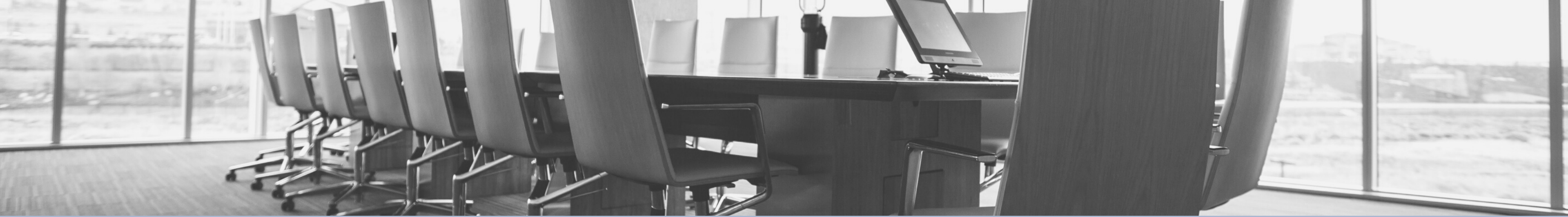




IMAGE CAPTIONING

TEAM NEOCOGNITRON



MEMBERS

ZERNAB SAEED

ALEX LOPEZ

DANIEL SINDELL

PRATAP KARKI

MICHAEL KWARTENG

IMAGE CAPTIONING

INTRODUCTION

- **Image Captioning** is the process of generating textual description of an image.
- It uses both **Natural Language Processing** and **Computer Vision** to generate the captions.
- **(NLP)** is a branch of artificial intelligence that helps computers understand, interpret and manipulate human language.
- Computer vision is a field of artificial intelligence that trains computers to interpret and understand the visual world.



MODEL OVERVIEW

A BACKGROUND

Image captioning has various applications such as recommendations in editing applications, usage in virtual assistants, for image indexing, for visually impaired persons, for social media, and several other natural language processing applications.

It has been demonstrated that deep learning models are able to achieve optimum results in the field of caption generation problems

Goal was to explore this field by leveraging the tools we learned in class to understand this concept

MODEL HIGHLIGHTS

DATA PREPARATION

process of cleaning and transforming raw data prior to processing and analysis. It is an important step prior to processing and often involves reformatting data, making corrections to data and the combining of data sets to enrich data.

ENCODER

The Convolutional Neural Network(CNN) can be thought of as an encoder. The input image is given to CNN to extract the features. The last hidden state of the CNN is connected to the Decoder.

DECODER

The Decoder is a Recurrent Neural Network(RNN) which does language modelling up to the word level. The first time step receives the encoded output from the encoder and also the <START> vector.

DATA

- IMAGENET DATASET
- CAPTIONS TEXTFILE
- CLEANING DATA
- MAPPING DATA
- TOKENIZING DATA

startseq dog sheep endseq

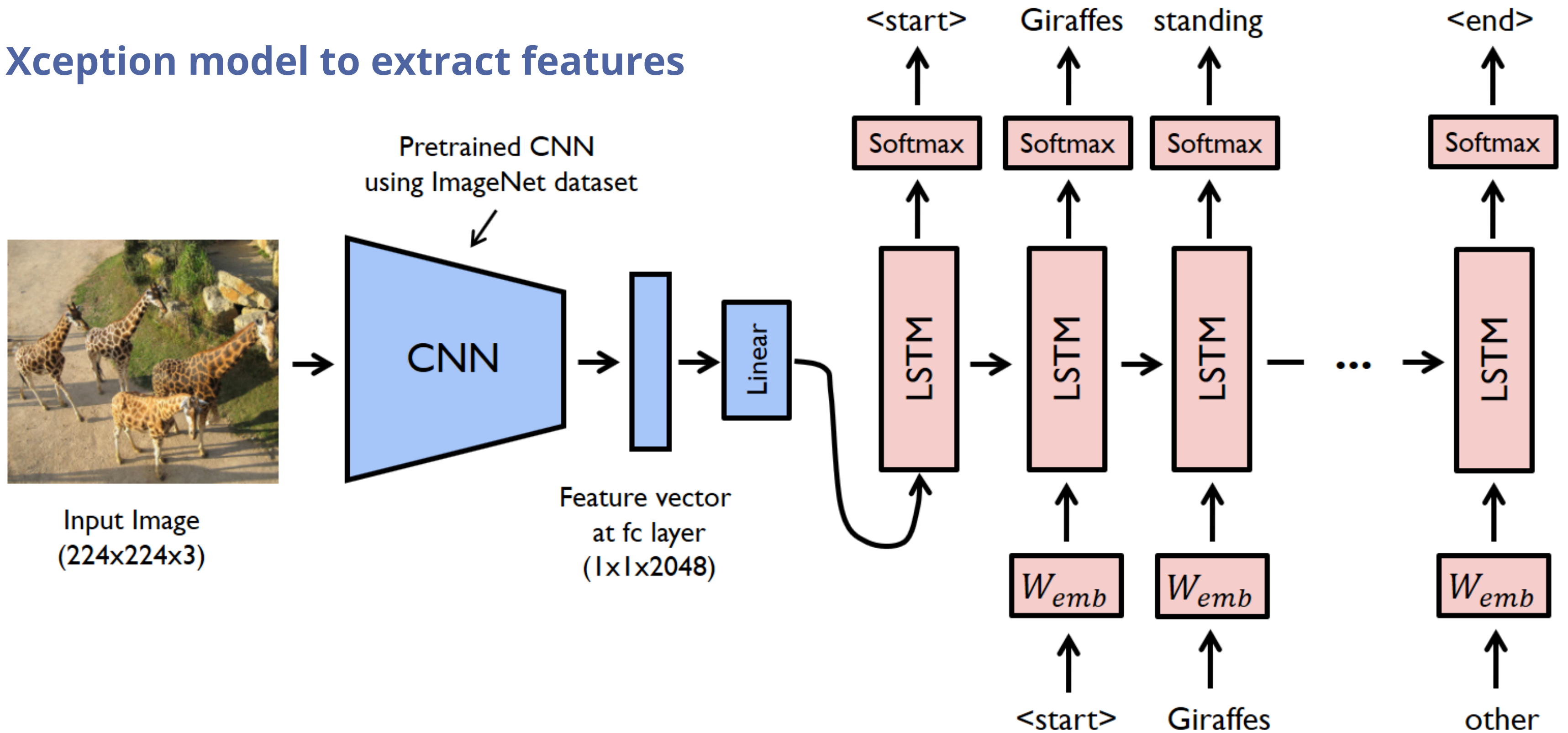


startseq two puppies playing endseq



ENCODER - DECODER ARCHITECTURE

- Xception model to extract features



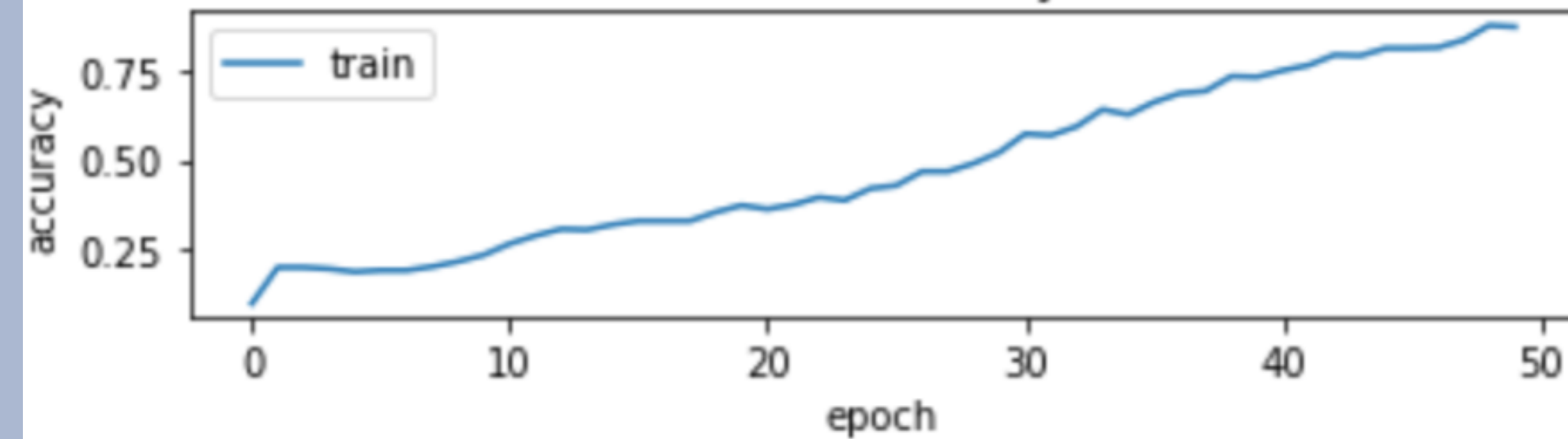
- Adam Optimizer and Categorical Cross Entropy

TRAINING AND TESTING

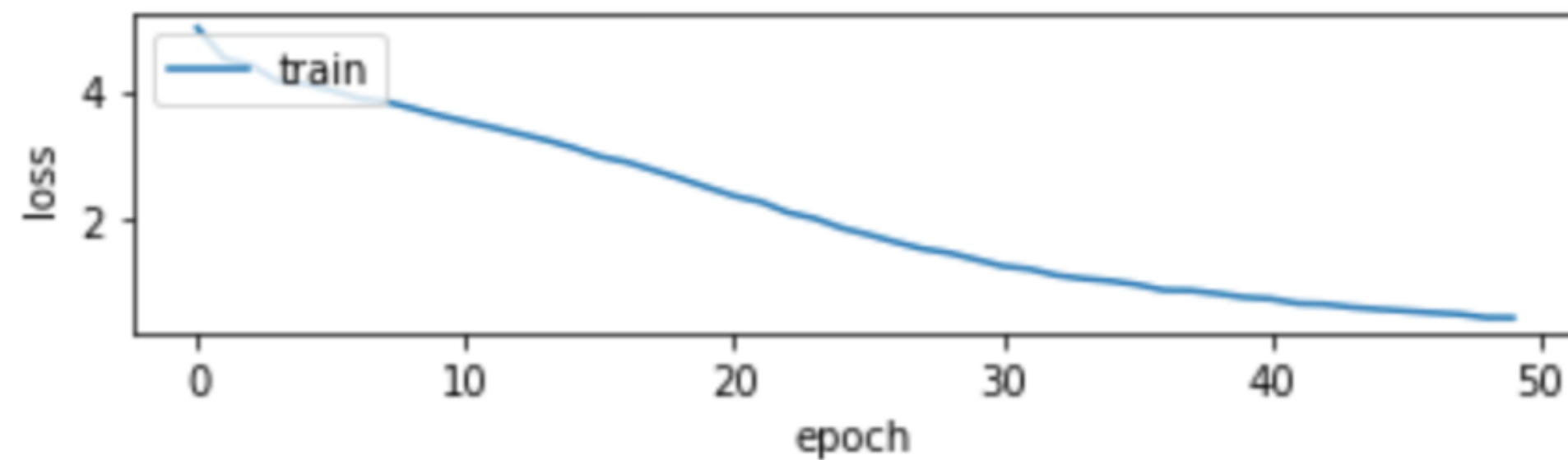
We were able to train our model on 50 epochs

Testing was not achieved

model accuracy



model loss

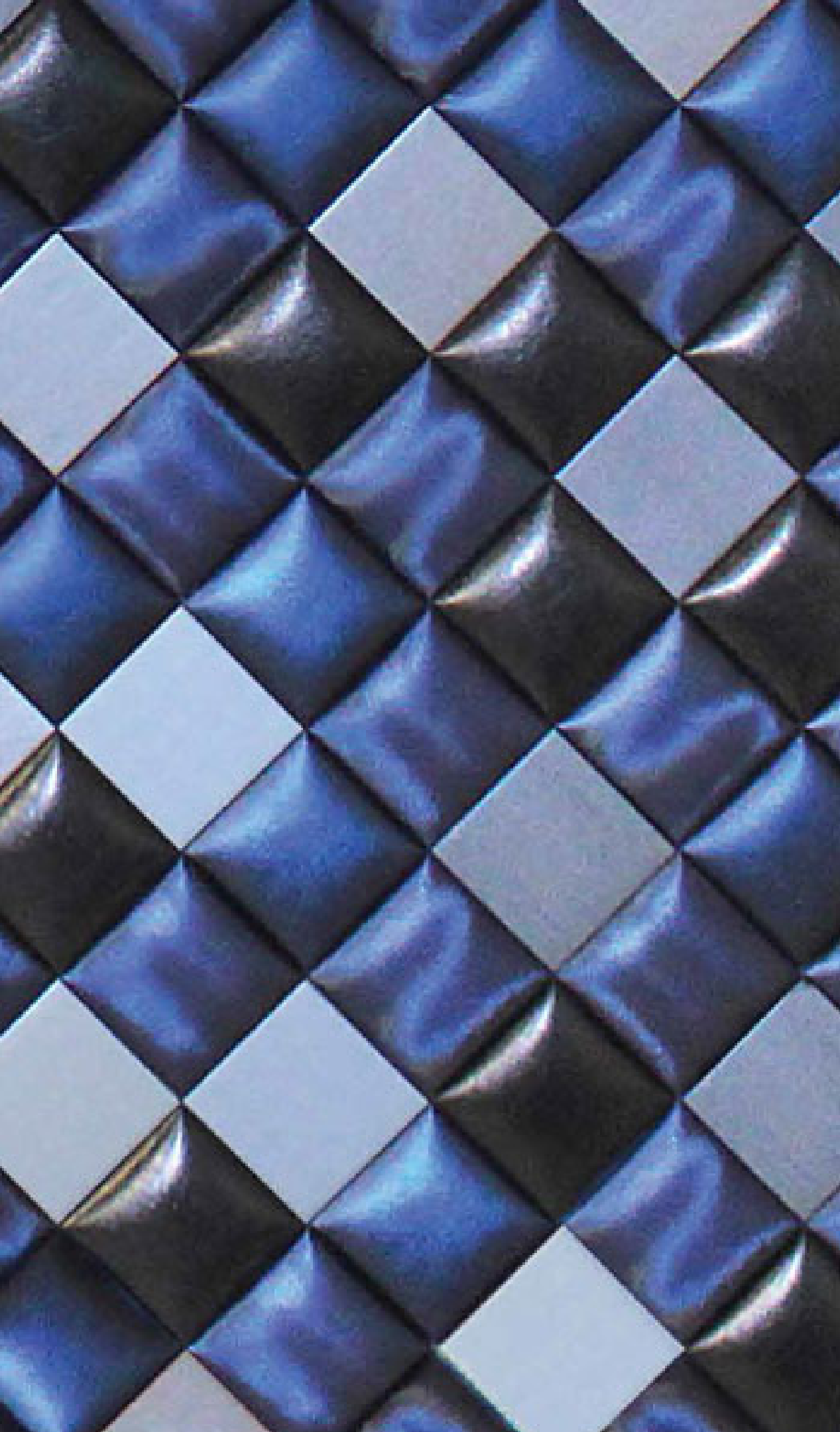


CURRENT GOALS

Continue working on the code to figure out why validation input is not testing

ADDITIONAL GOALS

Revise network and use BLEU Score to test whether the network achieves what we want it to



THANK YOU