

# ĐỀ CƯƠNG CHI TIẾT

Ngày thực hiện đề cương: 27/03/2025

## Thông tin đề tài

Tên đề tài	Nhận diện biểu cảm khuôn mặt trong điều kiện ánh sáng yếu sử dụng CNN nhẹ kết hợp kỹ thuật tăng cường dữ liệu thích ứng
Lĩnh vực chuyên ngành	Trí tuệ nhân tạo, Xử lý ảnh, Học sâu (Deep Learning).
Loại hình nghiên cứu	Nghiên cứu ứng dụng (Applied Research), Nghiên cứu phát triển (Developmental Research).
Người hướng dẫn	Th.s Đỗ Như Tài
Sinh viên thực hiện 1	Nguyễn Hữu Lộc (3122410213)
Sinh viên thực hiện 2	Mai Phúc Lâm (3122410207)
Sinh viên thực hiện 3	Văn Tuấn Kiệt (3122410202)
Sinh viên thực hiện 4	Nguyễn Đức Duy Lâm (3122410208)

## GIỚI THIỆU

Câu hỏi gợi ý	Lĩnh vực và nghiên cứu liên quan đã và đang phát triển như thế nào? Các vấn đề, và bài toán đặt ra cần giải quyết là gì?
Hướng dẫn	Giới thiệu tổng quan về đề tài – những vấn đề và lĩnh vực liên quan đến đề tài.

Ngày nay, công nghệ trí tuệ nhân tạo (AI) đã trở thành một yếu tố quan trọng trong việc thay đổi cách con người tiếp cận và giải quyết các vấn đề thực tiễn. Đặc biệt, nhận diện biểu cảm khuôn mặt (Facial Expression Recognition - FER) là một lĩnh vực có tiềm năng ứng dụng lớn trong giao tiếp người-máy, giám sát an ninh, và phân tích cảm xúc. Tuy nhiên, hiệu suất của các mô hình học sâu như CNN thường giảm đáng kể trong điều kiện ánh sáng yếu - một thách thức phổ biến trong các ứng dụng thực tế như camera giám sát ban đêm hoặc thiết bị nhúng.

Các nghiên cứu gần đây (2020-2025) đã đề xuất nhiều phương pháp để cải thiện FER trong điều kiện ánh sáng yếu, chẳng hạn như sử dụng GAN hoặc Retinex, nhưng các giải pháp này thường phức tạp và không khả thi với nguồn lực hạn chế. Do đó, bài toán đặt ra là liệu có thể xây dựng một hệ thống nhận diện biểu cảm đơn giản, hiệu quả, sử dụng CNN nhẹ (MobileNetV3) kết hợp với kỹ thuật tăng cường dữ liệu thích ứng hay không?

# TÌNH HÌNH NGHIÊN CỨU TRONG VÀ NGOÀI NƯỚC

Câu hỏi gợi ý	Lĩnh vực và nghiên cứu liên quan đã và đang phát triển như thế nào? Các vấn đề, và bài toán đặt ra cần giải quyết là gì?
Hướng dẫn	Tìm hiểu các nghiên cứu đã công bố gần nhất (5 năm trở lại) về lĩnh vực liên quan trong và ngoài nước. Phân tích các kết quả đóng góp, nhận xét các hạn chế còn tồn tại.

## Nghiên cứu ngoài nước

Trong những năm gần đây, nhận diện cảm xúc khuôn mặt (FER) nổi bật trong AI và thị giác máy tính nhờ học sâu. Các mô hình CNN như VGGNet (73,28% - Khaireddin et al.), ResNet, và ensemble (75,8% - Khanzada et al.) vượt kỷ lục trên FER2013. Ứng dụng mở rộng trong y tế, giáo dục, an ninh, và HCI, ví dụ web app thời gian thực (69,8%, 40ms). Nghiên cứu xử lý dữ liệu phức tạp qua dữ liệu phụ (CK+, JAFFE), tăng cường dữ liệu, và đa phương thức, đạt 99,26% trên CK+ (AA-DCN, 2024). Thách thức gồm thiếu dữ liệu đa dạng và vấn đề đạo đức, hướng tới tích hợp mốc khuôn mặt, CNN chú ý, và dữ liệu lớn như AffectNet.

## Nghiên cứu trong nước (Việt Nam)

Tại Việt Nam, nghiên cứu về nhận diện cảm xúc khuôn mặt tuy còn ở giai đoạn phát triển nhưng đã ghi nhận những kết quả đáng chú ý, tập trung vào ứng dụng thực tiễn: Công nghệ FER được áp dụng trong giám sát giao thông (phát hiện tài xế mệt mỏi), giáo dục (phân tích phản ứng học sinh qua video) và dịch vụ khách hàng (đánh giá mức độ hài lòng). Cũng đạt được những kết quả ấn tượng

## Phân tích kết quả đóng góp

### Đóng góp chung

- Phát triển công nghệ học sâu, nhận diện cảm xúc chính xác hơn.
- Ứng dụng đa dạng: y tế, giáo dục, an ninh, tiếp thị.
- Cung cấp tập dữ liệu chuẩn và hiểu biết về hành vi con người.

### Hạn chế chung

- Hiệu suất giảm trong điều kiện ánh sáng yếu do mất mát chi tiết và độ sáng thấp.
- Các yếu tố gây nhiễu trong môi trường không kiểm soát (che khuất, tư thế, v.v.).
- Hiệu suất tính toán hạn chế của các mô hình phức tạp, không phù hợp với thiết bị tài nguyên thấp.
- Thiếu dữ liệu ánh sáng yếu thực tế, dẫn đến việc phải sử dụng dữ liệu tổng hợp.

- Sự mất cân bằng dữ liệu và nhầm lẫn giữa các cảm xúc tương đồng.

## Kết luận

Nghiên cứu về nhận diện cảm xúc khuôn mặt trong và ngoài nước đã đạt được nhiều thành tựu quan trọng, từ cải tiến thuật toán đến ứng dụng thực tiễn. Tuy nhiên, để phát triển bền vững, cần giải quyết các vấn đề về quyền riêng tư, thiên vị thuật toán và dữ liệu đa dạng, đặc biệt tại Việt Nam, nơi tiềm năng nghiên cứu còn lớn nhưng nguồn lực còn hạn chế.

## MỤC TIÊU VÀ PHẠM VI NGHIÊN CỨU

<b>Câu hỏi gợi ý</b>	Mục tiêu nghiên cứu chính của đề tài là gì? Phạm vi nghiên cứu là gì?
<b>Hướng dẫn</b>	Đặt bài toán giải quyết và trình bày mục tiêu nghiên cứu chính của đề tài. Nêu phạm vi nghiên cứu của đề tài, bao gồm việc giới hạn phạm vi nghiên cứu và triển khai, các giả định ban đầu đối với nghiên cứu.

**Mục tiêu nghiên cứu:** Đề tài hướng đến giải quyết bài toán nhận diện biểu cảm khuôn mặt trong điều kiện ánh sáng yếu, một thách thức thực tiễn trong các ứng dụng như giám sát an ninh và giao tiếp người-máy. Các mục tiêu chính bao gồm:

1. Xây dựng mô hình CNN nhẹ (MobileNetV3) để nhận diện biểu cảm khuôn mặt trong điều kiện ánh sáng yếu.
2. Đề xuất và triển khai kỹ thuật tăng cường dữ liệu thích ứng dựa trên mức độ sáng của từng ảnh nhằm nâng cao hiệu suất mô hình.
3. Đánh giá hiệu quả của phương pháp đề xuất so với các phương pháp cơ bản không sử dụng tăng cường dữ liệu thích ứng.

**Phạm vi nghiên cứu:** Đề tài được triển khai và thử nghiệm trên tập dữ liệu FER-2013, với trọng tâm là xử lý điều kiện ánh sáng yếu. Nghiên cứu giới hạn trong việc sử dụng CNN nhẹ (MobileNetV3) và kỹ thuật tăng cường dữ liệu thích ứng, không mở rộng sang các mô hình phức tạp như GAN hay Retinex. Các giả định ban đầu bao gồm:

- Tập dữ liệu FER-2013 sau khi được biến đổi để mô phỏng ánh sáng yếu vẫn đủ đại diện cho bài toán thực tế.
- Kỹ thuật tăng cường dữ liệu thích ứng có thể cải thiện hiệu suất của mô hình CNN nhẹ so với phương pháp cố định.

# CÁC GIẢ THUYẾT VÀ CÁCH TIẾP CẬN

Câu hỏi gợi ý	Các giả thuyết đặt ra để giải quyết bài toán chính? Các cách tiếp cận để giải quyết bài toán đặt ra?
Hướng dẫn	Đặt ra những giả thuyết, hay vấn đề bài toán con cần phải giải quyết để đạt được mục tiêu nghiên cứu đề tài. Các cách tiếp cận (dự kiến) để giải quyết các giả thuyết, bài toán con đã đặt.

**Giả thuyết:** Để đạt được mục tiêu nghiên cứu nhận diện biểu cảm khuôn mặt trong điều kiện ánh sáng yếu, đề tài đặt ra các giả thuyết sau:

1. Có thể xây dựng một hệ thống nhận diện biểu cảm hiệu quả bằng cách sử dụng CNN nhẹ (MobileNetV3), đảm bảo tốc độ và hiệu suất trên thiết bị hạn chế tài nguyên.
2. Kỹ thuật tăng cường dữ liệu thích ứng, dựa trên phân tích mức độ sáng của từng ảnh, sẽ cải thiện độ chính xác của mô hình so với các phương pháp tăng cường dữ liệu cố định.
3. Hệ thống đạt được độ chính xác trên 70% và tốc độ suy luận dưới 0.1 giây/ảnh trên CPU khi kết hợp CNN nhẹ và tăng cường dữ liệu thích ứng.

**Bài toán con:**

1. Xây dựng kỹ thuật tăng cường dữ liệu thích ứng dựa trên mức độ sáng của từng ảnh để cải thiện chất lượng đầu vào cho mô hình.
2. Huấn luyện và tối ưu hóa mô hình MobileNetV3 để nhận diện biểu cảm trong điều kiện ánh sáng yếu.

**Cách tiếp cận:**

1. **Tiền xử lý dữ liệu:** Sử dụng tập dữ liệu FER-2013, giảm độ sáng ngẫu nhiên trong khoảng 20-80% để mô phỏng điều kiện ánh sáng yếu. Áp dụng kỹ thuật tăng cường dữ liệu thích ứng (ví dụ: gamma correction, contrast stretching) dựa trên histogram hoặc độ sáng trung bình của từng ảnh.
2. **Huấn luyện mô hình:** Fine-tuning mô hình MobileNetV3 trên tập dữ liệu đã được tăng cường, tối ưu hóa để đạt hiệu suất cao trong điều kiện ánh sáng yếu.
3. **Đánh giá:** So sánh hiệu suất (độ chính xác và tốc độ suy luận) giữa mô hình sử dụng tăng cường dữ liệu thích ứng và mô hình cơ bản không áp dụng kỹ thuật này.
4. **Công cụ:** Sử dụng Python với TensorFlow/Keras để huấn luyện mô hình và OpenCV để xử lý ảnh.

## TÀI LIỆU THAM KHẢO

<b>Câu hỏi gợi ý</b>	Các tài liệu tham khảo thuộc lĩnh vực nghiên cứu liên quan là gì?
<b>Hướng dẫn</b>	Liệt kê các tài liệu tham khảo liên quan đến lĩnh vực nghiên cứu của đề tài.

1. Amit Khanzada, Charles Bai, Ferhat Turker Celepcikay, (2020). Facial Expression Recognition with Deep Learning: Improving on the State of the Art and Applying to the Real World.
2. Shan Li and Weihong Deng, (2018). Deep Facial Expression Recognition: A Survey. *IEEE Transactions*.
3. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.
4. Howard, A. G., et al. (2019). MobileNetV3: Searching for Efficient Neural Network Architectures. *arXiv preprint arXiv:1905.02244*.
5. Goodfellow, I., et al. (2013). Challenges in Representation Learning: A Report on Three Machine Learning Contests (FER-2013 Dataset). *Neural Networks*, 64, 59-63.
6. Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2021). EnlightenGAN: Deep Light Enhancement Without Paired Supervision. *IEEE Transactions on Image Processing*, 30, 2340-2350.

## NGƯỜI THỰC HIỆN ĐỀ CƯƠNG

<b>Sinh viên thực hiện 1</b>	Nguyễn Hữu Lộc	3122410213
<b>Sinh viên thực hiện 2</b>	Mai Phúc Lâm	3122410207
<b>Sinh viên thực hiện 3</b>	Văn Tuấn Kiệt	3122410202
<b>Sinh viên thực hiện 4</b>	Nguyễn Đức Duy Lâm	3122410208

## XÁC NHẬN CỦA NGƯỜI HƯỚNG DẪN

<b>Người hướng dẫn</b>	Th.S Đỗ Như Tài
<b>Ý kiến của người hướng dẫn:</b>	

