TRƯỜNG ĐẠI HỌC SÀI GÒN

**KHOA CÔNG NGHỆ THÔNG TIN**

****----- 🙡 🕮 🙣 -----

**MÔN PHƯƠNG PHÁP NGHIÊN CỨU KHOA HỌC**

**Tên Bài: Phân tích luận văn tốt nghiệp**

Giảng viên hướng dẫn: **TS. Đỗ Như Tài**

Sinh viên thực hiện:

Tạ Hồng Quí – 3122410348

**TP.Hồ Chí Minh**

Nội dung

[1. Tên đồ án 3](#_Toc195879055)

[2. Tác giả, nơi và năm công bố 3](#_Toc195879056)

[3. Mục tiêu nghiên cứu 3](#_Toc195879057)

[4. Cơ sở lý thuyết 3](#_Toc195879058)

[5. Phương pháp thu thập và xử lý thông tin 3](#_Toc195879059)

[6. Kết quả đạt được 4](#_Toc195879060)

[7. Hạn chế của công trình 4](#_Toc195879061)

# 1. Tên đồ án

**Nhận diện cảm xúc mặt người sử dụng mạng học sâu có chú ý**

# 2. Tác giả, nơi và năm công bố

* **Tác giả:** Phạm Quí Luận
* **GVHD:** TS. Trần Tuấn Anh
* **Nơi công bố:** Trường Đại học Bách Khoa, ĐHQG-HCM
* **Năm công bố:** 12/2019

# 3. Mục tiêu nghiên cứu

* Phát triển mô hình học sâu hiệu quả để **phân lớp cảm xúc khuôn mặt** từ ảnh tĩnh trong môi trường thực tế phức tạp.
* Đề xuất và đánh giá mô hình **Residual Masking Network (RMN)** sử dụng cơ chế chú ý.
* Xây dựng và thử nghiệm **bộ dữ liệu VEMO** – gồm ảnh khuôn mặt người Việt để cải thiện độ phù hợp trong bối cảnh nội địa.

# 4. Cơ sở lý thuyết

* **Biểu cảm khuôn mặt** là một hình thức giao tiếp phi ngôn ngữ quan trọng.
* Lý thuyết từ **Ekman & Friesen** về 6 cảm xúc cơ bản (giận dữ, ghê tởm, sợ hãi, hạnh phúc, buồn bã, ngạc nhiên).
* **Mạng nơ-ron tích chập (CNN)** và cơ chế **chú ý (Attention Mechanism)** giúp mô hình học được các vùng ảnh quan trọng.
* **Hệ thống mã hóa cơ mặt (FACS)** được dùng để phân tích cấu hình cơ mặt.

# 5. Phương pháp thu thập và xử lý thông tin

* Dữ liệu từ:
  + Bộ **FER2013** (ảnh gán nhãn biểu cảm trong môi trường thực tế).
  + Bộ **VEMO**: Tác giả tự thu thập ảnh khuôn mặt người Việt.
* Dựng mô hình học sâu **Residual Masking Network** gồm các khối học tập và masking có khả năng chọn lọc vùng ảnh.
* Thực hiện:
  + Huấn luyện mô hình trên dữ liệu được chuẩn hóa.
  + So sánh với các mô hình khác như VGG19, ResNet18, Inception V3.

# 6. Kết quả đạt được

* **RMN đạt độ chính xác 76.82% trên tập FER2013**, vượt một số phương pháp hiện đại.
* Mô hình có hiệu quả tốt trên cả tập dữ liệu tự thu thập (VEMO).
* Trực quan hóa GradCAM chứng minh mô hình học được vùng biểu cảm đúng trọng tâm.
* Cải tiến RMN giúp tăng khả năng khái quát và tập trung vào vùng biểu cảm.

# 7. Hạn chế của công trình

* **Độ chính xác chưa vượt quá 80%** trên dữ liệu thực tế phức tạp.
* Dữ liệu VEMO chưa đủ lớn, **chưa được công bố công khai** để cộng đồng kiểm nghiệm.
* Chưa tích hợp thông tin thời gian (temporal), tức **chỉ hoạt động tốt với ảnh tĩnh**.
* Không có phân tích sâu về yếu tố văn hóa ảnh hưởng đến biểu cảm (mới dừng ở mức giả định phổ quát).