

# GÁN NHÃN CHO HÌNH ẢNH

Trương Thanh Luân

Trường ĐH Công Nghệ Thông Tin

## What ?

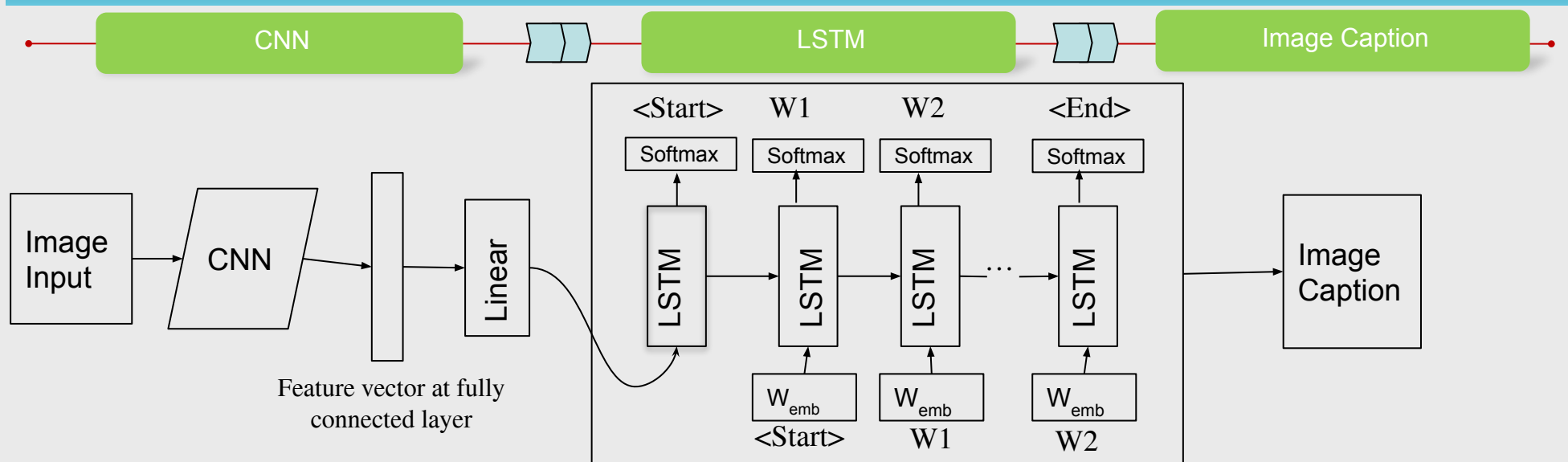
Kết hợp mô hình CNN và LSTM để gán nhãn cho hình ảnh:

- CNN được sử dụng để trích xuất các đặc trưng hình ảnh.
- LSTM sử dụng thông tin từ CNN để tạo mô tả cho hình ảnh.
- Gán nhãn cho hình ảnh.

## Why ?

- Giúp những người khiếm thính hoặc các em nhỏ có thể biết được cảnh vật xung quanh. Ví dụ: Image -> text -> voice.
- Quản lý và tìm kiếm hình ảnh theo nội dung hình ảnh.

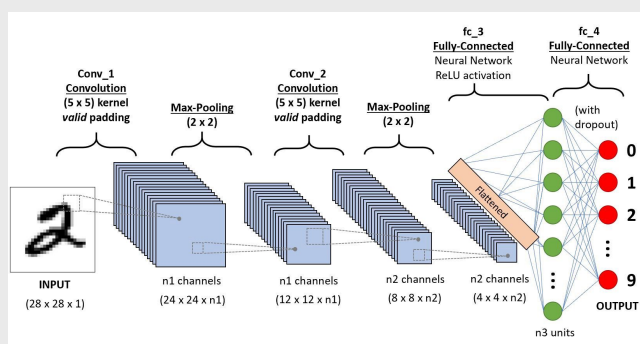
## Overview



## Description

### 1. Mô hình CNN

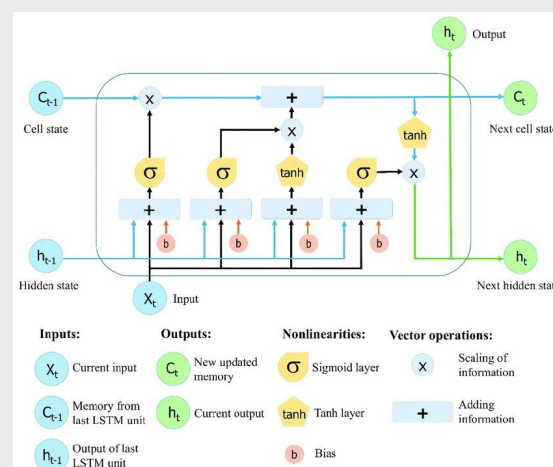
- Các lớp tích chập của CNN giúp tìm ra các đặc trưng từ hình ảnh. Ví dụ: cấu trúc như cạnh, góc, hoặc hình dạng của các đối tượng trong hình ảnh.
- Các đặc trưng hình ảnh sẽ được trích xuất từ InceptionV3, mô hình CNN được đào tạo trên bộ dữ liệu Imagenet. Sử dụng các lớp tích chập và lớp pooling, CNN có khả năng tự động học và trích xuất các đặc trưng quan trọng từ hình ảnh.



Hình 1. Kiến trúc mô hình CNN

### 2. Mô hình LSTM

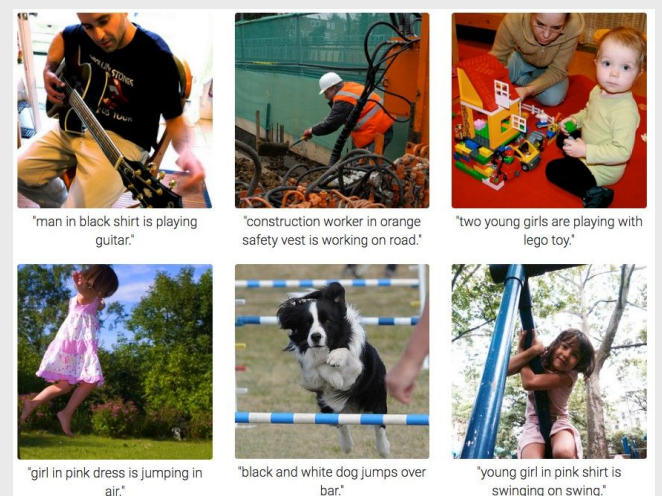
- Khả năng ghi nhớ và xử lý thông tin từ quá khứ.
- Phân tích chuỗi hình ảnh liên tiếp để dự đoán kết quả trong tương lai.
- Chống hiện tượng "Vanishing Gradient".



Hình 2. Kiến trúc mô hình LSTM

### 3. Kết quả mong đợi

- Độ chính xác cao.
- Tự động hóa quá trình gán nhãn.
- Xử lý hình ảnh phức tạp và chuỗi hình ảnh.
- Tự động hóa quá trình gán nhãn.



Hình 3. Ảnh được gán nhãn từ mô hình CNN và LSTM