

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

Nguyễn Trường Khoa Nguyễn^{1,2,3}

¹ Trường ĐH Công nghệ thông tin TPCHM

² University of Information Technology

³ Đại học Quốc gia TP HCM

What ?

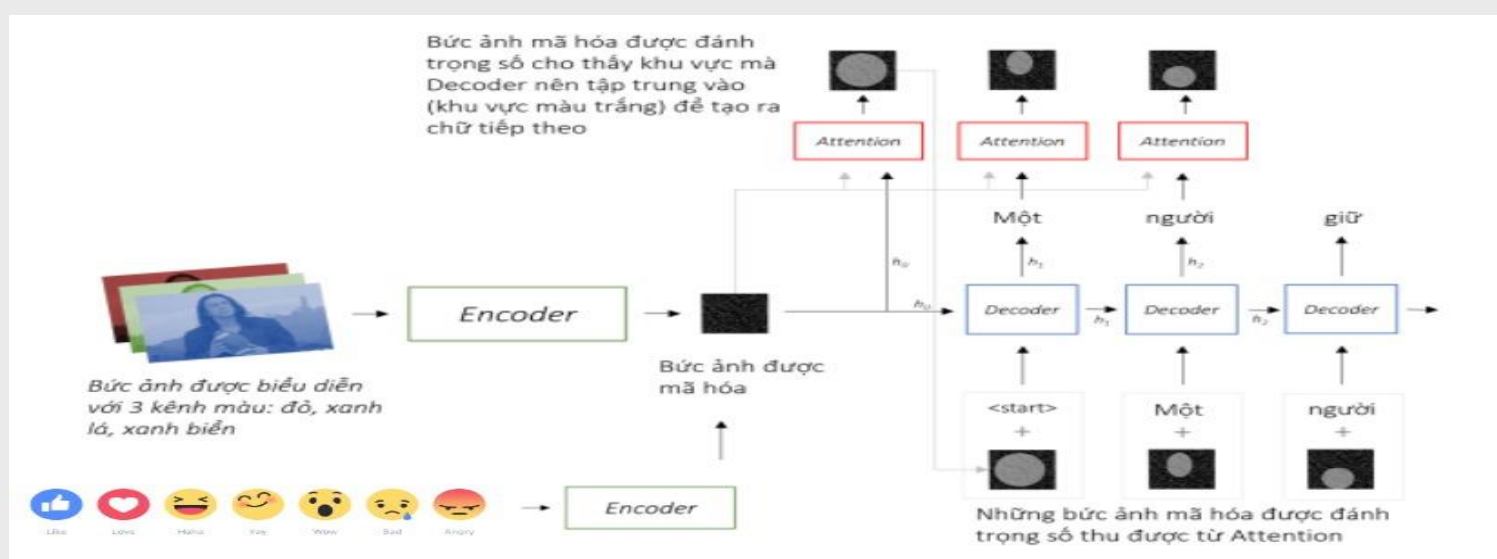
Đề xuất phương pháp xây dựng bộ dữ liệu mô tả cảm xúc tiếng Việt, và một mô hình cải tiến sử dụng thêm yếu tố cảm xúc để tạo mô tả ảnh.

- Xây dựng một bộ dữ liệu đa dạng câu mô tả mang yếu tố cảm xúc cho ảnh.
- Cải tiến mô hình Show Attend and Tell cho phép người dùng nhúng cảm xúc vào mô hình.
- Triển khai ứng dụng tích hợp mô hình hỗ trợ tương tác.

Why ?

- Bài toán tuy được nghiên cứu rộng rãi, nhưng các nghiên cứu thường tập trung vào việc tạo các câu mô tả tiếng Anh. Đồng thời câu tạo ra chỉ nói lên sự vật, hiện tượng của bức ảnh, không có tính cá nhân hóa để làm sinh động hơn.
- Tập trung xây dựng bộ dữ liệu tiếng Việt cho bài toán, và cải thiện mô hình để có thể tạo ra câu mô tả mang tính cá nhân hóa.

Overview



Description

1. Xây dựng bộ dữ liệu

- Xây dựng bộ dữ liệu "personality captions" gồm việc lựa chọn các cảm xúc sử dụng, hình ảnh từ các bộ dữ liệu Flickr30k, MSCOCO, YFCC100M.
- Nghiên cứu, chọn lọc các cảm xúc cơ bản để sử dụng (buồn, giận dữ, sợ hãi, ngạc nhiên).
- Tiến hành dán nhãn tiếng Việt cho hình ảnh với cảm xúc tương ứng.

PERSONALITY-CAPTIONS			
TẬP CHIA	TRAIN	VALID	TEST
Số lượng ảnh	186 858	5000	10 000
Số lượng chủ thích	186 858	5000	50 000
Số loại cảm xúc	215	215	215
Số lượng từ trung bình cho mỗi chủ thích	11	11	11
Số câu trung bình trên một cảm xúc	861	23	46

Bảng 1. Thống kê bộ dữ liệu xây dựng

2. Trích xuất đặc trưng ảnh và cảm xúc

- Tiền xử lý các dữ liệu ảnh (kích cỡ,...) và trích xuất đặc trưng của chúng qua việc sử dụng các mạng CNN như: ResNet, VGGNet, Inception, Densenet
- Sử dụng các mô hình nhúng từ phổ biến như: word2Vec, Glove để lấy đặc trưng của cảm xúc.
- Kết hợp đặc trưng của cảm xúc và hình ảnh thành đặc trưng mới để tiến tạo mô tả.

3. Tạo mô tả cho ảnh.

- Sử dụng LSTM, với đầu vào là đặc trưng được trích xuất của ảnh và cảm xúc.
- Kết hợp từ tạo ra với cơ chế Attention để làm đầu vào tạo từ nối tiếp theo.

4. Triển khai ứng dụng

- Tiến hành triển khai mô hình lên server trên nền tảng cloud thích hợp.
- Sử dụng các công nghệ như docker, react, ... Để thiết kế chương trình tương tác với mô hình.
- Chương trình cho phép người dùng upload ảnh cũng như chọn cảm xúc cho bức ảnh đó, và sẽ trả về cho người dùng câu mô tả mang yếu tố cảm xúc đã lựa chọn.