

Presented by Group 2

# Urdu (اردو) Handwritten OCR



# Introduction

What is an OCR?

Optical Character Recognition (OCR) technology  
converts:

Documents → Digital text

Handwritten Urdu OCR? Now, that's the challenge!





# Research Question

**URDU Handwritten OCR:**  
**Given a snippet of Urdu handwritten text, convert  
it to digital text**

# Dataset Example

کوں سوچ سکنا تھا، اے بندوستان اکٹریت اور انگریز حکمرانوں کی مشترکہ

Input image

کوں سوچ سکنا تھا کہ بندوستان اکٹریت اور انگریز حکمرانوں کی مشترکہ

Output image

# Collection and Division of Dataset



# Models Used

01

CNN

02

CNN+ Bi-  
directional  
LSTM

03

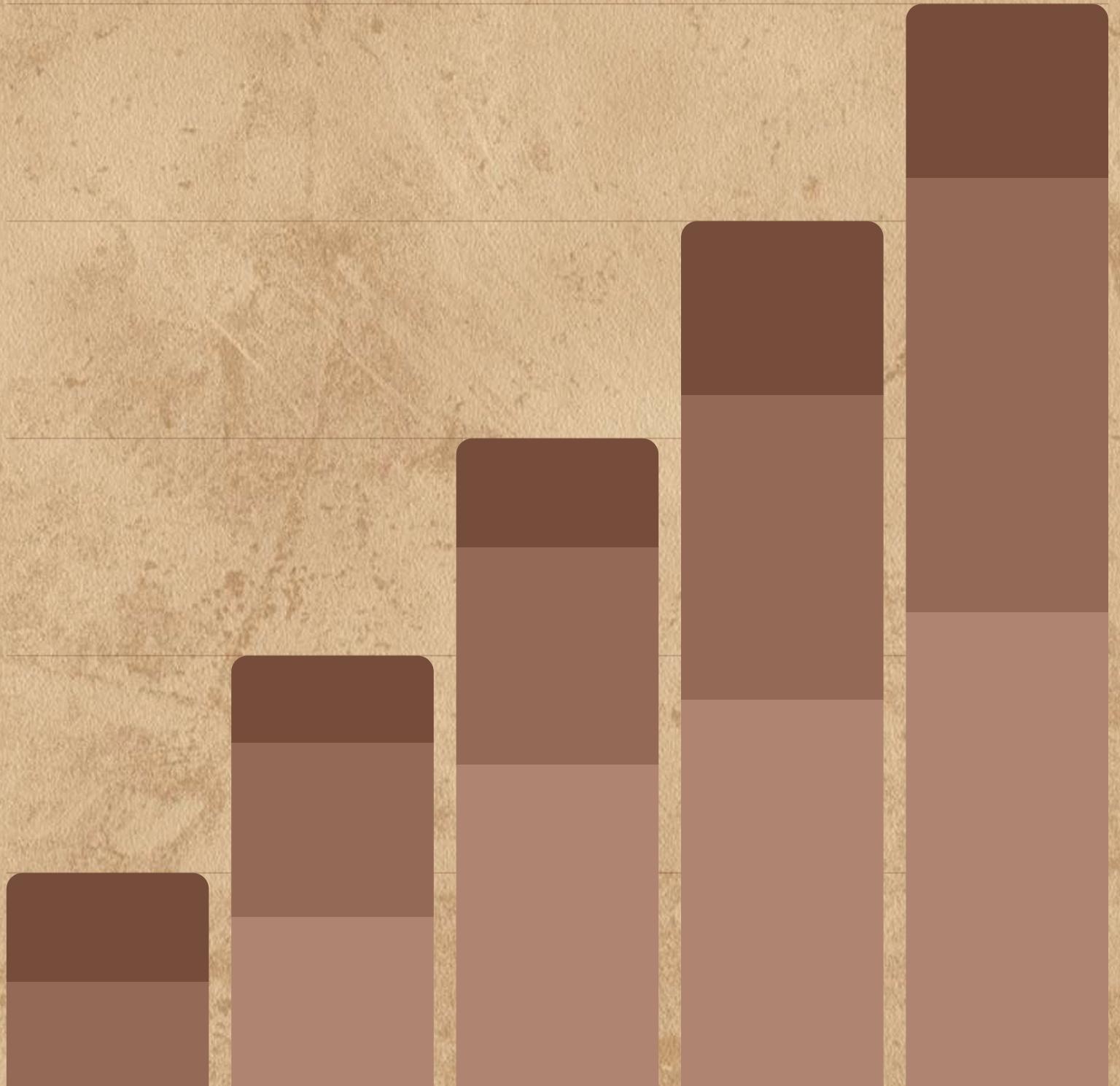
CNN+  
Transformer

# Evaluation Metrics

## CHARACTER ERROR RATE

total number of characters (n), including spaces, to the minimum number of insertions (i), substitutions (s) and deletions (d) of characters that are required to obtain the Ground Truth result. The formula to calculate CER is as follows:

$$\text{CER} = [ (i + s + d) / n ] * 100$$



# CNN Results

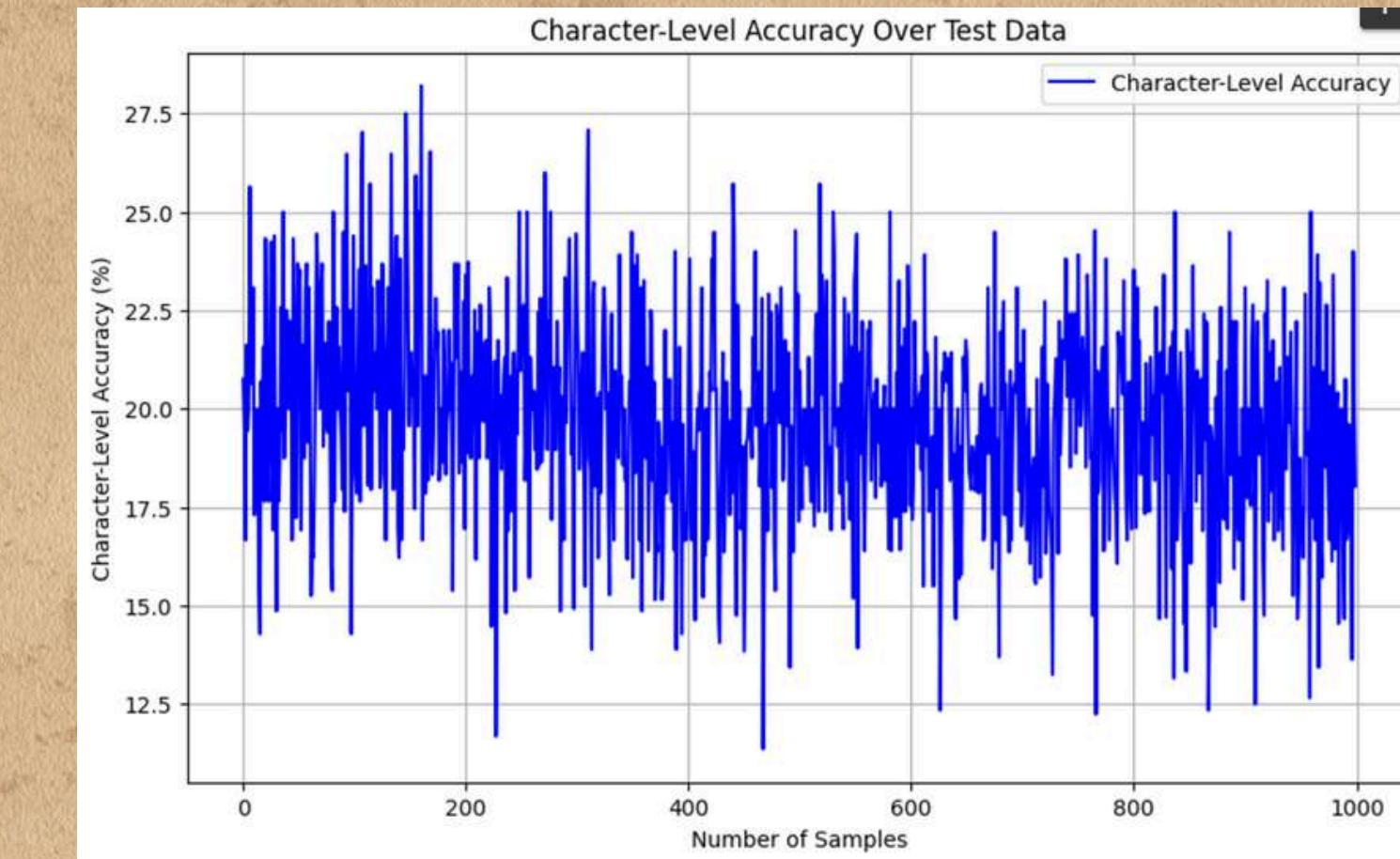
EPOCHS = 50 + 10

BATCH SIZE = 32

LEARNING RATE = 0.0001

OPTIMIZER = Adam

LOSS FUNCTION = CTC LOSS



Training  
Accuracy

40%

Testing  
Accuracy

20%

# CNN + BLSTM Results

EPOCHS = 50 + 30

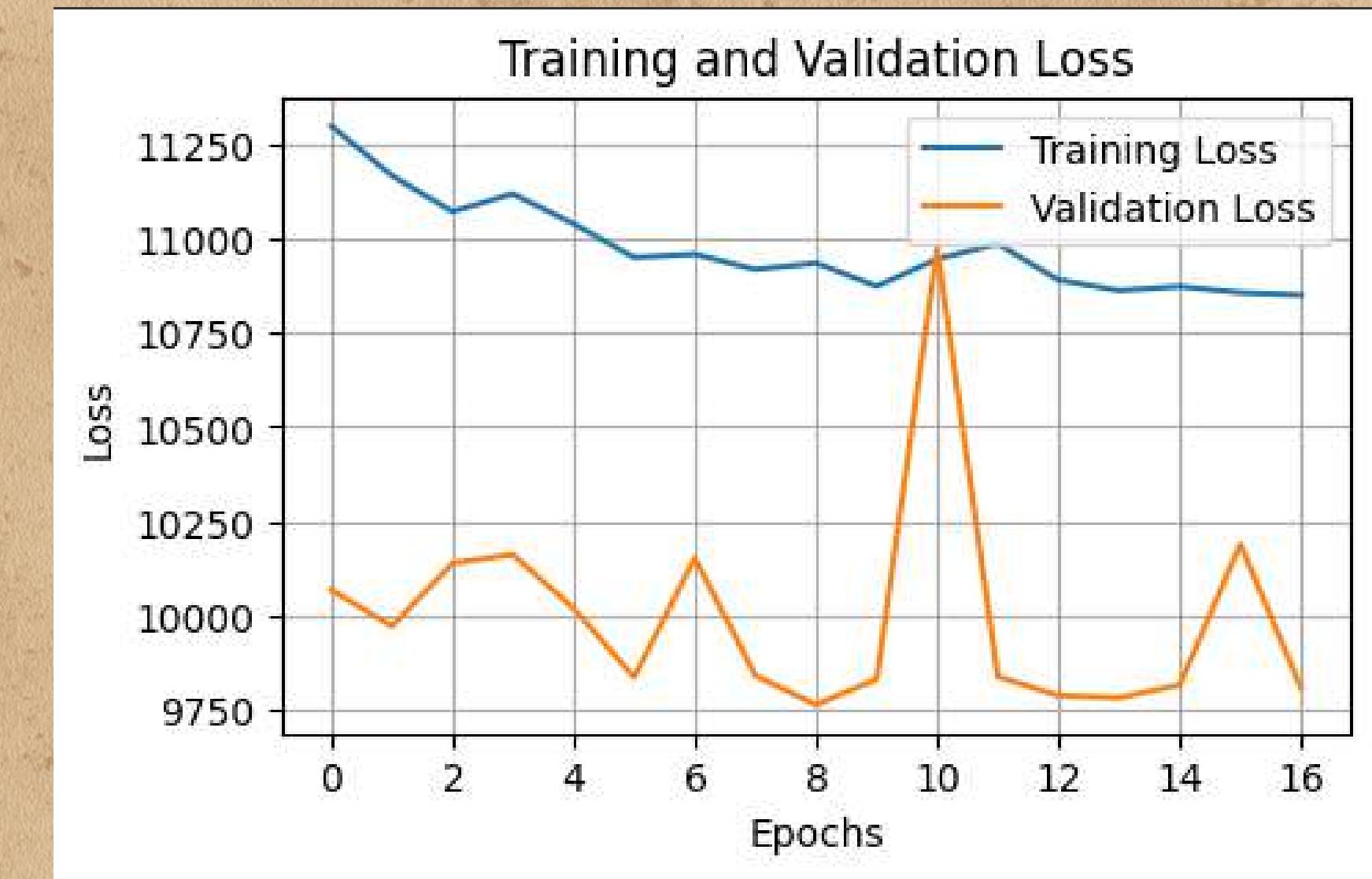
BATCH SIZE = 64

LEARNING RATE = 0.001

OPTIMIZER = EXPONENTIAL DECAY

LOSS FUNCTION = CTC LOSS

EARLY STOPPING PATIENCE = 3



Mean Character  
Accuracy for Training

25%

Mean Character  
Accuracy for Testing

~8%

# CNN + Transformer

EPOCHS = 50 + 30

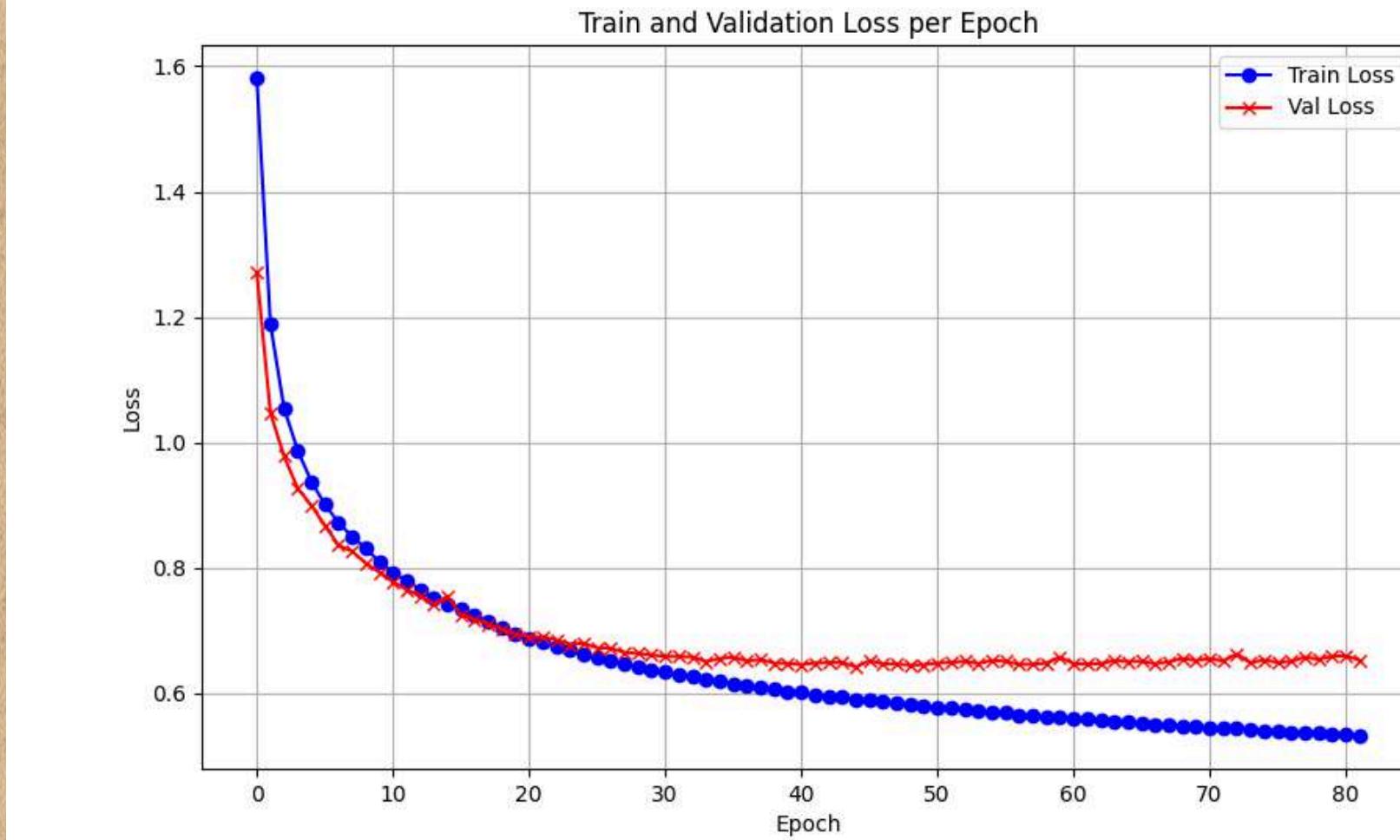
BATCH SIZE = 16

LEARNING RATE = 0.001

OPTIMIZER = Adam

LOSS FUNCTION = Default

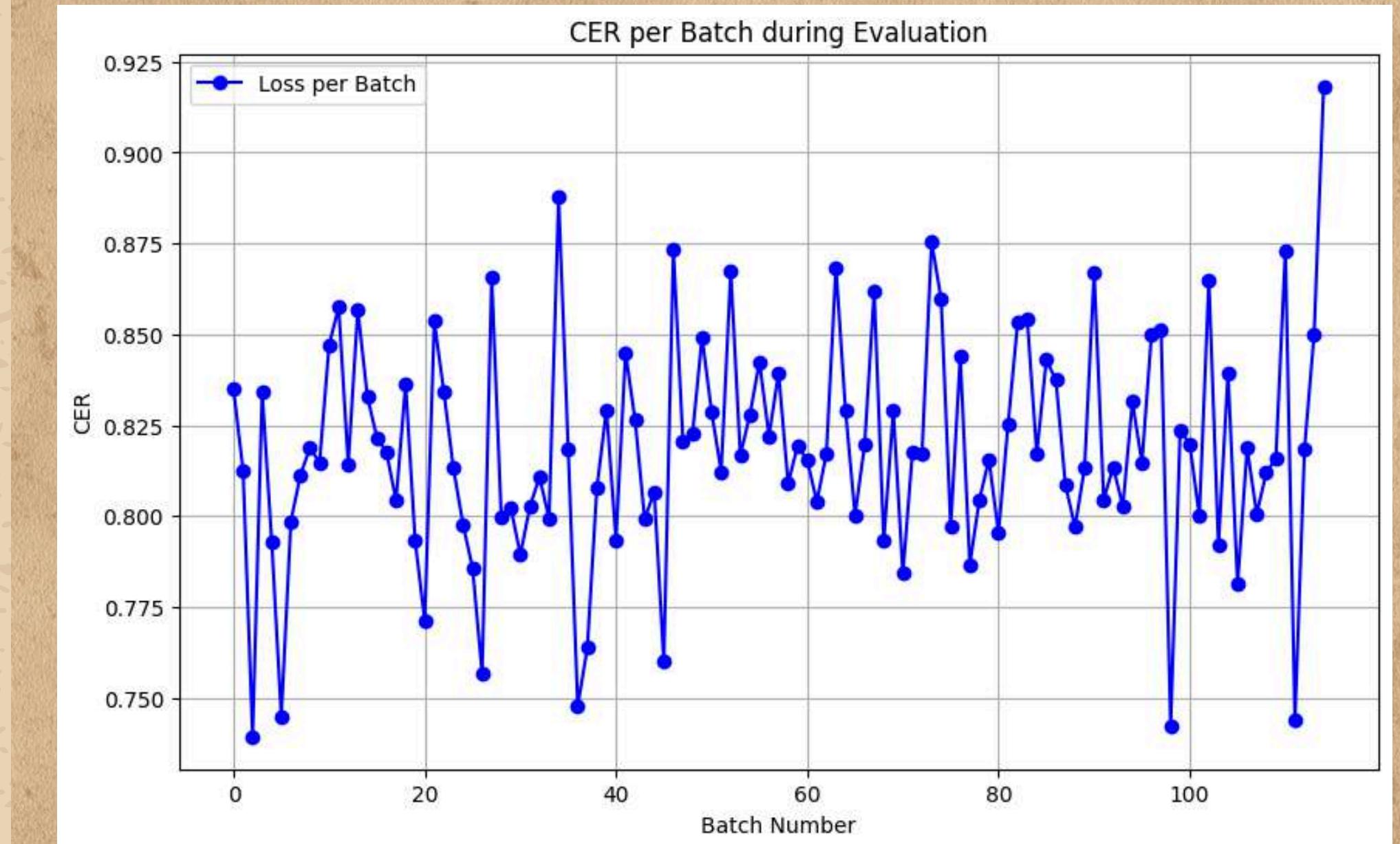
EARLY STOPPING PATIENCE = 3



Training Loss	Training Accuracy	Testing CER
0.5336	80.30%	0.8185

# CNN + Transformer

CER (Character Error Rate),  
calculated per batch.



# CNN + Transformer Model

The model utilizes:

- CNN - 5 convolution layers
- Transformer - encoder (Roberta) and decoder (GPT-2)
- CER - evaluation

# Comparison

	PRIOR RESULTS	OUR RESULTS
CER %	5.9% val, 6.2% test	50%
DATA SETS	8483 (handwritten images)	18,270 (handwritten images)
TYPE OF THE DATASET	The training and test data had variations (some were printed images some were handwritten)	Both Validation and Test images were handwritten

# Future work (Improvements)

01

Get a larger dataset, with more variations

02

Clean Your Dataset, neatly written without any additional noise.

03

Implement Advanced CNNs such as ResNet, DenseNet

04

Try different transformers such as Vision Transformers (ViTs)

# References

- [1]. A. Maqsood, N. Riaz, A. Ul-Hasan, and F. Shafait, "A Unified Architecture for Urdu Printed and Handwritten Text Recognition," National Center of Artificial Intelligence (NCAI), National University of Sciences and Technology (NUST), Islamabad, Pakistan. [Online].
- [2]. M. Pervaiz, "Arabic Handwritten OCR," GitHub repository, 2024. [Online]. Available: <https://github.com/MUZAMMILPERVAIZ/Arabic-Handwritten-OCR/tree/main> . [Accessed: Dec. 2, 2024].
- [3]. H. Bakri, "Handwriting Recognition with AI," GitHub repository, 2023. [Online]. Available: <https://github.com/harithbakri/handwritingrecognitionwithai>. [Accessed: Nov. 26, 2024].



# Thank you very much!

Presented by Group 2