IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

LITERATURE SURVEY ON OPTICAL MUSIC RECOGNITION USING MACHINE LEARNING

[1]Priyanka Chalak, [2]Sonal Fatangare, [3]Ashwini Bansode, [4]Ganesh Bhagat, [5]Vaishali Badde

[1] Student, [2] Asst. Prof, [3] Student, [4] Student, [5] Student

[1] BE Computer Engineering,

[1]RMDSSOE, Pune, India

Abstract: Sheet music transcription into a machine-readable format is the focus of optical music recognition (OMR). OMR helps with use cases in digital musicology by making it easier to analyze sheet music statistically or look for notational trends. After a musical composition has been visually expressed in a document with music notation, OMR begins. Recently, OMR has moved away from using traditional computer vision methods and toward a machine learning strategy. In this work, we address these flaws by: Offering a comprehensive definition of OMR and its relationship to related fields; Examining how OMR inverts the music encoding process to recover the musical notation and the musical semantics from documents; and proposing a taxonomy of OMR, with a novel taxonomy of applications being the most salient feature.

Index Terms **- Optical Music Recognition, Music Notation, Music Scores, Neural Network, Image Processing, Preprocessing, Optical Character Recognition.**

I. INTRODUCTION

The study of computationally reading musical notation in written texts is known as optical music recognition. A structure of notes in time can be thought of as music. This is the conceptualization we take into account for OMR even if it isn't the only one. While there are other ways to conceptualize music, this is the only one with a reliable, widely acknowledged visual language used to transmit it in writing. A note is a musical object that is defined by four parameters: pitch, duration, loudness, and timbre. Additionally, it has an onset: a placement onto the axis of time, which in music does not mean wall-clock time, but is measured in relative units called beats.3 Periods of musical time during which no note is supposed to be played are marked by rests, which only have an onset and a duration.

Optical music recognition software that provides information about the music notation document is referred to as document metadata extraction. Recently, a paradigm change in OMR has also been brought about by deep learning's effectiveness in enhancing text and speech recognition. This position paper tries to put one of these strategies into practice.

II. LITERATURE SURVEY

In 2018 Sanu Pulimootil Achankunju, et al [1] proposed system using optical music recognition (OMR) technology, which converts printed or sheet music into a machine-readable format, they collected the symbolic note information from the whole works of four well-known composers. The OMR data have a lot of extraction problems and are fairly noisy. To enable melody search on this noisy material and still achieve very high retrieval quality, developed a music search engine. also discuss the results of our tests using externally generated musical themes to evaluate the performance of our search engine.

Julia Adamska, et al [2] The server's responsibility is to extract audio from the image, create an mp3 file, and deliver it to the client programmer. The system's most important component is the programmer that implements the algorithm for note recognition. It is based on decision trees and the traits of the many symbols that were taken out of the image. The system utilizes Microsoft Azure, a cloud operating system, and is built on the Windows Phone 8 architecture. It makes it simple to archive images, notes that are recognized in the Music XML format, and produced mp3 files.

The server's responsibility is to extract audio from the image, create an mp3 file, and deliver it to the client programmed. The system's most important component is the programmed that implements the algorithm for note recognition. It is based on decision trees and the traits of the many symbols that were taken out of the image. The system utilizes Microsoft Azure, a cloud operating system, and is built on the Windows Phone 8 architecture. It makes it simple to archive images, notes that are recognized in the Music XML format, and produced mp3 files.[3]

In 2015 Apurva A. Mehta et al [4] are exposed to music in a variety of ways, including audible, visual, and another, less well-known format: written music. In some ways, music rules our lives. The computer system under discussion accepts images of piano music scores created using contemporary staff notation. Utilizing thresholding, stave lines, and hierarchical decomposition, score sheets are segmented. An established artificial neural network based on boosting method is used to recognize segmented symbols.

Kia Ng et al [5] are exposed to music in a variety of ways, including audible, visual, and another, less well-known format: written music. In some ways, music rules our lives. The computer system under discussion in this paper accepts images of piano music scores created using contemporary staff notation. Utilizing thresholding, stave lines, and hierarchical decomposition, score sheets are segmented. An established artificial neural network based on boosting method is used to recognize segmented symbols.

Music scanned from paper is converted using optical music recognition (OMR) technologies into a format that can be played or edited on a computer. The two steps of these systems are typically the recognition of the graphical symbols (such as note-heads and lines) and the musical interpretation and linkages of the symbols (such as the pitch and rhythm of the notes). It examines the second stage and provide a two-step methodology that allows for a cost-effective representation of the system's parsing rules.[6]

In 2014 Jorge Calvo-Zaragoza et al [7] There are also a number of recognition options for the two modalities—online, which uses the pen strokes—and offline, which uses the image created after drawing the symbol. To reach major conclusions about the recognition of these data, some experiments are included. It is anticipated that this study will serve as a reference point for subsequent work in the area of online handwritten music notation recognition.

Jorge Calvo-Zaragoza et al [8] Over less-than-ideal visual circumstances of musical scores, assess the performance of an end-to-end method that makes use of a deep convolutional recurrent neural network (CRNN). Therefore, another component of our contribution is Camera-PrIMuS, a collection of printed monophonic music scores that have been artificially altered to resemble camera-based realistic scenes, including aberrations like distorted illumination, rotation, and blur. Our findings demonstrate that the CRNN is capable of completing the task under these circumstances, getting an error around 2% at the music-symbol level, and so proving to be a ground-breaking contribution to the field of OMR systems.

Investigation of the application of end-to-end neural networks. This is accomplished by utilizing a neural model that combines the strengths of recurrent neural networks, which address the sequential aspect of the issue, with convolutional neural networks, which work on the input image. These models can be immediately trained from input photos and their accompanying transcripts into music symbol sequences by using the so-called Connectionist Temporal Classification loss function.[9]

In 2017 Jorge Calvo-Zaragoza et al [10] Investigation the application of end-to-end neural networks. This is accomplished by utilizing a neural model that combines the strengths of recurrent neural networks, which address the sequential aspect of the issue, with convolutional neural networks, which work on the input image. These models can be immediately trained from input photos and their accompanying transcripts into music symbol sequences by using the so-called Connectionist Temporal Classification loss function.

Discussion of efforts to convert printed music notation images into symbolic representations that can be used for playing, analysis, and rendering in the field of optical music recognition. There is a small core of symbols and primitives used in music notation that are composed in a rule-bound manner, but there are also numerous common deviations to these rules and a long tail of uncommon symbols.[11]

In 2018 Liang Chen et al [12] convert printed music notation images into symbolic representations that can be used for playing, analysis, and rendering in the field of optical music recognition. There is a small core of symbols and primitives used in music notation that are composed in a rule-bound manner, but there are also numerous common deviations to these rules and a long tail of uncommon symbols.

Arindam Chowdhury et al [13] The Lester S. Levy Collection of Sheet Music will be digitally transformed as part of the project (Milton S. Eisenhower Library, Johns Hopkins University). Phase One involved digitizing and building a database of text index records for the Levy Collection's music, lyrics, and color photographs of the album covers. In phase two, the digital music is transformed into computer-readable music notation format along with full-text lyrics, sound renditions are produced, and metadata is created to improve search possibilities.

Use of neural network-based cross-modality embedding spaces to tackle the following two sheet music-related problems: finding the right piece of sheet music from a database given a piece of music audio as a search query, and matching an audio recording of a piece with the corresponding images of sheet music.[14]

In 2004 Matthew J. Dovey [15] The mission of the three-year Online Music Retrieval and Searching (OMRAS) project, which started in 1999 and was supported by the JISC/NSF International Digital Library Initiative, was to look into the problems associated with retrieving information about polyphonic music. The accomplishments of OMRAS in pattern matching, document retrieval, and voice transcription are described here, along with some prototype work on how to integrate these approaches into library systems.

Hoda M. Fahmy et al [16] The Build-Weed-Incorporate programming approach for graph grammars is introduced, along with examples of how it can be used to interpret complicated diagrams where the interaction of spatially dispersed symbols plays a significant semantic role. Symbol recognition and high-level recognition are the two stages of diagram recognition. In-depth research has been done on symbol recognition in the literature. In this study, we assume a symbol recognizer exists, and use of graph grammar to piece together the information content of the diagram from the symbols and their spatial relationships.

novel method for matching musically comparable audio clips to scanned pages of sheet music and mapping them to a set of audio recordings. optical music recognition (OMR) and digital signal processing techniques are used to first convert the scanned images and audio recordings into a common feature representation. A direct comparison of the two different forms of data is made easier because to this shared representation. This makes it possible to search the audio library using scan-based searches.[17]

In 2003 Susan E. George [18] novel method for matching musically comparable audio clips to scanned pages of sheet music and mapping them to a set of audio recordings. optical music recognition (OMR) and digital signal processing techniques are used to first convert the scanned images and audio recordings into a common feature representation. A direct comparison of the two different forms of data is made easier because to this shared representation.

Online and Offline Visual Perception of Music Notation Recognition discusses how music notation is interpreted and used by computers in a variety of application scenarios. It covers work on representation languages, web-based applications, and image processing and pen-based computer research. The achievements, difficulties, and issues brought up by the computer's interpretation of this musical notation language are collected in this book.[19]

In 2007 Angelos P. Giotis et al [20] The rising number of options over the past ten years reflects the increased interest in employing word spotting to address the problem of document indexing. However, there are very few in-depth investigations that examine the different components of a word detection system. With regard to the connected works, this intends to review current approaches and fill in the gaps in a number of issues. The types of texts and underlying difficulties that word detection techniques handle is thoroughly discussed. usage of retrieval improvement strategies based on relevance feedback which increase the retrieved results after presenting the key components of a word spotting system.

III. ALGORITHMIC SURVEY

table 1: algorithmic survey

Sr. No.	Paper Title	Methodologies used and Performance Metrics	Advantages and Disadvantages
1.	Music Search Engine from Noisy OMR Data [1]	A new melody search engine that searches the noisy OMR data.	The result shows that the retrieval quality is good even though the score inputs were corrupted by the OMR process.
2	Mobile System for Optical Music Recognition and Music Sound Generation [2]	A REST API ensures transparent communication with the mobile application.	Process multiple music tracks
3	An Integrated Grammar-based Approach for Mathematical Expression Recognition [3]	Parsing Algorithm is used: Using this algorithm, we compute the most likely parse tree according to the proposed model.	statistical model for online handwritten math expression recognition. Disadvantages: restricted with spatial and geometric
4	Optical Music Notes Recognition for Printed Piano Music Score Sheet [4]	Algorithm is used: Recognition process is implemented using boosting based neural network classifier and compared to without using feature extraction method and along with the feature extraction method (i.e., PCA using SVD).	Pitch identification is carried out for polyphony music so it can also work for monophony music. Disadvantages: System cannot work for more diverse score sheets and instruments.
5	Big Data Optical Music Recognition with Multi Images and Multi Recognizers [5]	Big Data approach to harnessing datasets by aligning and combining the results of multiple versions of the same score, processed with multiple technologies.	Improve the accuracy of musical score digitization.
6	A Music Notation Construction Engine for Optical Music Recognition [6]	Models such as Fast R-CNNs, Faster R-CNNs, Single Shot Detectors (SSD) were used to detect musical objects	musical object detection, musical symbol reconstruction-on and encoding the musical knowledge into a machine-readable file in OMR pipeline
7	Recognition of Pen-Based Music Notation [7]	deep learning methods and end-to-end learning	Handwritten sheet music dataset MUSCIMA Can be used

8	Neural end-to-end Optical Music Recognition on Realistic Monophonic Scores [8]	CNN, RNN, CRNN's are used in this Neural end-to-end OMR	Main aim of this organization is to catalog the location of musical sources
9	End-to-end Neural Optical Music Recognition of Monophonic Scores [9]	CNN, RNN, CRNN's are used in this Neural end-to-end OMR	Main aim of this organization is to catalog the location of musical sources
10	Handwritten Music Recognition for Mensural Notation [10]	Develop the machine learning, holistic 2paradigm and propose a system based on deep Convolutional and Recurrent Neural Networks	Consider a neural approach based on CRNN, for the task of HMR in Mensural notation
11	Human-guided Recognition of Music Score Images [11]	We compare our system with the commercial OMR system, Smart-Score, in terms of both efficiency and accuracy	Define a simple communication channel between the user and recognition engine
12	Optical Music Recognition and Human-in-the-loop Computation [12]	Identify error add pixels/models Constraints	create symbolic representations with accuracy near that of published music scores
13	An Efficient end-to-end Neural Model for Handwritten Text Recognition [13]	Deep convolutional network with a recurrent Encoder-Decoder network to map an image to a sequence of characters corresponding to the text present in the image	Offline handwritten text recognition from images is an important problem for enterprises attempting to digitize large volumes of hand-marked scanned documents/reports.
14	Learning Audio–Sheet Music Correspondences for Cross-Modal Retrieval and Piece Identification [14]	CNN to learn correspondences directly between images of sheet music and their respective audio counterparts	Matching musical audio directly to sheet music, without any higher-level abstract representation
15	Overview of The OMRAS Project: Online Music Retrieval and Searching [15]	prototype work in how to implement these techniques into library systems	OMRAS has achieved in pattern matching, document retrieval, and audio transcription.
16	A Graph grammar programming style for recognition of music notation [16]	Develop general methods for the construction of robust graph grammars	for solving picture processing problems
17	Automatic Mapping of Scanned Sheet Music to Audio Recordings [17]	using the OMR extraction results instead of the "clean" MIDI data	Solve the problem of mapping sheet music to audio recordings
18	Online Pen-Based Recognition Of Music Notation with Artificial Neural Networks [18]	Work on computer assisted notation-based music data entry	Handwritten sheet music dataset MUSCIMA Can be used
19	Wavelets For Dealing with Super-Imposed Objects in Recognition of Music Notation [19]	Method of image filtering using two-dimensional wavelets to separate the super-imposed objects	Evaluate how beneficial wavelet image filtering might be to the OMR process
20	A Survey of Document Image Word Spotting Techniques [20]	Markov models (HMMs), conditional random fields (CRFs), neural networks (NNs)	Word spotting for indexing documents available all over the world, written in various scripts or fonts

IV. PURPOSED WORK

The goal of optical music recognition is to convert sheet music into a machine-readable format (OMR). By making it simpler to analyze sheet music statistically or look for notational trends, OMR supports use cases in digital musicology.

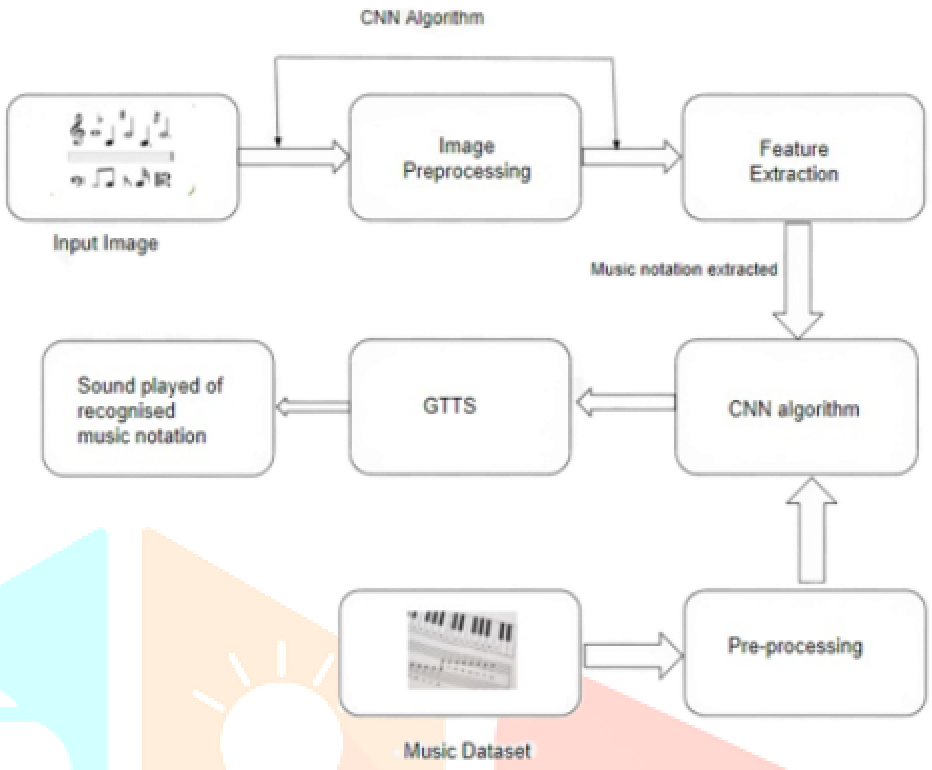

fig 1.: system architecture

In proposed system shown in fig.1, the system developed such that in which first image taken as an Input dataset and then apply the CNN algorithm on that image by which preprocess of the image is done and by feature extraction music notation from the image. After this comparison of this extracted notation with the predefined music dataset from the database is done. Then CNN algorithm will pass the accurate music notation to the GTTS (Python Library) Then GTTS recognized the sent input from CNN algorithm and create/play of that music notation.

V. REFERENCES

[1] Sanu Pulimootil Achankunju. Music search engine from noisy OMR data. In 1st International Workshop on Reading Music Systems, pages 23–24, Paris, France, 2018.

[2] Julia Adamska, Mateusz Piecuch, Mateusz Podgórski, Piotr Walkiewicz, and Ewa Lukasik. Mobile system for optical music recognition and music sound generation. In Computer Information Systems and Industrial Management, pages 571–582, Cham, 2015. Springer International Publishing.

[3] Francisco Álvaro, Joan-Andreu Sánchez, and José-Miguel Benedí. An integrated grammar-based approach for mathematical expression recognition. Pattern Recognition, 51:135–147, 2016.

[4] Apurva A. Mehta and Malay S. Bhatt. Optical music notes recognition for printed piano music score sheet. In International Conference on Computer Communication and Informatics, Coimbatore, India, 2015.

[5] Kia Ng, Alex McLean, and Alan Marsden. Big data optical music recognition with multi-images and multi recognisers. In EVA London 2014 on Electronic Visualisation and the Arts, pages 215–218. BCS, 2014.

[6] David Bainbridge and Tim Bell. A music notation construction engine for optical music recognition. Software: Practice and Experience, 33(2):173–200, 2003.

[7] Jorge Calvo-Zaragoza and Jose Oncina. Recognition of pen-based music notation: The HOMUS dataset. In 22nd International Conference on Pattern Recognition, pages 3038–3043. Institute of Electrical & Electronics Engineers (IEEE), 2014.

[8] Jorge Calvo-Zaragoza and David Rizo. Camera-primus: Neural end-to-end optical music recognition on realistic monophonic scores. In 19th International Society for Music Information Retrieval Conference, pages 248–255, Paris, France, 2018.

[9] Jorge Calvo-Zaragoza and David Rizo. End-to-end neural optical music recognition of monophonic scores. Applied Sciences, 8(4), 2018.

[10] Jorge Calvo-Zaragoza, Alejandro Toselli, and Enrique Vidal. Handwritten music recognition for mensural notation: Formulation, data and baseline results. In 14th International Conference on Document Analysis and Recognition, pages 1081–1086, Kyoto, Japan, 2017.

[11] Liang Chen, Rong Jin, and Christopher Raphael. Human-guided recognition of music score images. In 4th International Workshop on Digital Libraries for Musicology. ACM Press, 2017.

[12] Liang Chen and Christopher Raphael. Optical music recognition and human-in-the-loop computation. In 1st International Workshop on Reading Music Systems, pages 11–12, Paris, France, 2018.

[13] Arindam Chowdhury and Lovekesh Vig. An efficient end-to-end neural model for handwritten text recognition. In 29th British Machine Vision Conference, 2018.

[14] Matthias Dorfer, Jan Hajic jr., Andreas Arzt, Harald Frostel, and Gerhard Widmer. Learning audio–sheet music correspondences for ˇ cross-modal retrieval and piece identification. Transactions of the International Society for Music Information Retrieval, 1(1):22–33, 2018.

[15] Matthew J. Dovey. Overview of the OMRAS project: Online music retrieval and searching. Journal of the American Society for Information Science and Technology, 55(12):1100–1107, 2004.

[16] Hoda M. Fahmy and Dorothea Blostein. A graph grammar programming style for recognition of music notation. Machine Vision and Applications, 6(2):83–99, 1993.

[17] Christian Fremerey, Meinard Müller, Frank Kurth, and Michael Clausen. Automatic mapping of scanned sheet music to audio recordings. In 9th International Conference on Music Information Retrieval, pages 413–418, 2008.

[18] Susan E. George. Online pen-based recognition of music notation with artificial neural networks. Computer Music Journal, 27(2):70– 79, 2003.

[19] Susan E. George. Wavelets for dealing with super-imposed objects in recognition of music notation. In Visual Perception of Music Notation: On-Line and Off Line Recognition, pages 78–107. IRM Press, Hershey, PA, 2004.

[20] Angelos P. Giotis, Giorgos Sfikas, Basilis Gatos, and Christophoros Nikou. A survey of document image word spotting techniques. Pattern Recognition, 68:310–332, 2017.