

# TRƯỜNG ĐẠI HỌC ĐẠI NAM

ĐỀ TÀI: Hệ thống nhận diện và đếm số lượng sinh viên trong lớp học

Học phần: TTCN5 - Triển khai, phát triển ứng dụng AI và IoT

Giảng viên hướng dẫn: Th.S Lê Trung Hiếu và Kỹ sư Nguyễn Thái Khánh

Nguyễn Hoàng Liêm  
Đại học Đại Nam  
MSV: 1671020178

Trần Hiếu Nghĩa  
Đại học Đại Nam  
MSV: 1671020223

Phạm Văn Duy  
Đại học Đại Nam  
MSV: 1671020061

Nguyễn Minh Hiếu  
Đại học Đại Nam  
MSV: 1671020108

**Tóm tắt nội dung**—Bài báo cáo này trình bày mô hình nhận diện số sinh viên trong lớp học nhóm chúng em phát triển để nhận dạng số người qua hình ảnh sử dụng YOLOv8. Cuối cùng, thu được kết quả với độ chính xác tương đối

## I. GIỚI THIỆU

Đề tài "Nhận diện số sinh viên trong lớp" nhằm hỗ trợ việc quản lý số lượng sinh viên có mặt và vắng mặt trong lớp học, giúp nâng cao hiệu quả giám sát, giảm thiểu các rủi ro và đảm bảo an toàn cho sinh viên trước những người lạ. Việc áp dụng công nghệ nhận diện khuôn mặt không chỉ giúp tự động hóa quá trình điểm danh mà còn góp phần tăng cường bảo mật trong môi trường học đường. Chúng em lựa chọn đề tài này với mong muốn nghiên cứu và tìm ra giải pháp tối ưu để cải thiện những vấn đề trên, đồng thời nâng cao kiến thức và kỹ năng trong lĩnh vực công nghệ trí tuệ nhân tạo (AI) và thị giác máy tính.

## II. ỨNG DỤNG AI VÀ IoT

### A. MỤC TIÊU

Vận dụng kiến thức lập trình, thiết kế giải pháp IoT cho hệ thống giữa thiết bị và mạng internet. Vận dụng kiến thức lập trình, AI trong triển khai phát triển phần mềm. Thiết kế lập trình mạch điều khiển IoT với các thiết bị cảm biến giao tiếp dữ liệu thông qua Back-End. Thiết kế hệ thống phần mềm Back-End, Front-End giao tiếp với thiết bị IoT kết hợp với AI cho giải pháp thông minh

### B. Ứng dụng AI và IoT

#### 1. AI (Trí tuệ nhân tạo) là gì?

Trí tuệ nhân tạo (AI - Artificial Intelligence) là công nghệ giúp máy móc có khả năng học hỏi, suy nghĩ, phân tích và ra quyết định tương tự con người. AI bao gồm nhiều lĩnh vực như:

- Machine Learning (ML):** Máy học từ dữ liệu để đưa ra dự đoán hoặc tự động hóa tác vụ.
- Computer Vision:** Nhận diện hình ảnh, khuôn mặt, vật thể trong ảnh và video.
- Xử lý ngôn ngữ tự nhiên (NLP - Natural Language Processing):** Giúp máy hiểu và phản hồi ngôn ngữ con người (ví dụ: ChatGPT, Google Assistant).
- Hệ thống chuyên gia (Expert Systems):** Hỗ trợ ra quyết định dựa trên kiến thức chuyên môn.

#### 2. IoT (Internet of Things) là gì?

Internet vạn vật (IoT - Internet of Things) là mạng lưới các thiết bị thông minh có thể kết nối internet, thu thập và trao đổi dữ liệu. Các thiết bị IoT có thể hoạt động tự động và được điều khiển từ xa.

Ví dụ về ứng dụng IoT:

- Nhà thông minh (Smart Home):** Điều khiển đèn, điều hòa, camera giám sát qua điện thoại.
- Giao thông thông minh:** Cảm biến giám sát lưu lượng xe, hệ thống đèn giao thông tự động điều chỉnh.
- Nông nghiệp thông minh:** Cảm biến đo độ ẩm đất, nhiệt độ môi trường để tưới tiêu tự động.
- Chăm sóc sức khỏe:** Đồng hồ thông minh theo dõi nhịp tim, huyết áp, phát hiện dấu hiệu bệnh lý.

#### 3. Sự kết hợp giữa AI và IoT (AIoT - Artificial Intelligence of Things)

Khi AI kết hợp với IoT, hệ thống không chỉ thu thập dữ liệu mà còn tự động phân tích và đưa ra quyết định thông minh hơn. AIoT giúp các thiết bị IoT học hỏi từ dữ liệu và hoạt động hiệu quả hơn.

#### Ứng dụng AIoT thực tế:

- Camera an ninh thông minh:** Nhận diện khuôn mặt, phát hiện người lạ và gửi cảnh báo.
- Hệ thống giám sát môi trường:** Dự đoán ô nhiễm không khí, phát hiện khí độc hại.
- Xe tự hành:** Kết hợp cảm biến IoT và AI để phát hiện vật cản, nhận diện biển báo giao thông.
- Robot tự động trong nhà máy:** Kiểm tra sản phẩm lỗi, tối ưu hóa dây chuyền sản xuất.

#### 4. Lợi ích của AI và IoT

- Tự động hóa:** Giảm sức lao động của con người, tăng hiệu suất làm việc.
- Tiết kiệm chi phí:** Tối ưu hóa quy trình sản xuất, giảm thiểu lãng phí.
- Cải thiện chất lượng cuộc sống:** Ứng dụng trong y tế, nhà thông minh giúp nâng cao tiện nghi.
- Tăng cường bảo mật:** Camera AI giúp nhận diện khuôn mặt, phát hiện kẻ xâm nhập.

#### 5. Kết luận

AI và IoT là hai công nghệ quan trọng đang thay đổi cách con người làm việc và sinh hoạt. Khi kết hợp lại (AIoT), chúng tạo

ra những hệ thống thông minh có thể học hỏi và thích nghi với môi trường, mang lại hiệu quả cao hơn trong nhiều lĩnh vực như công nghiệp, y tế, giao thông và nhà thông minh.

### III. GIỚI THIỆU CHUNG VỀ YOLOV8

Nhận diện đối tượng là một trong những lĩnh vực quan trọng nhất của thị giác máy tính, với vô số ứng dụng trong đời sống hàng ngày như giám sát an ninh, xe tự hành, y tế, và nhiều lĩnh vực khác. Trong số các mô hình hiện đại, dòng YOLO (*You Only Look Once*) đã chứng minh được hiệu quả vượt trội nhờ vào khả năng nhận diện nhanh chóng và chính xác trong thời gian thực.

Sau nhiều phiên bản cải tiến, mới đây, YOLOv8 đã được ra mắt với nhiều nâng cấp đáng kể cả về hiệu suất lẫn độ chính xác. YOLOv8 kế thừa những điểm mạnh của các phiên bản trước, đồng thời tinh chỉnh kiến trúc để trở nên nhẹ hơn, nhanh hơn và phù hợp hơn với nhiều ứng dụng thực tế. Sự ra đời của YOLOv8 đánh dấu một bước tiến quan trọng trong lĩnh vực nhận diện đối tượng bằng trí tuệ nhân tạo.

### IV. SỰ TIẾN HÓA CỦA YOLO QUA CÁC PHIÊN BẢN

YOLO lần đầu tiên được giới thiệu vào năm 2016 bởi Joseph Redmon. Kể từ đó, mô hình này đã trải qua nhiều cải tiến, với mỗi phiên bản mới đều mang lại những đột phá về hiệu suất.

- **YOLOv1 - YOLOv3 (2016-2018):** Các phiên bản đầu tiên tập trung vào tốc độ xử lý nhanh và khả năng nhận diện chính xác trong thời gian thực. YOLOv3 cải tiến thuật toán để đạt độ chính xác cao hơn.
- **YOLOv4 (2020):** Cải thiện đáng kể hiệu suất bằng cách sử dụng CSPDarknet53 làm *backbone* và tối ưu các chiến lược huấn luyện.
- **YOLOv5 (2020):** Phiên bản được Ultralytics phát triển với trọng tâm là tính dễ sử dụng, hỗ trợ PyTorch mạnh mẽ và khả năng triển khai linh hoạt.
- **YOLOv6, YOLOv7 (2022):** Các phiên bản này tiếp tục tối ưu tốc độ và độ chính xác, với nhiều cải tiến về kiến trúc. YOLOv7 đặc biệt mạnh mẽ trong việc cân bằng giữa tốc độ xử lý và độ chính xác.
- **YOLOv8 (2023):** Phiên bản mới nhất với những cải tiến vượt bậc về tốc độ, độ chính xác và tính linh hoạt trong nhiều tác vụ khác nhau như phát hiện đối tượng, phân đoạn hình ảnh, theo dõi đối tượng...

Vậy điều gì làm cho YOLOv8 trở thành một lựa chọn tối ưu trong năm 2023 và xa hơn? Hãy cùng tìm hiểu sâu hơn về những nâng cấp đáng chú ý của mô hình này.

### V. NHỮNG CẢI TIẾN QUAN TRỌNG CỦA YOLOV8

#### A. Kiến Trúc Mạng Neuron Tối Ưu Hóa

Một trong những nâng cấp quan trọng nhất của YOLOv8 chính là kiến trúc mạng neuron mới, giúp cải thiện đáng kể tốc độ và độ chính xác. YOLOv8 sử dụng một mô hình CNN (*Convolutional Neural Network*) được tinh chỉnh để phù hợp hơn với bài toán nhận diện đối tượng.

- **Backbone mạnh mẽ:** YOLOv8 sử dụng một *backbone* tiên tiến giúp trích xuất đặc trưng hình ảnh hiệu quả hơn,

giảm tải tính toán nhưng vẫn giữ được chất lượng dự đoán cao.

- **Neck tối ưu:** Mô hình kết hợp *Feature Pyramid Network (FPN)* và *Path Aggregation Network (PAN)* để xử lý thông tin từ nhiều cấp độ đặc trưng khác nhau.
- **Head cải tiến:** YOLOv8 sử dụng chiến lược *anchor-free*, giúp tối ưu hóa việc phát hiện đối tượng mà không cần dựa vào các hộp giới hạn (*anchor boxes*), giảm thiểu sai số trong việc xác định kích thước của đối tượng.

Những thay đổi này giúp YOLOv8 không chỉ nhanh hơn mà còn có thể nhận diện đối tượng tốt hơn, ngay cả trong các tình huống phức tạp như ánh sáng yếu, vật thể bị che khuất một phần hay nhiều đối tượng xuất hiện đồng thời trong khung hình.

#### B. Độ Chính Xác Cao Hơn

Với mỗi phiên bản mới, mục tiêu chính của YOLO là cải thiện *mean Average Precision (mAP)* – chỉ số đánh giá độ chính xác trong việc phát hiện đối tượng. YOLOv8 có độ chính xác *mAP@50* cao hơn YOLOv7, trong khi vẫn giữ được tốc độ xử lý tối ưu.

Phiên bản	FPS (Tốc độ)	mAP@50	Kích thước mô hình
YOLOv5	Cao	48-50%	Nhỏ
YOLOv6	Cao hơn	50-52%	Trung bình
YOLOv7	Rất cao	53-55%	Lớn hơn
YOLOv8	Cao nhất	55-57%	Tối ưu nhất

Bảng I

SO SÁNH HIỆU SUẤT GIỮA CÁC PHIÊN BẢN YOLO GẦN ĐÂY.

#### C. Hỗ Trợ Nhiều Tác Vụ Khác Nhau

YOLOv8 không chỉ giới hạn trong nhận diện đối tượng (*Object Detection*), mà còn có thể áp dụng cho nhiều tác vụ khác như:

- **Instance Segmentation:** Phân vùng đối tượng trong hình ảnh với độ chính xác cao.
- **Object Tracking:** Theo dõi đối tượng trong video theo thời gian thực.
- **Pose Estimation:** Dự đoán tư thế của con người, ứng dụng trong thể thao, y tế...
- **Edge AI:** Có thể triển khai trên thiết bị nhúng với tài nguyên hạn chế như Raspberry Pi, Jetson Nano...

Điều này giúp YOLOv8 trở thành một mô hình linh hoạt, phù hợp với nhiều bài toán thực tế.

### VI. ỨNG DỤNG CỦA YOLOV8

Nhờ vào những cải tiến mạnh mẽ, YOLOv8 có thể được sử dụng trong nhiều lĩnh vực khác nhau:

- **Giám sát an ninh:** Nhận diện khuôn mặt, phát hiện người xâm nhập trái phép.
- **Giao thông thông minh:** Đếm số lượng xe cộ, nhận diện biển báo giao thông.
- **Y tế:** Phát hiện bất thường trong hình ảnh y khoa, hỗ trợ chẩn đoán bệnh.
- **Tự động hóa công nghiệp:** Kiểm tra chất lượng sản phẩm trong nhà máy.

**Ưu điểm của YOLO so với các mô hình khác:**

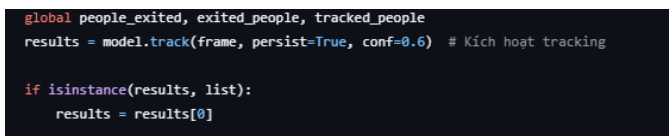
- **Tốc độ cao:** YOLO có thể nhận diện và phân loại nhiều đối tượng trong ảnh chỉ với một lần xử lý, phù hợp với các ứng dụng thời gian thực như xe tự lái, giám sát an ninh, và nhận diện động vật.
- **Độ chính xác cao:** Với kiến trúc CNN tối ưu, YOLO có khả năng phát hiện các đặc trưng quan trọng trong ảnh một cách chính xác, ngay cả khi đối tượng bị che khuất một phần hoặc xuất hiện trong môi trường phức tạp.
- **Khả năng nhận diện nhiều đối tượng cùng lúc:** YOLO chia hình ảnh thành các lưới nhỏ và dự đoán nhiều hộp giới hạn (bounding boxes) cùng lúc, giúp nó có thể nhận diện nhiều loài động vật hoặc vật thể khác nhau trong cùng một ảnh.

## VII. BÀI THỰC HÀNH



Hình 1. Bounding box chỉ lấy phần đầu của người

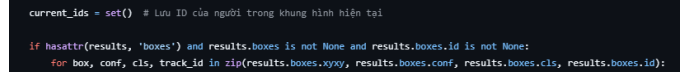
Hàm này giúp cắt một phần bounding box chỉ lấy phần đầu của người trong ảnh/video. Điều này hữu ích trong các bài toán như nhận diện khuôn mặt, đoán biểu cảm hoặc theo dõi chỉ phần đầu của đối tượng. Đặc biệt, phương pháp này có thể giảm tải tính toán cho hệ thống khi chỉ cần tập trung vào khu vực nhỏ hơn trong ảnh.



Hình 2. Nhận diện và theo dõi người trong khung hình

- **Sử dụng mô hình YOLO:** Hệ thống áp dụng YOLO để theo dõi các đối tượng trong khung hình một cách hiệu quả và chính xác.
- **conf=0.6:** Chỉ nhận diện đối tượng nếu độ tin cậy (confidence) lớn hơn 60%. Điều này đảm bảo loại bỏ các dự đoán có độ tin cậy thấp, giảm thiểu lỗi nhận diện.

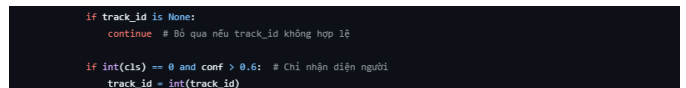
- **persist=True:** Giữ lại thông tin ID của các đối tượng qua các khung hình, giúp theo dõi liên tục và không bị gián đoạn.
- **if isinstance(results, list):** Đảm bảo results không phải danh sách rỗng để tránh lỗi trong quá trình xử lý.



Hình 3. Duyệt qua tất cả các đối tượng được phát hiện

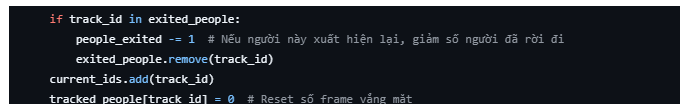
Duyệt qua từng đối tượng bằng:

- **box:** Tọa độ bounding box  $(x_1, y_1, x_2, y_2)$ , xác định vị trí và kích thước của đối tượng trong ảnh.
- **conf:** Độ tin cậy của dự đoán, phản ánh mức độ chính xác mà mô hình đưa ra.
- **cls:** Loại đối tượng, với 0 là người (hoặc các lớp khác tùy thuộc vào cấu hình huấn luyện).



Hình 4. Xử lý nếu đối tượng là người

- Nếu **track\_id** không tồn tại, bỏ qua đối tượng. Điều này ngăn chặn việc xử lý các đối tượng không được gán ID, giảm thiểu lỗi trong quá trình theo dõi.
- Xử lý tập trung vào các đối tượng đã được xác định rõ ràng, đảm bảo dữ liệu đầu ra có tính nhất quán và chính xác.



Hình 5. Cập nhật danh sách người theo dõi

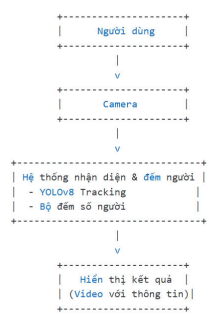
- Nếu một người từng rời đi xuất hiện lại, giảm biến **people\_exited**. Điều này giúp duy trì dữ liệu chính xác về số lượng người đã rời khỏi khung hình.
- Thêm **track\_id** vào **current\_ids** để đánh dấu rằng đối tượng này đang có mặt trong khung hình.
- Đặt lại số frame vắng mặt của đối tượng về 0 trong **tracked\_people**, đảm bảo hệ thống không nhầm lẫn với các đối tượng khác hoặc dữ liệu cũ.

## VIII. KẾT QUẢ

### A. Thiết lập thử nghiệm

Thí nghiệm được thực hiện bằng cách sử dụng camera real-time để nhận diện sinh viên trong lớp học với các thông số cụ thể như sau:

- **Mô hình:** YOLOv8n (phiên bản nhẹ của YOLOv8, được tối ưu cho video real-time). Mô hình này có khả năng xử lý nhanh nhưng vẫn đảm bảo độ chính xác cao.
- **Nguồn dữ liệu:** Video trực tiếp từ camera real-time, mô phỏng tình huống thực tế trong lớp học với nhiều sinh viên và điều kiện ánh sáng đa dạng.
- **Môi trường thử nghiệm:** Hệ thống được chạy trên GPU tầm trung, đảm bảo điều kiện gần giống với các ứng dụng thực tế.

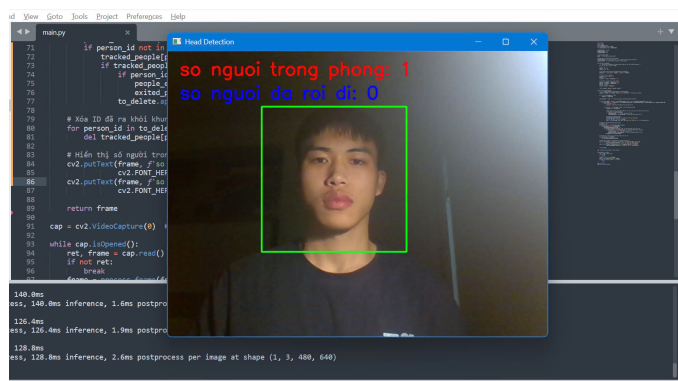


Hình 6. Sơ đồ dữ liệu

Hình 6 mô tả sơ đồ luồng dữ liệu trong hệ thống, bao gồm việc thu thập dữ liệu từ camera, xử lý qua mô hình YOLOv8, và hiển thị kết quả nhận diện theo thời gian thực.

### B. Kết quả thực nghiệm

Kết quả thực nghiệm được ghi nhận qua quá trình xử lý video từ camera real-time. Hình dưới đây minh họa kết quả nhận diện sinh viên trong lớp học:



Hình 7. Kết quả nhận diện

Kết quả thực nghiệm cho thấy mô hình YOLOv8 đạt độ chính xác cao, đặc biệt khi kết hợp các kỹ thuật tiền xử lý ảnh và điều chỉnh ngưỡng tự tin hợp lý. Một số đặc điểm nổi bật của kết quả bao gồm:

- **Độ chính xác:** Mô hình nhận diện chính xác các đối tượng trong hầu hết các tình huống, bao gồm khi đối tượng bị che khuất một phần hoặc trong điều kiện ánh sáng không đồng đều.

- **Tốc độ xử lý:** Hệ thống đạt tốc độ trung bình 23 FPS (khung hình mỗi giây), đảm bảo khả năng nhận diện theo thời gian thực và phù hợp với các ứng dụng giám sát lớp học.
- **Khả năng mở rộng:** Hệ thống có thể dễ dàng tích hợp thêm các tính năng khác như đếm số lượng sinh viên, nhận diện biểu cảm hoặc phân loại hành vi trong tương lai.

Những kết quả này chứng minh rằng YOLOv8 là một lựa chọn hiệu quả trong việc nhận diện đối tượng trong môi trường thực tế, đặc biệt là trong các ứng dụng liên quan đến giáo dục và giám sát.

## IX. KẾT LUẬN

Dự án đã triển khai thành công hệ thống nhận diện sinh viên trong lớp học dựa trên mô hình YOLOv8. Hệ thống không chỉ đạt độ chính xác cao mà còn đảm bảo tốc độ xử lý nhanh với trung bình 23 FPS, giúp theo dõi và điểm danh sinh viên một cách tự động và hiệu quả. Đây là bước tiến quan trọng trong việc ứng dụng trí tuệ nhân tạo vào lĩnh vực giáo dục và quản lý.

Một số đặc điểm nổi bật của dự án:

- **Hiệu quả thực tế:** Hệ thống hoạt động ổn định trong các tình huống thực tế, nhận diện tốt ngay cả khi có sự che khuất hoặc ánh sáng không đồng đều.
- **Tính tiện lợi:** Hệ thống giúp giảm bớt công việc thủ công trong việc điểm danh, đồng thời tăng tính minh bạch và chính xác.
- **Khả năng mở rộng:** Dự án đã xây dựng nền tảng vững chắc để phát triển thêm các tính năng nâng cao trong tương lai.

### Hướng phát triển trong tương lai

Để tối ưu hóa và mở rộng ứng dụng của hệ thống, một số hướng phát triển khả thi bao gồm:

- **Theo dõi danh tính sinh viên:** Tích hợp công nghệ nhận diện khuôn mặt để xác định danh tính từng sinh viên, phục vụ mục đích điểm danh chính xác hơn.
- **Lưu trữ dữ liệu lịch sử:** Xây dựng cơ sở dữ liệu lưu trữ thông tin về lịch sử điểm danh, cho phép giáo viên và nhà quản lý theo dõi thông tin lâu dài.
- **Tích hợp hệ thống quản lý trường học:** Kết nối hệ thống với các nền tảng quản lý trường học (School Management Systems) để tự động hóa quy trình báo cáo và quản lý.
- **Cải thiện độ chính xác:** Nâng cấp mô hình với các phương pháp huấn luyện tiên tiến hơn, kết hợp dữ liệu đa dạng để tăng độ chính xác trong nhiều môi trường khác nhau.
- **Phát hiện hành vi:** Bổ sung khả năng nhận diện hành vi bất thường trong lớp học, giúp cải thiện môi trường học tập và an ninh.

Tóm lại, dự án không chỉ chứng minh hiệu quả của mô hình YOLOv8 trong bài toán nhận diện sinh viên mà còn mở ra nhiều tiềm năng phát triển trong các ứng dụng giáo

dục hiện đại. Việc tiếp tục cải tiến và mở rộng hệ thống sẽ mang lại giá trị thực tiễn cao, đóng góp tích cực vào sự đổi mới trong ngành giáo dục.

## X. TÀI LIỆU THAM KHẢO

### TÀI LIỆU

- [1] J. Redmon et al., “You Only Look Once: Unified, Real-Time Object Detection,” *CVPR*, 2016.
- [2] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” *arXiv preprint arXiv:2004.10934*, 2020.
- [3] Ultralytics, “YOLOv5 Documentation,” [Online]. Available: <https://github.com/ultralytics/yolov5>. [Accessed: Mar. 2025].
- [4] Ultralytics, “YOLOv8: Next-Generation Object Detection and Segmentation,” [Online]. Available: <https://github.com/ultralytics/ultralytics>. [Accessed: Mar. 2025].
- [5] N. Wojke, A. Bewley, and D. Paulus, “Simple Online and Realtime Tracker with a Deep Association Metric,” *ICIP*, 2017.
- [6] T. Lin et al., “Focal Loss for Dense Object Detection,” *ICCV*, 2017.
- [7] T. Lin et al., “Microsoft COCO: Common Objects in Context,” *ECCV*, 2014.
- [8] G. Bradski, “The OpenCV Library,” *Dr. Dobb’s Journal of Software Tools*, 2000.
- [9] X. Zhang et al., “Edge AI: On-Demand Deep Learning Model Inference at the Edge,” *IEEE Network*, 2020.