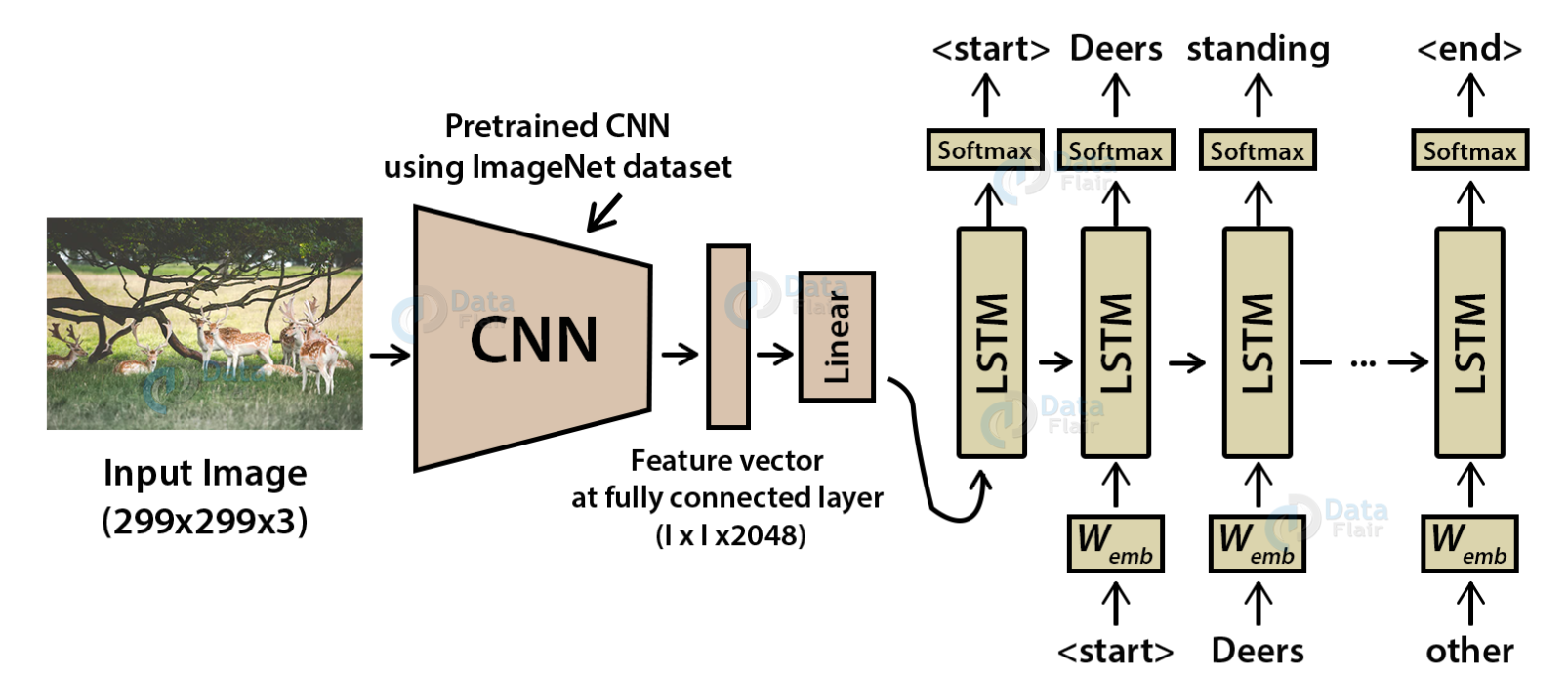
**Yêu cầu:**

Ở bài tiểu luận này chúng ta sẽ giải quyết bài toán Image captioning - kết hợp computer vision và NLP.

Một model kiểu mẫu cho bài toán này có thể được biểu diễn như sau:



Một lời giải cho bài toán này sẽ dựa vào một tutorial từ tensorflow: <https://www.tensorflow.org/tutorials/text/image_captioning>. Đây cũng là nhiệm vụ cơ bản của bài tiểu luận này.

Ngoài ra, nhiệm vụ mở rộng của bài toán sẽ bao gồm việc khảo sát và thực thi những kiến trúc khác khả thi ở hai giai đoạn CNN và RNN/LSTM: Attention, Transformer ...

Hạn nộp bài là vào 23h55 03/01/2022. Các bạn chạy trên Google Colab (runtime GPU), chọn hyperparameters, training and testing

Các bạn nén **tất cả code (bao gồm jupyter notebook), models và báo cáo** vào một file zip có dạng : ho\_ten\_MSSV.zip

1. Tìm hiểu lý thuyết

Image captioning là gì ?

RNN là gì?

LSTM là gì?

Sau đó chọn RNN/LSTM để chạy thử, chẳng hạn: Attention, Transformer ...

Lên plan ngày nào chạy những gì …

Chủ nhật xong

1. Chạy thử tutorial

<https://www.tensorflow.org/tutorials/text/image_captioning>

Vào đường link trên và chạy bằng google colab. Sau đó đọc hiểu code.

Thứ hai xong

1. Khảo sát

Mục tiêu: 2 ngày khảo sát một param

Thứ 3 + thứ 4:

Thứ 5 + Thứ 6:

Thứ 7:

chủ nhật:

1. Viết báo cáo:

Thứ hai