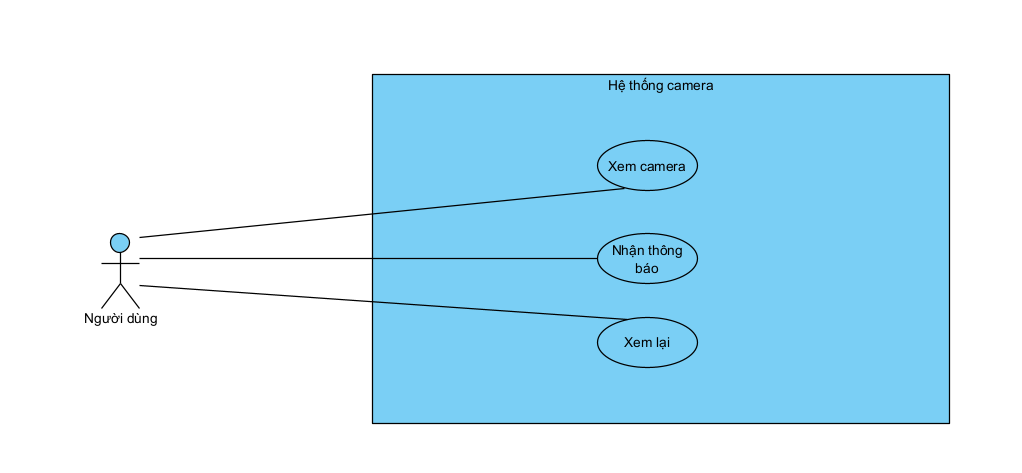
### 2.1. Xem video trực tiếp (Live Stream)

**Mô tả**: Hiển thị video real-time từ camera trên web

* **Tính năng**:
  + Stream video mượt mà
  + Hiển thị skeleton overlay (khung xương người) trên video
  + Hiển thị trạng thái: "Đang giám sát", "Phát hiện ngã", "Bình thường"
* **Giao diện**:
  + Ô video lớn ở giữa
  + Thông tin camera ở góc (tên, trạng thái)

### 2.2. Xem lại sự kiện ngã (Playback)

* **Mô tả**: Xem lại video của các sự kiện ngã đã được ghi lại
* **Tính năng**:
  + Danh sách các sự kiện (thời gian, loại ngã)
  + Click vào để xem video clip (10-20 giây)
  + Nút phát/tạm dừng cơ bản
  1. **Yêu cầu phi chức năng**
* Giao diện thân thiện, dễ sử dụng với người dùng.
* Ổn định, xử lý hiệu quả.
* Khả năng lưu trữ dữ liệu và phục hồi khi gặp sự cố.
* Yêu cầu phân quyền người dùng.
  1. **Use Case Diagram tổng quát**



*Hình 1: Use case tổng quan hệ thống*

# 