

BÁO CÁO CUỐI KỲ

Môn học

CS2205.CH1501 -

PHƯƠNG PHÁP LUẬN NCKH

Giảng viên

PGS.TS. LÊ ĐÌNH DUY

Thời gian

03/2021 - 06/2021

----- *Trang này cố tình để trống* -----

HƯỚNG DẪN


Yêu cầu:

- *Bước 1: Chọn File/Make a copy để tạo ra một file theo template mẫu https://docs.google.com/document/d/1pu86lH6STGaVk2JH70n3jWx8qt9Eue_imVTQhg3s/. Đặt tên tập tin này là: CS2205.CH1501.RM.FinalReport.MSHV*
- *Bước 2: Điền các thông tin về đề cương đề tài vào file GDocs trên. Tối đa 6 trang.*
- *Bước 3: Copy toàn bộ nội dung đề cương đề tài và Paste vào cuối tập tin này (tránh ghi đè lên nội dung của HV khác).*
- *Bước 4: Nộp bài (Turn in) theo yêu cầu trên Classroom. Chọn Add or Create và chọn Link đến file Google ở trên. Lưu ý đặt quyền Anyone with the link - Viewer. Trong phần Private Comment, cung cấp thông tin của github repos, thông tin các thành viên của nhóm và các ghi chú khác nếu có. Lưu một phiên bản pdf của đề cương trên github repos*

Lưu ý:

- *Việc tuân thủ các hướng dẫn, các yêu cầu theo mẫu là bắt buộc và được đánh giá trong điểm tổng kết của đồ án môn học.*
- ***Deadline: 25/07/2021***

----- *Trang này cố tình để trống* -----

Họ và tên (IN HOA)	TRẦN BÌNH HẬU
Ảnh	
Số buổi vắng	0
Bonus	Ghi số lần comment trên Google Classroom: 22
Tên đề tài (VN)	NHẬN DIỆN CẢM XÚC NGƯỜI HỌC TRONG LỚP
Tên đề tài (EN)	RECEIVING THE EMOTION OF STUDENTS IN CLASS
Giới thiệu	<ul style="list-style-type: none"> ● <i>Bài toán: hỗ trợ đánh giá cảm xúc học viên trong lớp học, từ đó cải thiện chất lượng giảng dạy cho phù hợp.</i> ● <i>Lý do chọn đề tài:</i> <ul style="list-style-type: none"> ○ <i>Nhu cầu đánh giá chất lượng giảng dạy trong lớp học truyền thống và lớp học trực tuyến tăng mạnh;</i> ○ <i>Các giải thuật trong thị giác máy tính được cải thiện lớn về độ chính xác, hiệu năng xử lý làm tăng tính khả thi của đề tài. Ngoài ra giá thành thiết bị camera rẻ là yếu tố dễ mở rộng, triển khai giải pháp;</i> ● <i>Mô tả bài toán:</i>

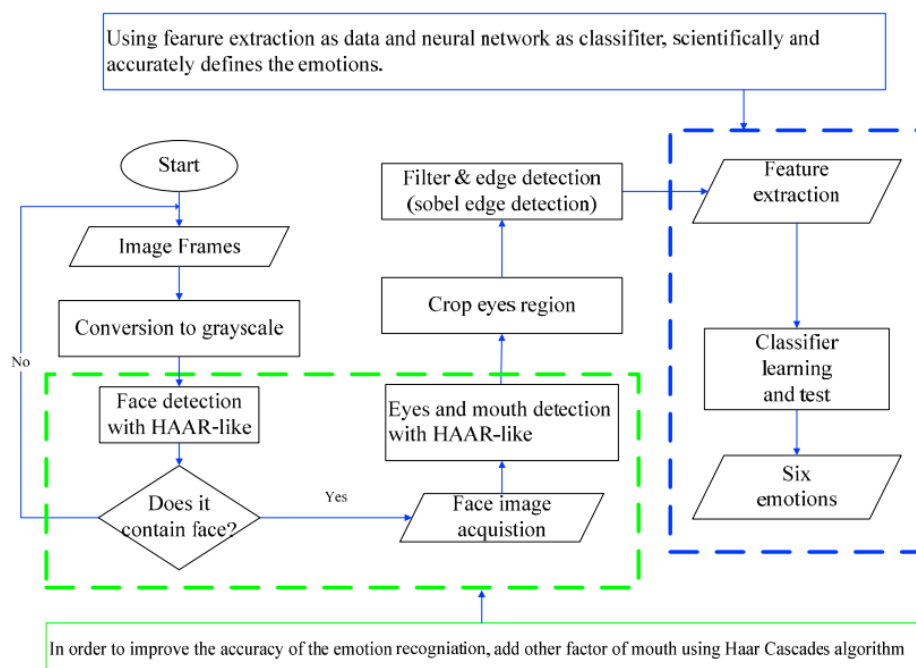
- Căn cứ: dữ liệu camera được lưu trữ trên hệ thống; dữ liệu cảm xúc từ bộ datasets chuẩn.
- Input: các frame image/video gương mặt học viên từ camera máy tính hoặc camera trong lớp học.
- Output: thông tin tỷ lệ cảm xúc của học viên trong lớp (phấn khởi/hào hứng, buồn ngủ/chán nản, khó chịu/giận dữ, trung tính/bình thường).
- Ràng buộc: camera phải bắt được cạnh mắt và miệng của gương mặt;
- Phạm vi: giải pháp phù hợp cho lớp học truyền thống và tích hợp được cho phần mềm dạy học trực tuyến.
- Hình ảnh minh họa kết quả



Mục tiêu

- Phân tích cảm xúc đạt độ chính xác trên 80% so với làm survey sau buổi học.
- Giải pháp có thể triển khai được cho lớp học truyền thống và lớp học trực tuyến (tích hợp vào phần mềm Learning Management System - LMS).

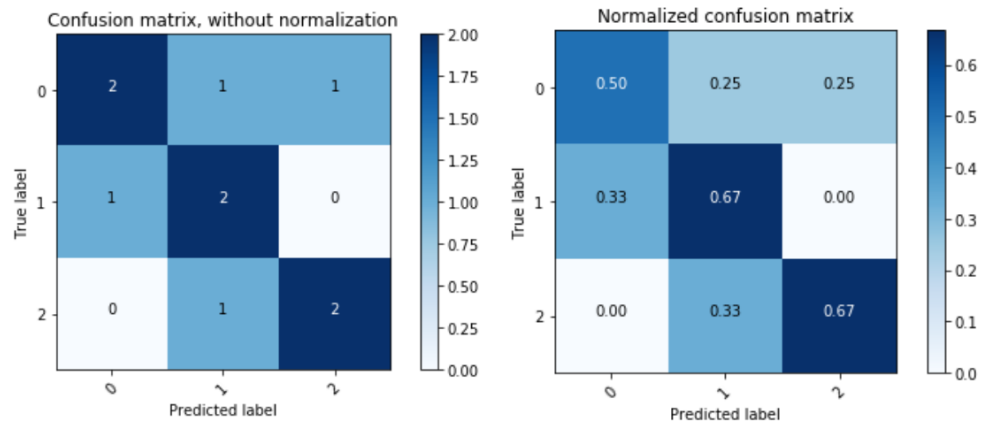
	<ul style="list-style-type: none"> • Tài nguyên máy chủ khi chạy phần mềm tăng 20% tương ứng 200 camera đổ dữ liệu về đồng thời. Tài nguyên phía máy tính/điện thoại người dùng tăng 1% khi sử dụng phần mềm học tập trực tuyến.
<p>Nội dung và phương pháp thực hiện</p>	<ul style="list-style-type: none"> • Nội dung 1: xây dựng giải pháp phân tích cảm xúc <ul style="list-style-type: none"> ○ Thu thập/đánh nhãn dữ liệu để xây dựng mô hình. ○ Train model, validation model và test model. ○ Thực hiện survey ngẫu nhiên các buổi học truyền thống và trực tuyến để so sánh. • Phương pháp 1: thực nghiệm và so sánh <ul style="list-style-type: none"> ○ Thu thập dữ liệu, đánh nhãn nhờ đội ngũ chuyên gia quản lý giáo dục đại học. ○ Dùng CNN (Keras và VGG16) để build model. <div data-bbox="507 1032 1425 1581"> <p>Diagram (a) illustrates the process of emotion recognition using CNN. It starts with an input image or frame from a video, which is processed by a CNN for background removal to extract a face vector. This vector is then used for expressional vector extraction, followed by a CNN for classification to predict the emotion (e.g., Mood: Happy).</p> <p>Diagram (b) shows a face with key points labeled: LE, LY, LF, RE, RF, RN, RP, CP, CH, and a vertical line through the center.</p> </div> <ul style="list-style-type: none"> ○ Thực hiện survey trong 50 buổi học ngẫu nhiên để so sánh với hệ thống. ○ Giải pháp thực hiện: phân tích góc cạnh mắt và miệng để nhận diện cảm xúc



- *Nội dung 2: tính khả thi của giải pháp cho lớp học trực tiếp và truyền thống*
 - *Thử nghiệm đối với camera doom, ptz đặt tại lớp học, mỗi lớp 2 camera ở 2 góc từ phía bảng.*
 - *Thử nghiệm đối với camera điện thoại, laptop, camera máy tính rời.*
 - *Tích hợp được vào hệ thống quản lý lớp học hoặc hệ thống LMS.*
- *Phương pháp 2: thực nghiệm, so sánh, phân tích đánh giá*
 - *Thử nghiệm giải pháp phân tích cảm xúc đối với tất cả các loại camera phổ thông (tích hợp sẵn, chi phí thấp từ vài trăm ngàn đến 1-2 triệu).*
 - *Viết API để các hệ thống quản lý khác có thể sử dụng giải pháp này, chỉ cần gọi API là có thể sử dụng được.*
- *Nội dung 3: nâng cao hiệu năng sử dụng hệ thống*
 - *Cải thiện tỷ lệ sử dụng tài nguyên hệ thống*
- *Phương pháp 3:*

	<ul style="list-style-type: none">○ Cân chỉnh camera với độ phân giải phù hợp để giảm dung lượng dữ liệu truyền về hệ thống nhưng vẫn đảm bảo nhìn rõ mặt.○ Dữ liệu phục vụ input cho hệ thống được phân tách riêng và chỉ tập trung khu vực mặt.○ Giải thuật được cải tiến liên tục để đảm bảo độ chính xác, tốc độ và sử dụng tài nguyên.○ Dữ liệu video được xử lý trực tiếp trên RAM trước khi lưu trữ kết quả xuống hệ thống lưu trữ.○ Hệ thống lưu trữ (SAN) sử dụng công nghệ SSD, kết nối FC 16Gbps để tăng tốc độ read/write dữ liệu.○ Sử dụng hệ thống giám sát vận hành phần mềm (hệ thống monitoring).												
Kết quả dự kiến	<ul style="list-style-type: none">● Kết quả:<ul style="list-style-type: none">○ Thuật toán cho độ chính xác cao, tốc độ thực thi nhanh.○ Giải pháp triển khai, mở rộng, có thể áp dụng cho lớp học truyền thống và trực tuyến, chạy được đa nền tảng.○ Hiệu năng cao, ít tốn năng lượng, tốc độ xử lý nhanh, tối ưu dữ liệu truyền và lưu trữ.● Thuật toán: CNN (VGG16) để phân loại và nhận diện cảm xúc.● Bộ dữ liệu:<ul style="list-style-type: none">○ FER2013 (35,887 grayscale images từ Google), MMI (2900 videos), MultiPie(hơn 750k images captured).○ Bộ dữ liệu tự xây dựng (500 videos từ camera).● So sánh giữa các phương pháp:<table><thead><tr><th>Authors</th><th>Datasets</th><th>The architecture used</th><th>Recognition rate</th></tr></thead><tbody><tr><td>Mollahosseini et al</td><td>MultiPie, MMI, FER2013</td><td>CNN</td><td>94.7%, 77%, 61.1%</td></tr><tr><td></td><td>Bộ dữ liệu tự xây dựng</td><td>CNN</td><td>92%</td></tr></tbody></table>	Authors	Datasets	The architecture used	Recognition rate	Mollahosseini et al	MultiPie, MMI, FER2013	CNN	94.7%, 77%, 61.1%		Bộ dữ liệu tự xây dựng	CNN	92%
Authors	Datasets	The architecture used	Recognition rate										
Mollahosseini et al	MultiPie, MMI, FER2013	CNN	94.7%, 77%, 61.1%										
	Bộ dữ liệu tự xây dựng	CNN	92%										

- *Đánh giá thuật toán dựa trên Confusion matrix, F1 Score, MAP.*



- *Github: <https://github.com/hautb15/CS2205.CH1501.git>*

**Tài liệu
tham
khảo**

- [1] L. Đ. Duy, Phát Triển Ứng Dụng Camera Thông Minh, Hồ Chí Minh: NXB Đại học CNTT, 2018.
- [2] P. R. Dachapally, "Facial emotion detection using convolutional neural networks and representational autoencoder units," arXiv preprint arXiv:1706.01509, 2017.
- [3] S. Ghosh, Hiware, Kaustubh, Ganguly, Niloy, Mitra, Bivas and De, Pradipta, "Emotion detection from touch interactions during text entry on smartphones," International Journal of Human-Computer Studies, vol. 130, no. 1071-5819, pp. 47-57, 2019.
- [4] Kumari, Rina, Ashok, Nischal, Ghosal, Tirthankar and Ekbal, Asif, "Misinformation detection using multitask learning with mutual learning for novelty detection and emotion recognition," Information Processing & Management, vol. 58, no. 5, p. 102631, 01 09 2021.
- [5] N. Mehendale, "Facial emotion recognition using convolutional neural networks (FERC)," SN Applied Sciences, vol. 2, no. 3, p. 446, 18 02 2020.

[6] Mustaqeem, and Kwon, Soonil, "Att-Net: Enhanced emotion recognition system using lightweight self-attention module," *Applied Soft Computing*, vol. 102, p. 107101, 01 04 2021.

[7] Zhao, Jianfeng, Mao, Xia and Chen, Lijiang, "Speech emotion recognition using deep 1D & 2D CNN LSTM networks," *Biomedical Signal Processing and Control*, vol. 47, pp. 312-323, 01 01 2019.