

BÁO CÁO ĐỒ ÁN CUỐI KỲ

Môn học: CS2205 - PHƯƠNG PHÁP LUẬN NCKH

Lớp: CS2205.MAR2024

GV: PGS.TS. Lê Đình Duy

Trường ĐH Công Nghệ Thông Tin, ĐHQG-HCM



TẠO CÂU MÔ TẢ TIẾNG VIỆT MANG YẾU TỐ CẢM XÚC CHO ẢNH DỰA TRÊN TIẾP CẬN HỌC SÂU

Nguyễn Trường Khoa Nguyên-230201004

Tóm tắt

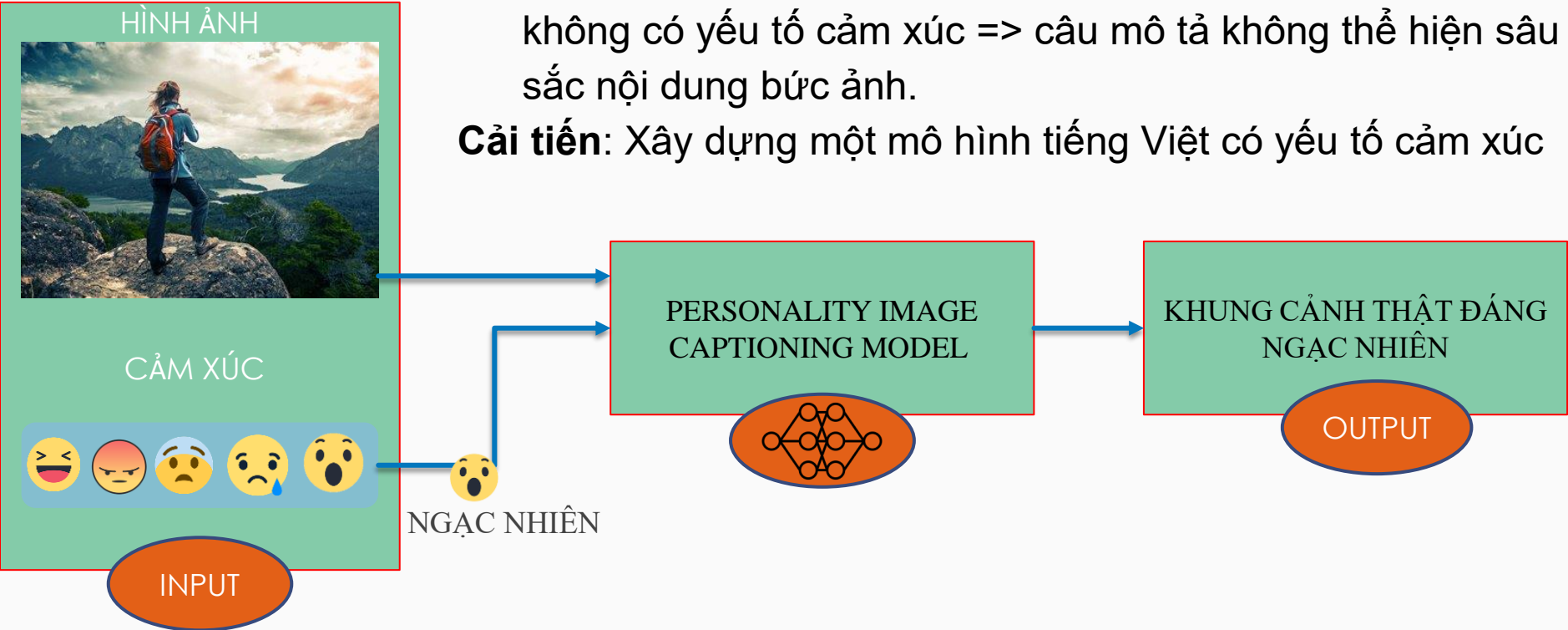
- Lớp: CS2205.MAR2024
- Link Github: <https://github.com/nguyennguyentruong/CS2205>
- Link YouTube video:
- Ảnh + Họ và Tên: Nguyễn Trường Khoa Nguyễn



Giới thiệu

- Các mô hình tạo mô tả ảnh thường làm về tiếng Anh và không có yếu tố cảm xúc => câu mô tả không thể hiện sâu sắc nội dung bức ảnh.

Cải tiến: Xây dựng một mô hình tiếng Việt có yếu tố cảm xúc

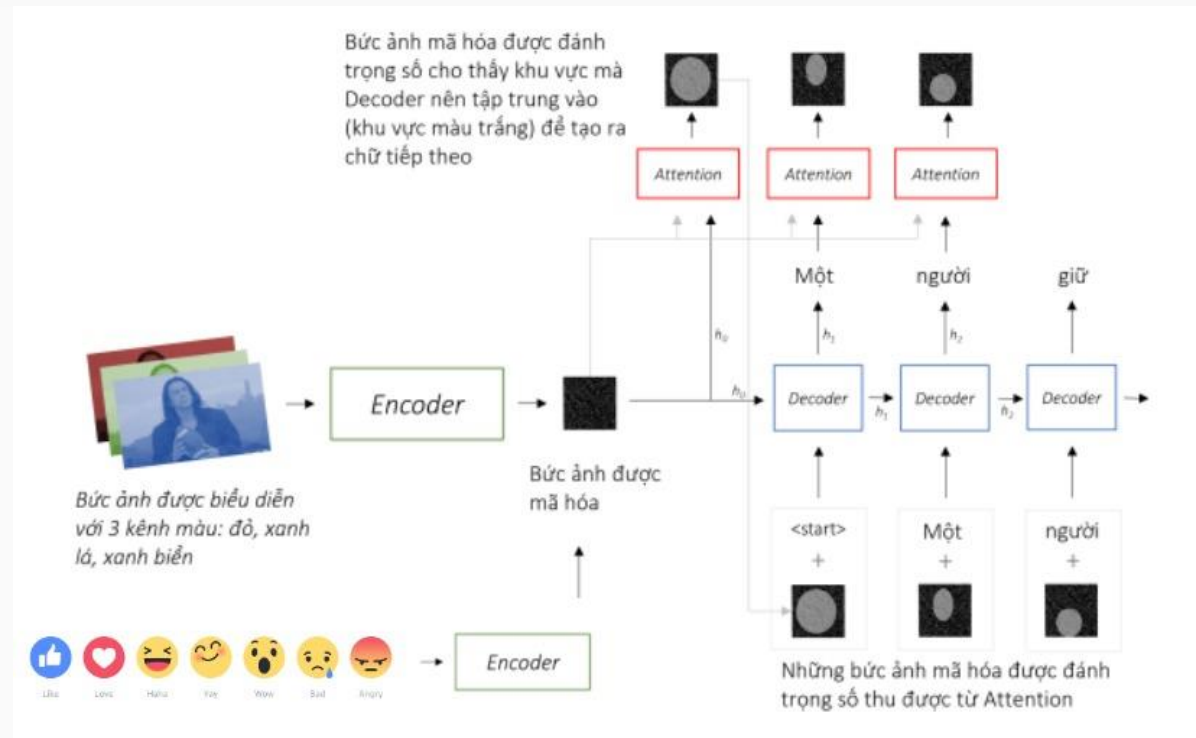


Mục tiêu

- Xây dựng bộ dữ liệu câu mô tả cảm xúc tiếng Việt.
- Xây dựng một mô hình tự động tạo câu mô tả cho ảnh bằng tiếng Việt, nhằm hỗ trợ việc hiểu nội dung hình ảnh và cải thiện tính hiệu quả của các mô hình học máy.
- Xây dựng ứng dụng tích hợp mô hình để người dùng có thể tương tác.

Nội dung và Phương pháp

Sử dụng mô hình
show attend and tell
và thêm layer encode
cảm xúc



Nội dung và Phương pháp

- Tìm hiểu kiến trúc của mô hình Show Attend and Tell và cách xử lý dữ liệu ảnh và văn bản.
- Thiết kế layer mới trong mô hình để nhúng các yếu tố cảm xúc.
- Xây dựng bộ dữ liệu "personality captions" dựa trên phương pháp phân loại của Facebook và sử dụng API chuyển đổi từ tiếng Anh sang tiếng Việt.
- Nghiên cứu các độ đo để đánh giá mô hình.
- Huấn luyện và đánh giá mô hình với bộ dữ liệu mới.
- Triển khai ứng dụng tích hợp mô hình

Kết quả dự kiến

- Mô hình thành công trên tập dữ liệu, tạo câu mô tả phù hợp với nội dung và cảm xúc.
- Xây dựng bộ dữ liệu tiếng Việt đa dạng, phản ánh sự đa dạng của đối tượng và hành động trong ảnh, cũng như tính cảm xúc và cá nhân hóa của người mô tả.
- Triển khai thành công ứng dụng tích hợp mô hình.

Tài liệu tham khảo

- [1]. Shuang Bai and Shan An. “A Survey on Automatic Image Caption Generation”. in: Neurocomputing 311 (may 2018). doi: 10.1016/j.neucom.2018.05.080
- [2]. Nicholas Becherer and others. “Improving optimization of convolutional neural networks through parameter fine-tuning”. in: Neural Computing and Applications (2017), pages 1–11.
- [3]. Oriol Vinyals and others. “Show and Tell: A Neural Image Caption Generator”. in: CoRR abs/1411.4555 (2014). arXiv: 1411.4555. url: <http://arxiv.org/abs/1411.4555>.
- [4]] K. Shuster, S. Humeau, H. Hu, A. Bordes, and J. Weston, “Engaging image captioning via personality,” in CVPR, 2019