



Image Captioning in Vietnamese Language

Supervisor:

Dr. Phan Việt Anh

Trainees:

Nguyễn Minh Dũng

Nguyễn Duy Nhất

Lê Công Pha

Võ Minh Tâm

Feb 25th, 2021

CONTENTS

- I. Overview
- II. Dataset
- III. Approach
- IV. Results

I. OVERVIEW

Image Captioning?

1. A woman with a racket goes to hit a tennis ball.
2. The tennis player is ready to hit the ball.



Src: MS COCO Dataset

I. OVERVIEW

Image Captioning?

Vietnamese?

1. A woman with a racket goes to hit a tennis ball.
2. The tennis player is ready to hit the ball.



Src: MS COCO Dataset

I. INPUT - OUTPUT



Input

Vận động viên tennis nữ
đang vung vợt đỡ bóng.

Output

I. APPLICATIONS




**Be the Eyes
of Blind or
Visually Impaired
People**

Src: Internet

II. DATASET

UIT-ViLC (UIT-Vietnamese Image Captioning)

Authors	Quan Hoang Lam, Quang Duy Le, Kiet Van Nguyen, Ngan Luu-Thuy Nguyen (The UIT NLP Group, University of Information Technology - VNU HCM) [1]	
Ori-src	MS COCO - sportball	
Train set	2695 imgs	3850 imgs 5 captions/img 10-15 tokens/caption mostly
Val set	924 imgs	
Test set	231 imgs	

II. Dataset



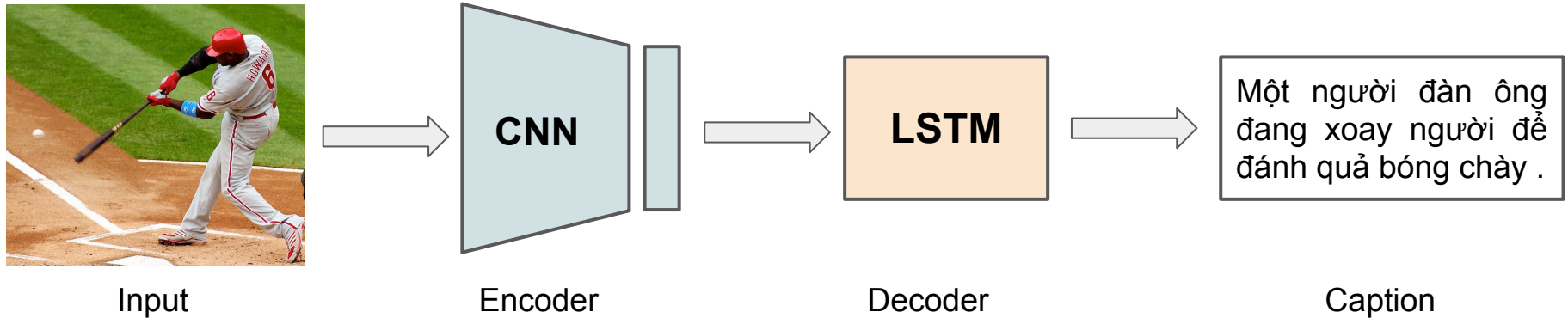
1. Các cầu thủ bóng đá **ddang** di chuyển ở trên sân để tranh bóng.
2. Một nam thủ môn đang **truwowc** trên sân để bắt bóng.
3. Các cầu thủ bóng đá đang thi đấu ở trên sân.
4. Một cầu thủ áo xanh đang khụy gối ở cạnh quả bóng.
5. Những cậu bé đang chơi bóng đá ở trên sân.

III. APPROACH

- **Preprocessing mistakes:**
 - Grammar
 - Spelling
 - Extra spaces
 - Punctuation
 - Vietnamese's accent signs

III. APPROACH

- **Model architecture**



IV. RESULT

MODEL	BLUE_1	BLUE_2	BLUE_3	BLUE_4	ROUGE_I	CIDEr



... Image_Index: 1351
Image_ID: 53360

cap_id 18553: Một người đang giơ vợt lên cao để đánh quả bóng tennis .

cap_id 4074: Một người đàn ông đang nghiêng người để đánh quả bóng tennis .

cap_id 18056: Một người phụ nữ đang chuẩn bị đánh quả bóng tennis .

cap_id 13625: Một người phụ nữ đang nghiêng người để đánh quả bóng tennis .

cap_id 13386: Một người đàn ông đang chuẩn bị đánh quả bóng tennis .

