NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

School of Computer Science and Engineering
College of Engineering

Mining Big Spatial Data

Student: Ng Zhi Kai **Supervisor: Cong Gao**

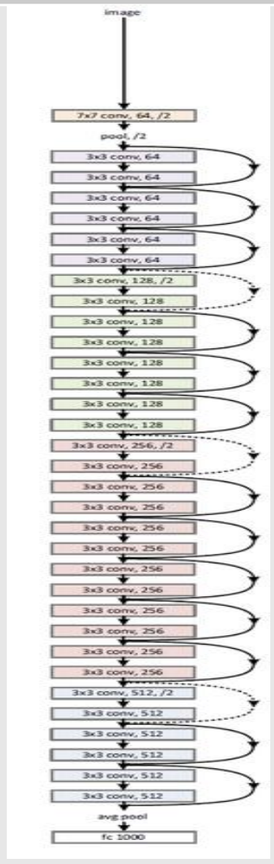
Project Objectives:

Proposed a road network representation by integrating image and traditional data sources, aiming to enhance performance across diverse downstream task.

Comparison of Image Encoders

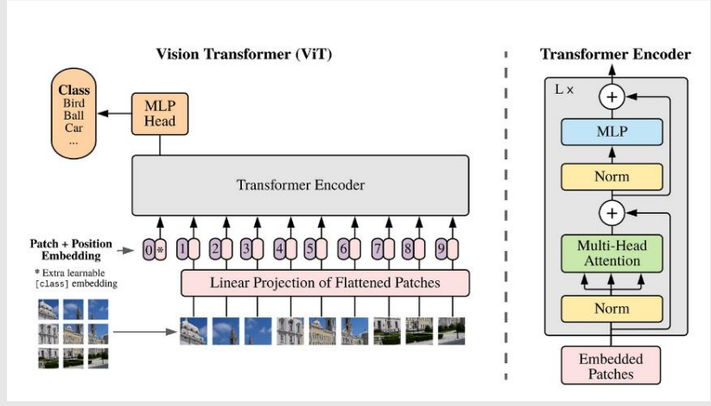
ResNet-101

- Convolutional Neural Network (CNN)
- Capture complex features within images
- Employs skip-connections



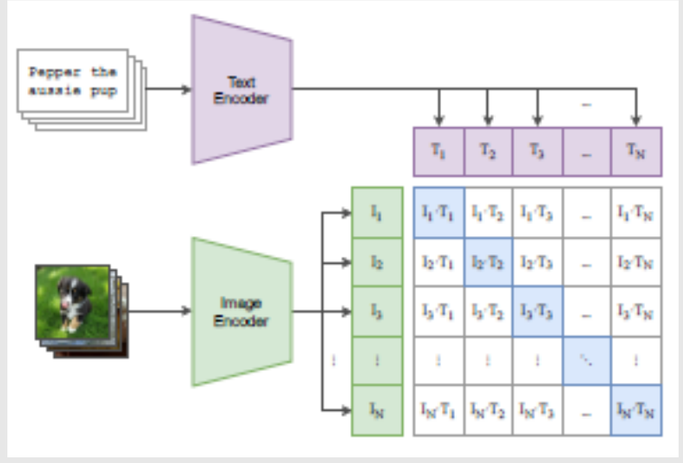
Vision Transformer(ViT)

- Patch embeddings
- Spatial embeddings
- Transformer encoder



Contrastive Language-Image Pretraining (CLIP)

- Combines vision and language
- Vision Encoder
- Text Encoder
- Contrastive objective



Additional Improvements:

- Aggregation of Embeddings
- Graph Neural Network(GNN)
- Point of Interest(POI)
- Road Features

Proposed Architecture:

1. Image Encoder
2. Text Encoder
3. Aggregation
4. Graph Neural Network(GNN)
5. Regression/ Classification model
6. Combining images with POI and road features

Down Stream Tasks:

1. Road Type Classification
2. Average Speed Prediction

Proposed Architecture