Image Captioning using Deep Learning

Mohammed Zoaib Mustafa

Project Context

Objectives

- Computer vision
- Natural language

Scope

- Data preparation
- Model evaluation

Libraries

- **TensorFlow/Keras:** Framework for building deep learning models
- **NumPy:** Library for numerical computations
- **Matplotlib:** Visualization tool for plotting
- **NLTK:** Natural Language Toolkit for text processing
- **Pickle:** Data serialization utility
- **tqdm:** Progress bar for loops

Model Architecture

Encoder

VGG16 extracts features, providing a rich representation of the input images.

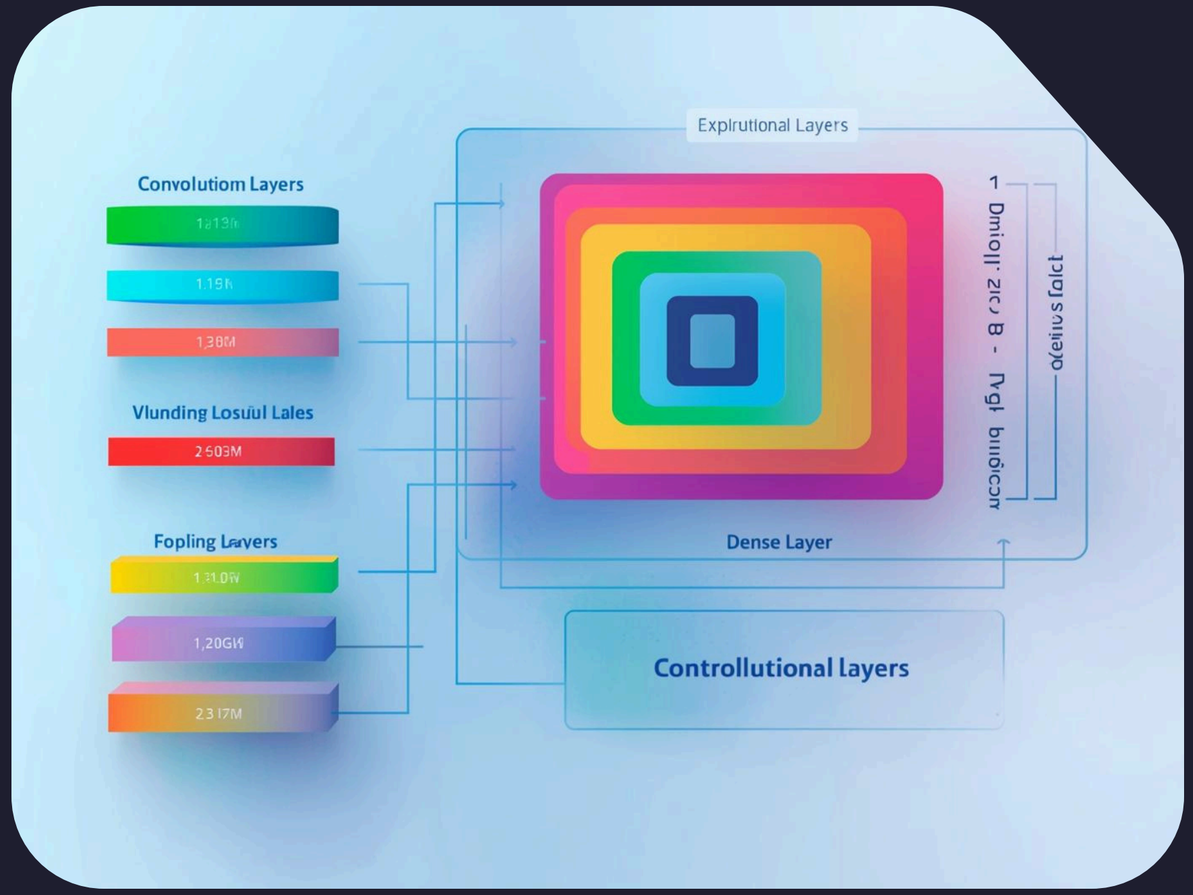
Decoder

LSTM generates captions by interpreting image features and producing textual descriptions.

Model Architecture Overview

Feature Extractor using VGG16 Pretrained Model

VGG16 extracts 4096-dimensional image features effectively



Data Preparation

Extraction

Preprocessing images helps extract features for effective model training.

Preprocessing

Cleaning and tokenizing captions prepares data for accurate text generation.

Data Preparation

Cleaning

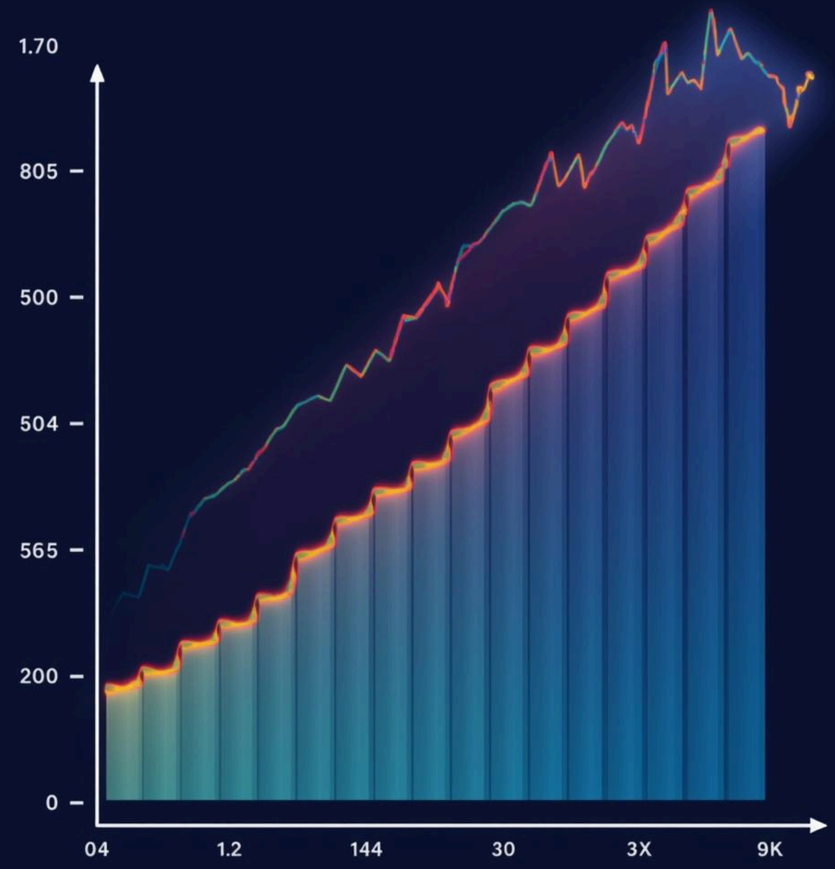
Lowercasing and punctuation removal ensure uniformity for effective text processing.

Tokenization

Converting words into integers simplifies sequence modeling for the neural network.

Training

- **20 epochs** for comprehensive learning
- Batch size of **32** for efficient processing
- Utilized a **custom data generator** for feeding image-text pairs
- Loss function: **Categorical Crossentropy** for optimizing predictions
- Optimizer: **Adam** for effective convergence



Model Evaluation

Metric

BLEU Score measures accuracy

BLEU-1

N-gram overlap of unigrams

BLEU-2

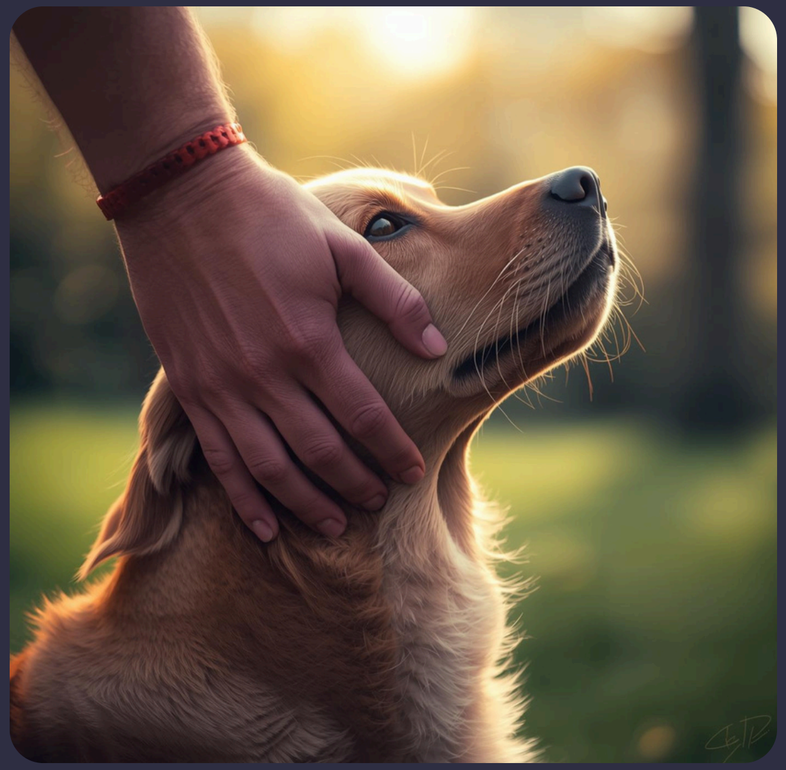
N-gram overlap of bigrams

Significance

Low scores indicate challenges

Results and Predictions

Model outputs for various images analyzed



Painting a Rainbow

Illustrative example of model predictions

Skier Analyzing Artwork

A moment of reflection in creativity

Conclusion and Key Takeaways

Email
zoaibmustafa2004@gmail.com

instagram
@_zoaib_

Phone
+91 7207190304