# Recent Achievements of Deep Learning on Recognition of Modern Japanese Magazines

LE DUC ANH　　データ科学研究系　特任助教

Introduction

✚ Since historical documents are an invaluable resource for historians in exploring social aspects, lifestyles, even weather in the previous era, many countries have been preserved their historical documents

✚ Document analysis and recognition can speed up the transcription process.

✚ We improve textline recognition by enlarge small training set by parallel text images (original Kindai and Japanese Font)

✚ We propose a distance based objective function adapt between original Kindai images and Japanese Font images.

System Overview

Input image

Layout analysis
paragraph, line, word

Text Recognition
By Transformer

明治十六年九月再版
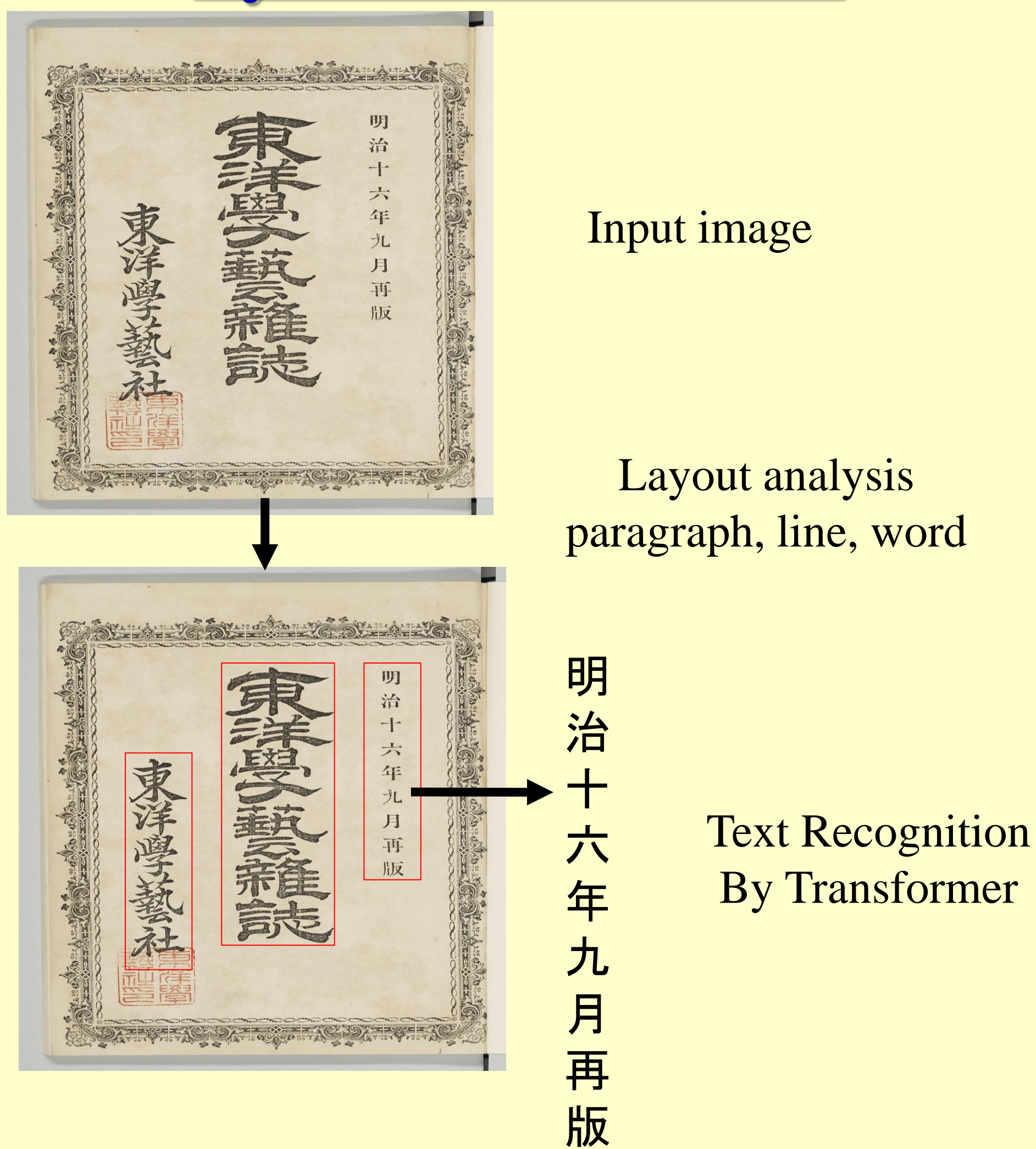
Dataset & results

✚ We employed The dataset from The National Diet Library (3997 pages) for training and Shisou dataset (922 pages) for evaluation.

✚ For parallel text line images, we employ Noto Sans and Noto Serif to generate textline images. As the result, we increase NDL dataset by 2 and 3 times

✚ The number of categories is 5,398 which contains many character categories that do not use in current Japanese character system.

The sizes of datasets to train Transformer OCR

Dataset	# images	# parallel textline images
NDL dataset	103,256	N/A
P1	206,512	103,256
P2	413,024	206,512

Performance of Text Recognition (Character Error Rate)

Method	NDL dataset	P1	P2
Transformer OCR	19.68	19.55	19.59
Transformer OCR with Adaptation	N/A	17.88	17.36

Example of recognition result

https://github.com/ducanh841988/Kindai-OCR

System Architecture

Example of parallel text image
(Left: Kindai textline,
Right: Font textline)

Structure of the Transformer for recognizing Kindai documents
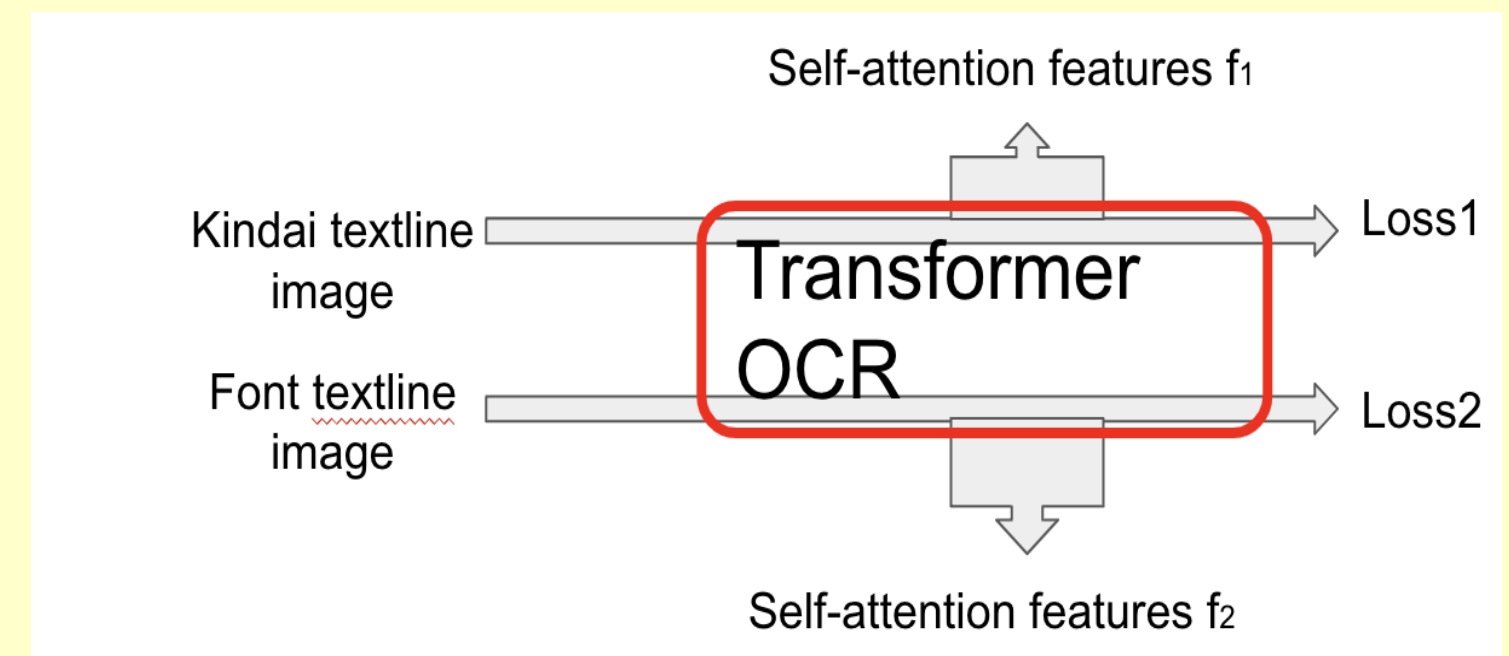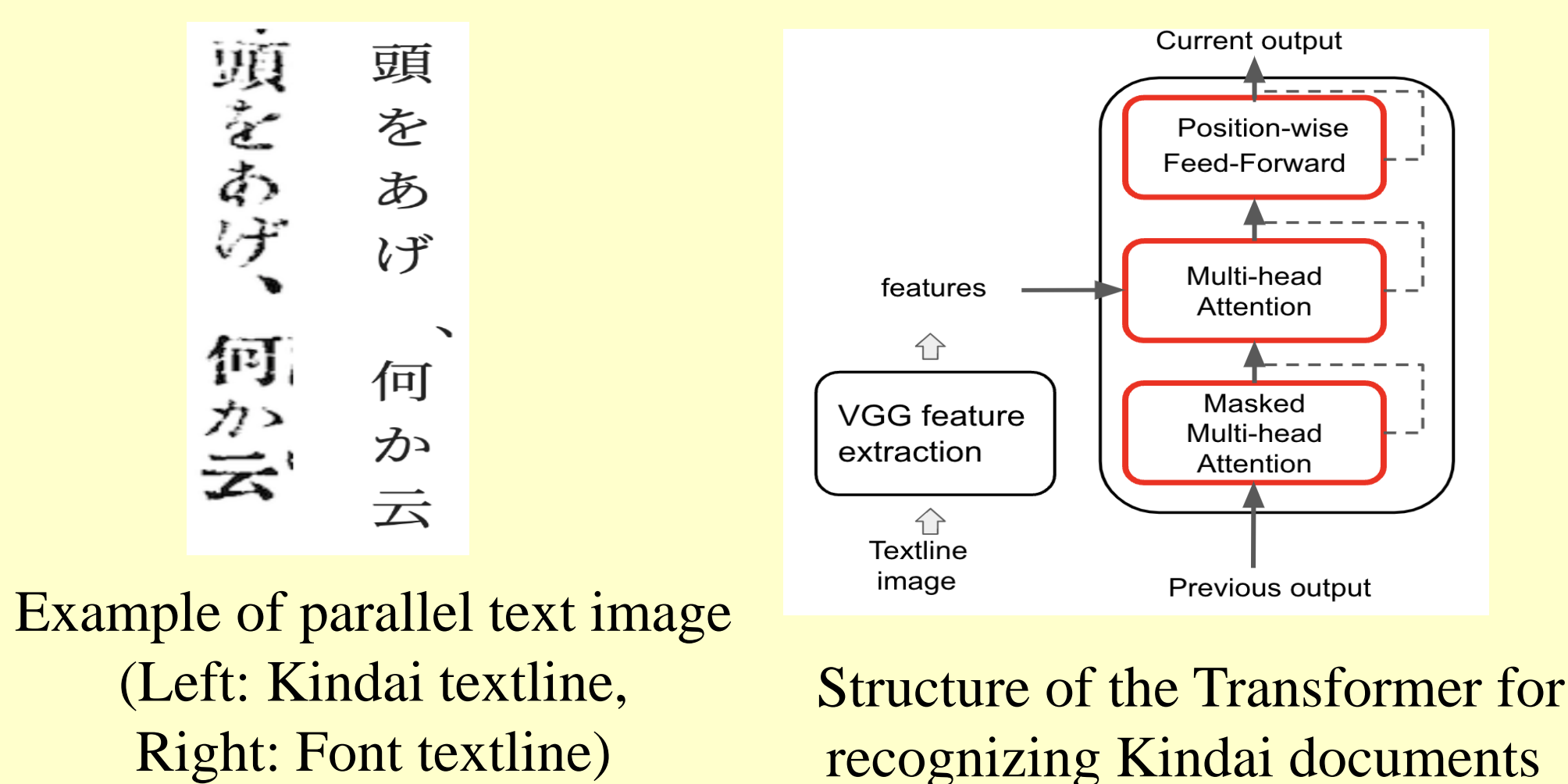
$$loss = Loss_1 + Loss_2 + \alpha * Distance(f_1, f_2)$$

Self-attention feature distance based loss for adaptation