



# Image Captioning via Vision and Language Transformers

Luke Davidson - [davidson.lu@northeastern.edu](mailto:davidson.lu@northeastern.edu)  
Kishore Pagidi - [pagidi.k@northeastern.edu](mailto:pagidi.k@northeastern.edu)  
Nand Dave - [nand.da@northeastern.edu](mailto:nand.da@northeastern.edu)

April 20, 2023

# Motivation

## What is Image Captioning?

- The combination of computer vision and natural language processing techniques to automatically generate accurate captions of unobserved input images

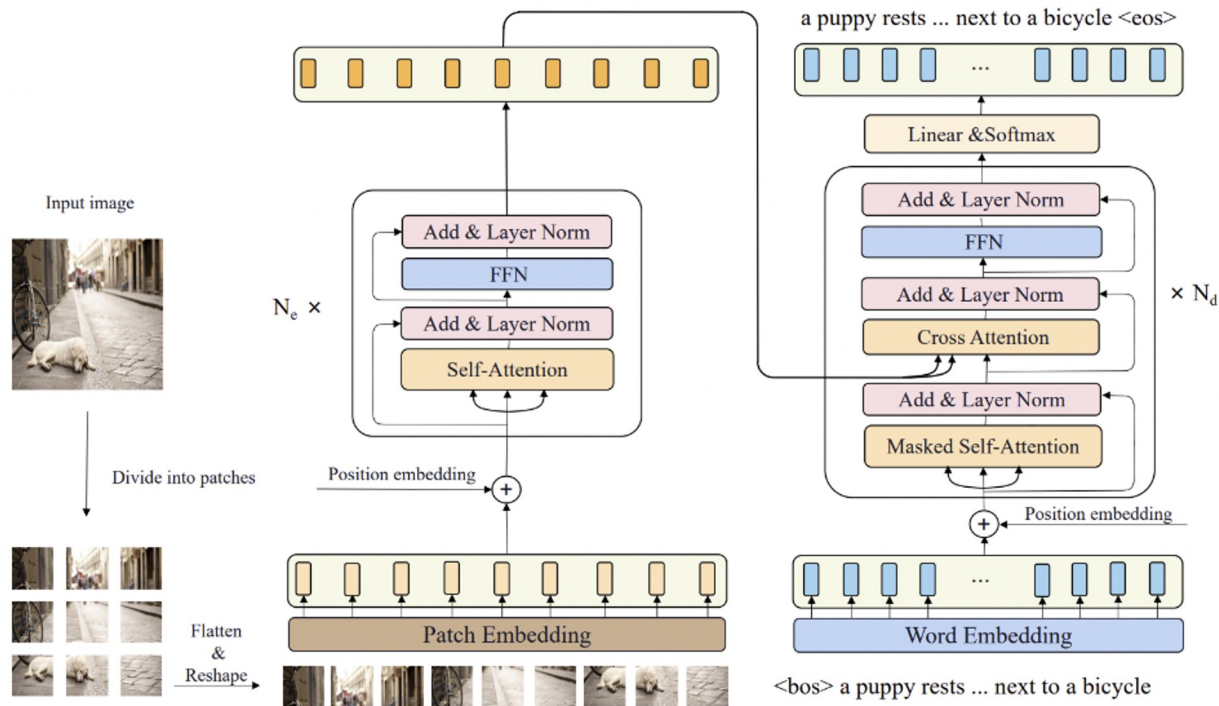
## Use Cases

1. Marketing and media
  - a. Automated captioning of social media posts or campaigns
2. Production
  - a. Recommendation systems for image and video editing
3. Assistance Services
  - a. Aiding visually impaired individuals in describing a live view or still image of something



the young boys are practicing tricks on their skate boards.  
a boy doing a trick on a skateboard on a rail.  
this is a skateboarder doing a dangerous trick.  
a boy on a skateboard going down a handrail at a set of stairs outside.  
a boy skateboards down a hand rail while two others watch.

# Approach



# Results



- Present results
- Describe dataset
- Graphs, plots, performance measures used

# Conclusion



**Takeaways**

**Achievements**

**Future Work**