

End-To-End Full-Page Optical Music Recognition for Mensural Notation

23rd International Society for Music Information Retrieval
Bengaluru, India, 2023

Antonio Ríos-Vila, José M. Iñesta, Jorge Calvo-Zaragoza

UI for Computer Research, University of Alicante, Alicante, Spain

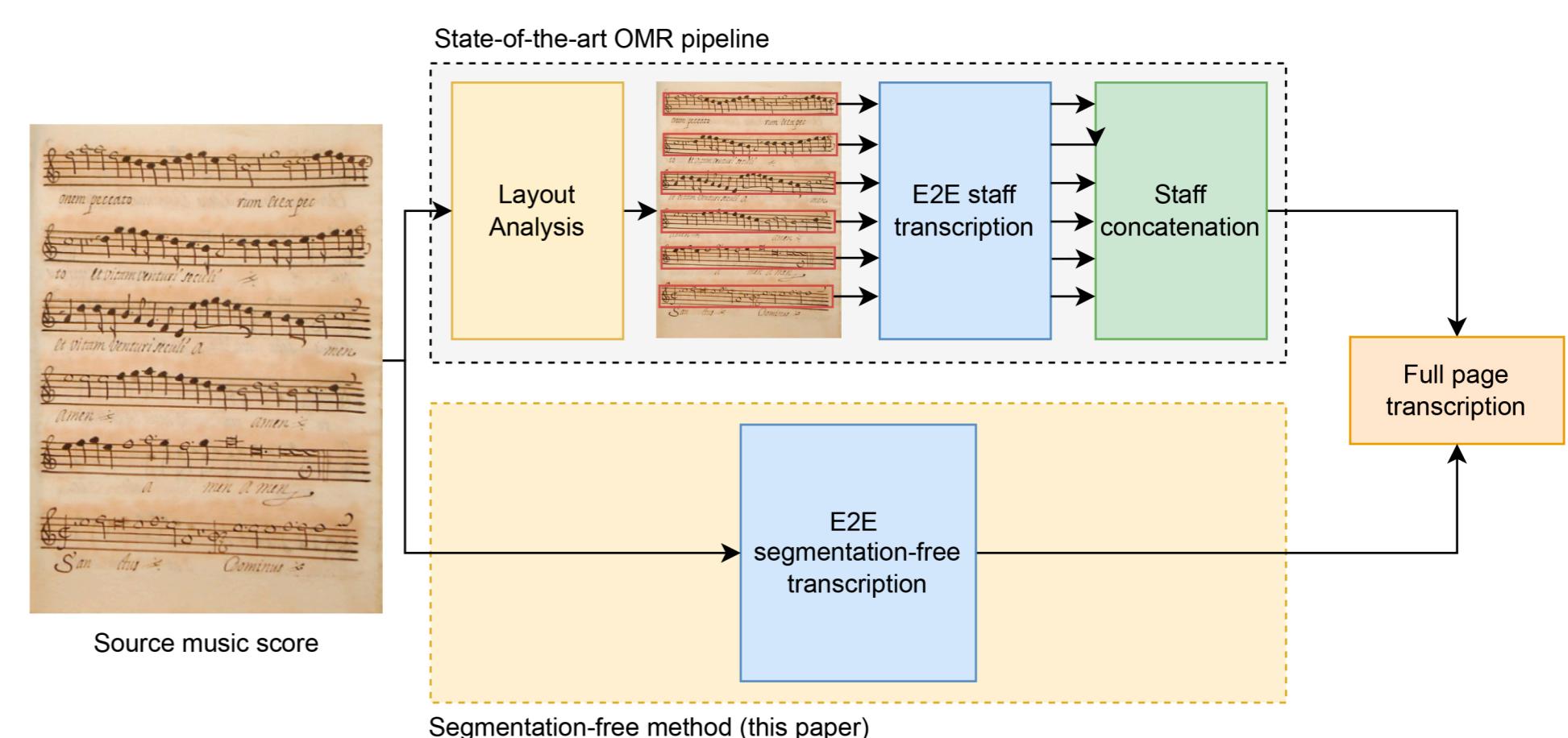
Universitat d'Alacant
Universidad de Alicante

GENERALITAT
VALENCIANA
Conselleria de Innovació,
Universitats, Ciència
y Sociedad Digital



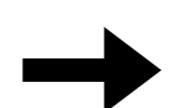
1. Introduction

- Full page OMR relies on **Layout Analysis + Transcription** workflows
- Two independent models trained separately
Assume an accumulated error
- Data labeling:
Define a bounding box (region) for each stave
Write the correspondent transcription for each box
Corpora has to be manually labeled twice
- **Need of joining both models to prevent these inconveniences**

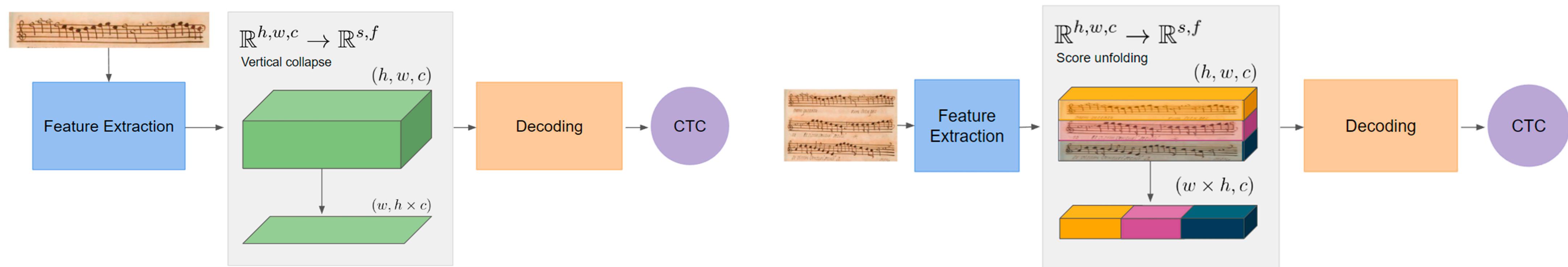


2. Extending End-to-End OMR Transcription Systems

Single-stave transcription system



Full page adaptation



3. Experimental setup

Corpora

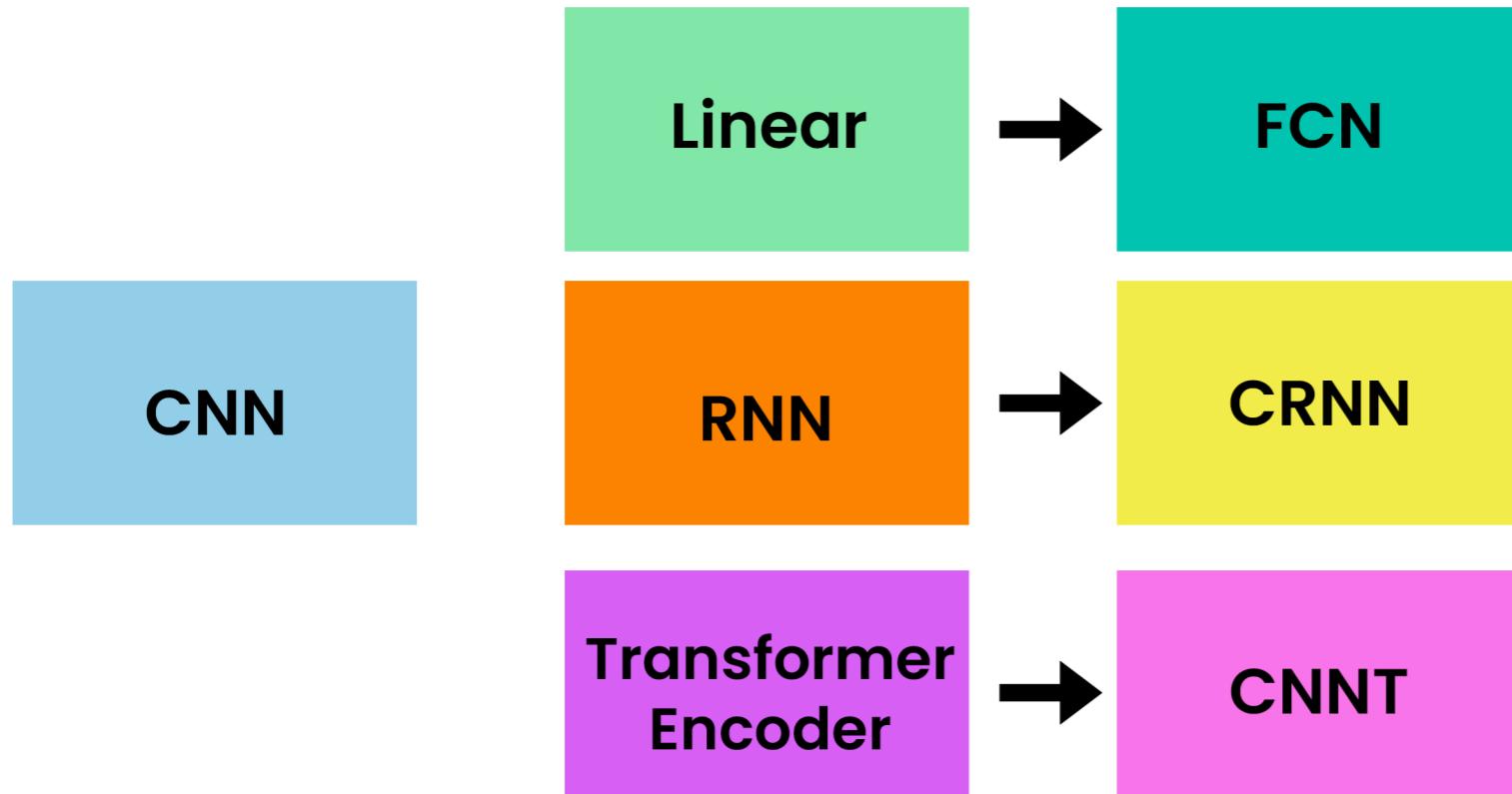


SEILS



CAPITAN

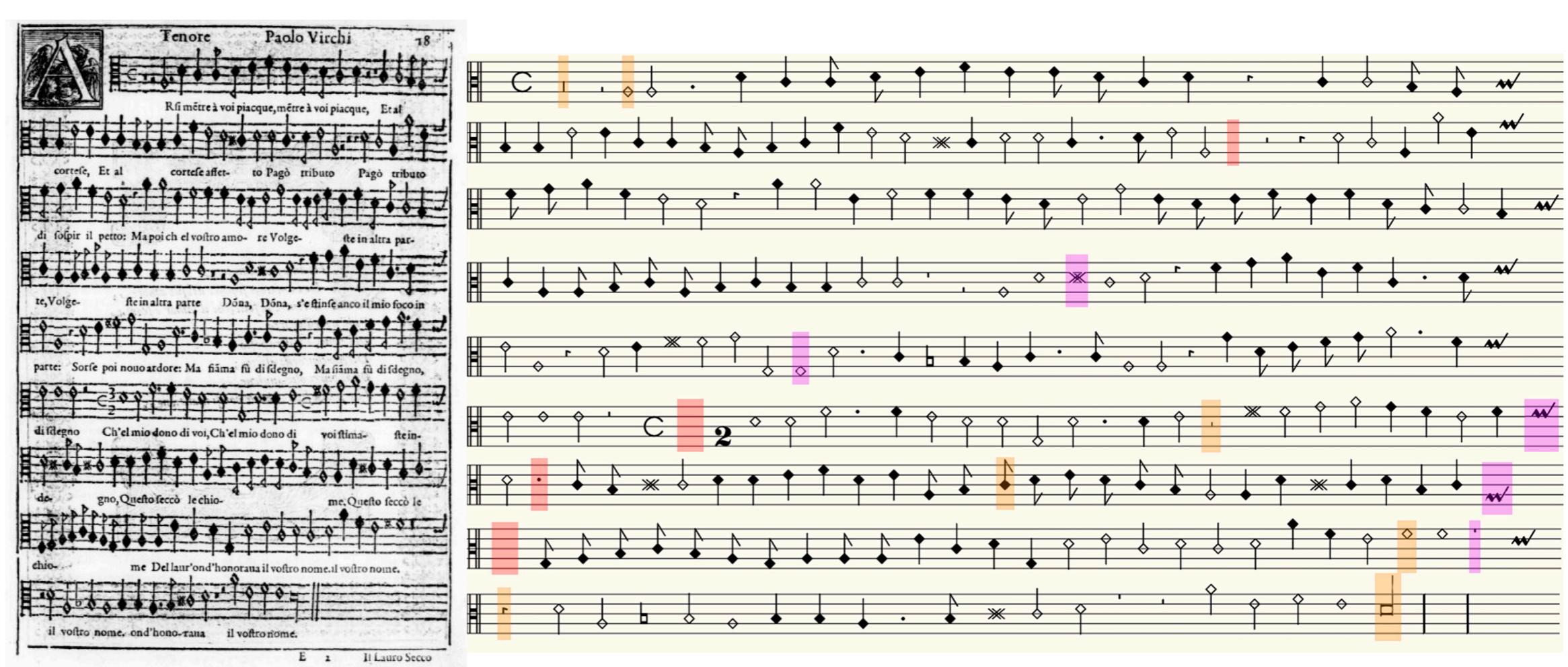
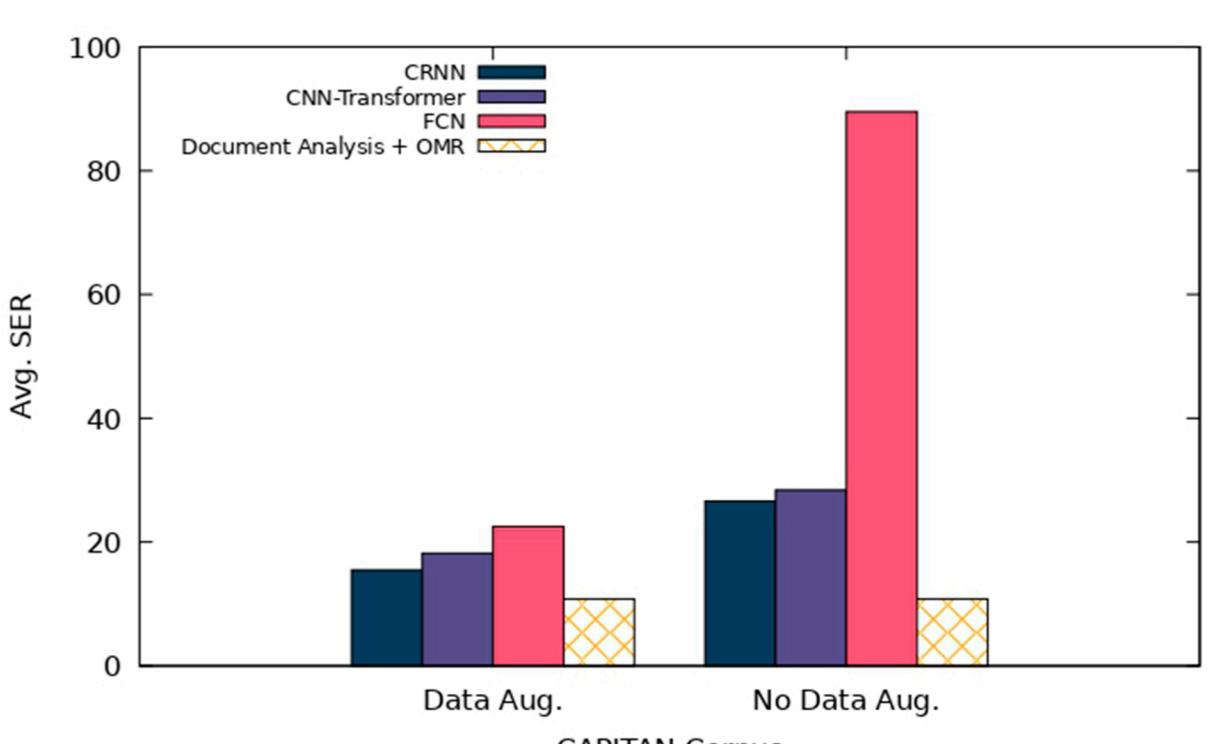
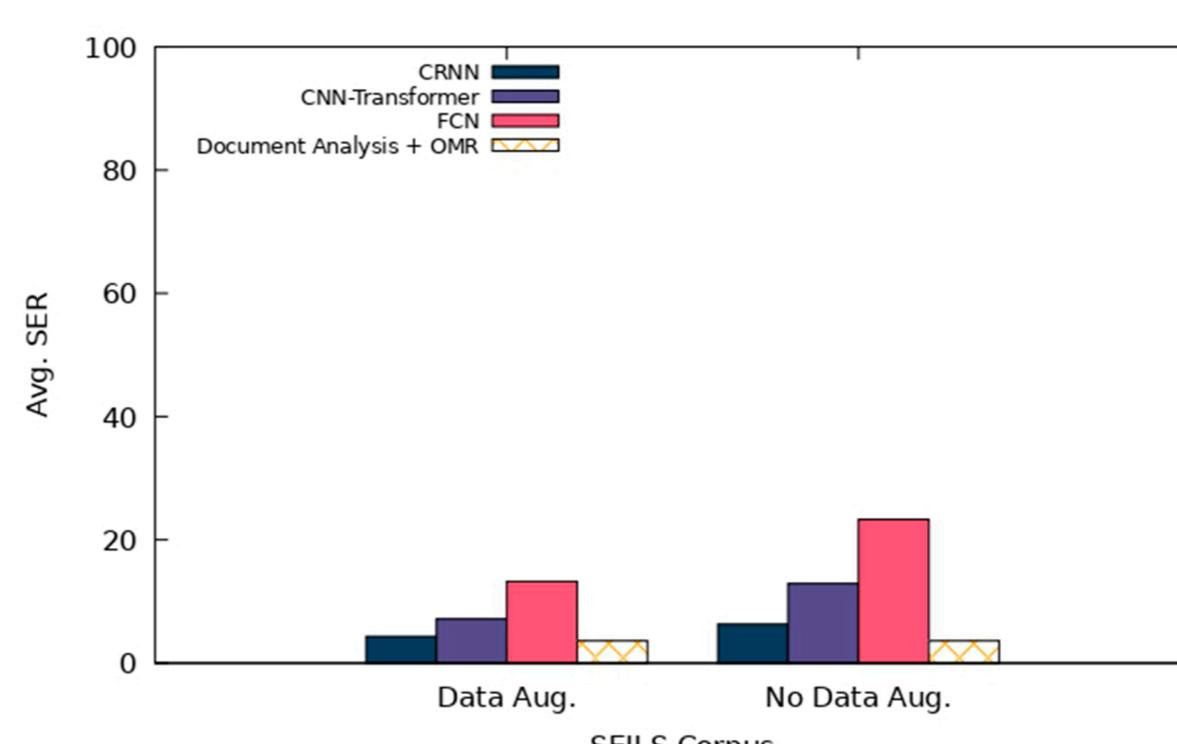
Neural Implementation



Metric

$$\text{Symbol Error Rate (SER)} = \frac{\text{Edit Distance } (h, t)}{\text{Length}(t)}$$

4. Results



5. Conclusions

- Presented a **layout analysis-free** E2E transcription system
- **Competitive results** against the state-of-the-art models with less labeling cost
- Sequence processing improves the model performance
- Model **limited to monophonic scores**
- Dependency on data amount (study of Transfer learning and SSL)