MUSIC READER/PLAYER (TRANSCRIBER)

PRIYANK KOUL (PES2201800381)
ISHAAN SAMANT (PES2201800036)
SHEVGOOR ADITHYA KAMATH (PES2201800023)
ARIF SIDIQ WANI (PES2201800260)

INTRODUCTION

The fields of machine learning and deep learning have undergone enormous transformations in the past few years and have brought about many useful applications in many areas. One area of interest is Optical Music Recognition (OMR). According to Wikipedia, OMR is a field of research that investigates how to computationally read music notation in documents. **The goal of OMR is to teach the computer to read and interpret sheet music and produce a machine-readable version of the written music score.**

SCOPE

- Digital sheet music is handy for professional musicians.
- Instead of hauling reams of paper around, all they need is to digitize machine-readable music.
- The key is an artificial neural network that uses deep learning methods to teach itself how to read sheet music and determine its content.

LITERATURE SURVEY
TAKING UP THE TOPIC

PUBLICATIONS/PAPERS

- https://web.stanford.edu/class/ee368/Project_Spring_1415/Reports/Stramer.pdf
- https://arxiv.org/pdf/1908.03608.pdf
- https://www.mdpi.com/2076-3417/9/13/2645/pdf
- https://www.mdpi.com/2076-3417/8/4/606

BLOGS/ARTICLES

- https://towardsdatascience.com/i-built-a-music-sheet-transcriber-heres-how-74708fe7c04c
- https://www.naratek.com/en/artificial-intelligence/2019/Using-deep-learning-to-create-a-digital-sheet-music-stand.php

LITERATURE SURVEY
INFERENCES

- One of the previously mentioned papers analyzed how OMR inverts the music encoding process to recover the musical notation and the musical semantics from documents. Additionally, it was discussed how deep learning affects modern OMR research, as opposed to the traditional pipeline.
- In recent years, with the development of computer vision technology, the convolutional neural network (CNN) based on deep learning technology has been increasingly applied to OCR and has achieved good results. So, the object-detection based approaches were prevalent initially.
- Other approaches used sequence-to-sequence architecture to translate OMR issues into translation problems. In the absence of context, instead of training a single segmentation symbol, the entire line of music is translated simultaneously. These approaches mainly use convolutional neural networks (CNN) and recurrent neural networks (RNN).

DATASET

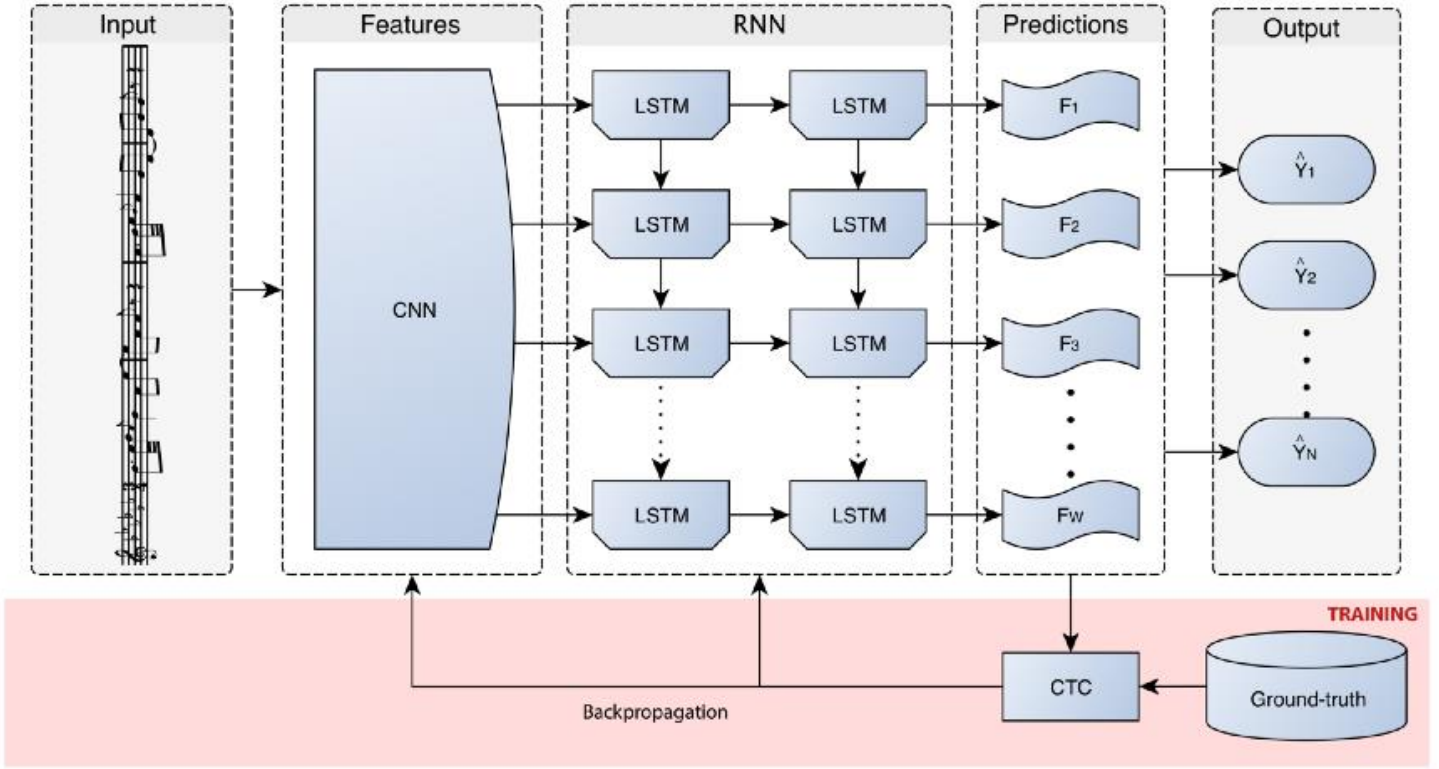
- The dataset that we intend to use for this project is the Printed Images of Music Staves (PrIMuS) dataset.

- It contains 87678 real music incipits which are represented by five files:Plaine and Easie code , image with rendered score, musical symbolic representation of incipit in Music Encoding Initiative format (MEI) and semantic encoding and finally agnostic encoded files.

- The semantic encoding is an on purpose simplified encoding.

- The agnostic encoding is a sequence containing graphical symbols with their position in the staff without their Musical meaning.

GRAPHICAL SCHEME/ARCHITECTURE



SOFTWARE FRAMEWORKS

- Python
- Tensorflow
- Convolution Recurrent Neural Networks
- Connectionist Temporal Classification loss function(CTC Learning)
- Semantic and Agnostic model (for different notes)
- Bidirectional Long Short Term Memory (BLSTM)

MINIMUM HARDWARE REQUIREMENTS

- RAM:- 8GB and above.
- CPU:- i5 8th generation / Ryzen 3000 series and above.
- GPU:- NVIDIA MX150 4GB.

  (It takes around 1 second per score in a general-purpose computer like an Intel Core i5-2400 CPU at 3.10 GHz with 4GB of RAM, and without speeding-up the computation with GPUs.)

DELIVERABLES

- Model that performs OMR on digital images of a musical sheet.
- Feature extraction of the musical notes from the image.
- Model that outputs the sequence of a given set of notes.
- Translating musical scores on monophonic notation to ABC scores .

THANK YOU