Multimedia Tools and Applications (2020) 79:6383–6408
https://doi.org/10.1007/s11042-019-08200-0

OMR metrics and evaluation: a systematic review

Luciano Mengarelli[1] · Bruno Kostiuk[1] · João G. Vitório[1] · Maicon A. Tibola[1] · William Wolff[1] · Carlos N. Silla Jr.[1]

Received: 9 September 2018 / Revised: 3 June 2019 / Accepted: 8 September 2019 /
Published online: 14 December 2019

Abstract

Music is rhythm, timbre, tones, intensity and performance. Conventional Western Music Notation (CWMN) is used to generate Music Scores in order to register music on paper. Optical Music Recognition (OMR) studies techniques and algorithms for converting music scores into a readable format for computers. This work presents a systematic literature review (SLR) searching for metrics and methods of evaluation and comparing for OMR systems and algorithms. The most commonly used metrics on OMR works are described. A research protocol is elaborated and executed. From 802 publications found, 94 are evaluated. All results are organized and classified focusing on metrics, stages, comparisons, OMR datasets and related works. Although there is still no standard methodology for evaluating OMR systems, a good number of datasets and metrics are already available and apply to all the stages of OMR. Some of the analyzed works can give good directions for future works.

Keywords OMR · Metrics · Evaluation · Review · Optical music recognition · Music scores · image processing

✉ Luciano Mengarelli
luciano.mengarelli@gmail.com

Bruno Kostiuk
brunokostiuk@hotmail.com

João G. Vitório
joaobracaioli@gmail.com

Maicon A. Tibola
maicoaugusto@yahoo.com.br

William Wolff
wolff@wdw.com.br

Carlos N. Silla Jr.
carlos.sillajr@gmail.com

[1] Postgraduate Program in Computer Science — PPGIa — Polytechnic School, Pontifícia Universidade Católica do Paraná - PUCPR, Av. Imaculada Conceição, 1155 - Prado Velho, Curitiba, Parana 80215-901, Brazil

1 Introduction

When we talk about music, several thoughts cross our minds like rhythm, timbre, tones, intensity and also musicians playing, in other words, following a list of steps on their instruments to reach that sound. Conventional Western Music Notation (CWMN) is used to compose documents called “Music Scores” (Fig. 1) in order to describe those steps allowing historical register and music sharing. This notation is formed by symbols that indicate the tones, the duration and the way to perform [58].

Music scores are usually written on paper and can deteriorate over time. That’s why there is a great concern in saving and restoring them. The digitalisation has been used as a solution to preserve music scores, facilitating also the distribution and digital processing of them. “The development of general image processing methods for object recognition has contributed to the development of several important algorithms for Optical Music Recognition (OMR)” [82]. OMR has been a research area since the mid 1960s that study techniques and algorithms for converting or interpreting the music score into a readable format for computers [20].

An OMR system (Fig. 2), including also the image pre-processing as part of it, can be divide in [81]:

1. Stage 1 (S1 - See Fig. 3): Image pre-processing (binarization, noise removal, blurring, deskewing, etc);
2. Stage 2 (S2 - See Fig. 4): Recognition of musical symbols from images. Usually also subdivided in:
   (a) Staff lines detection and removal;
   (b) Symbols primitives segmentation;
   (c) Symbols recognition;
3. Stage 3 (S3): Reconstruction of the musical information in order to identify the music notation;
4. Stage 4 (S4): Construction of a model containing the symbolical presentation of the music score.

One of the most important algorithms in the image pre-processing stage (S1) is the binarization. The scanned image is analyzed to determine what information belongs to the music score (musical symbols and lines) and what should be disregarded (background and noise), reducing the amount of information to be processed [82].

The Staff Detection and Removal is a procedure with great impact for the OMR, since it isolates the music symbols through the removal of the staff lines of the score and can

Fig. 1 Music Score Sample

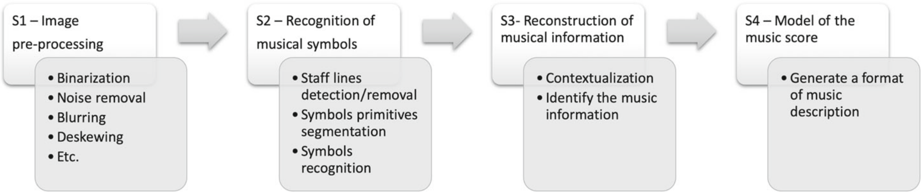

Fig. 2 OMR Stages [81]

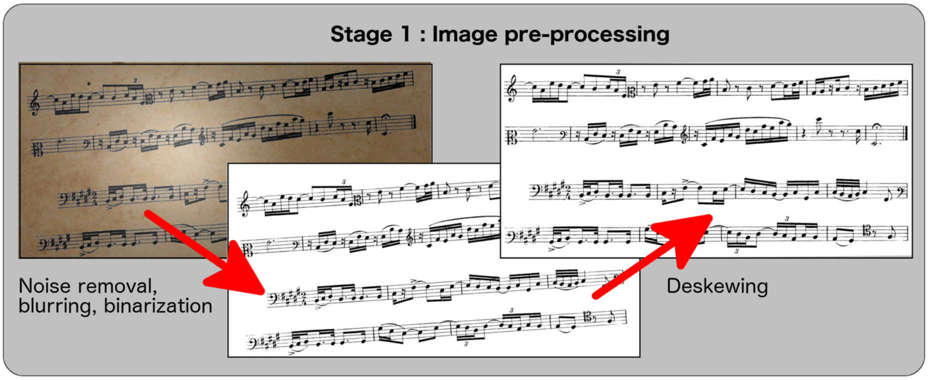

Fig. 3 OMR Stage 1

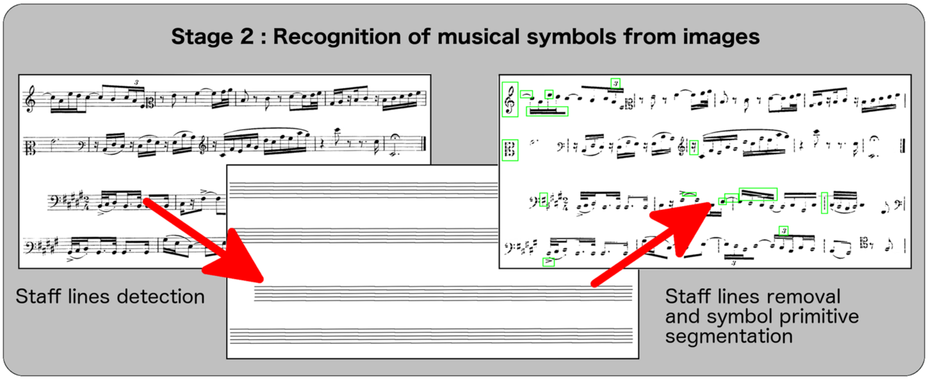

Fig. 4 OMR Stage 2

sometimes result in loss of information for the next steps of the process. For this reason, there are some authors who suggest OMR algorithms that do not remove the staff lines [82].

Segmentation and recognition of symbols extracts elementary graphic symbols like dots, rests, note heads and stems (see Table 1), usually using a classification task and occasionally with an alternative approach. All objects extracted are then classified again now considering size, bounding, linking, connections and other features in order to recognize higher level elements of the musical notation. There is a great variety of algorithms for the later task [81].

The later two stages are often interconnected. Here is where graphical and syntactic rules are used to contextualize the information transforming the symbols recognized on the previous stage in music information. Once all the music information is produced, the last stage is responsible for generating a format of music description, like MIDI files and more recently MusicXML is the more often used format [82].

Currently there is a great deal of work on OMR, commercial and free OMR systems. The question is: Which are the best algorithms? Which full system is more accurate? The problem is that nobody knows this answer. Therefore, it is of great importance to evaluate OMR systems through metrics and comparisons [12].

This work presents a systematic literature review searching for metrics and methods of evaluation and comparing OMR systems and algorithms. Through the research it was reported that each metric was applied in some specific procedure of the OMR. Section 1.2 describes the most commonly used metrics on OMR works. Section 2 explains the research protocol, the objectives, the used databases and how the research was executed and conduced. Section 3 organizes the information obtained from the research and displays metrics, stages, comparisons, OMR datasets and related works. Section 4 discuss some of the results and highlights some works.

1.1 Challenges

Optical music recognition area presents several challenges, for example: staff removal and symbol segmentation [66], ambiguities during graphical primitives detection and classification [4], imbalance of classes and overlapping elements [43]. Concerning handwritten scores, more challenges appear: high variability of handwritten styles, pens, papers and traces [4]. Considering the use of digital cameras, there are still: paper and staves bending, no parallelism between paper and camera resulting in size variations for elements, shadow and light influence [63]. Another important matter to be considered is that different errors during the OMR process can cause distinct levels of impact on final results. Music notation symbols can be organized considering their information type:

> "Clefs (Fig. 5a) determine the pitches for each line and space of the staff (Fig. 5b), accidentals (Fig. 5e) temporarily modify the pitch of following notes. The pitch of notes itself (Fig. 5c) is indicated by their vertical placement on the staff, and their appearance affects the relative duration. Ornaments (Fig. 5f) change the pitch pattern of individual notes. Rests (Fig. 5d) indicate a relative duration of silence. Dynamics (Fig. 5g) signify the varying loudness. Articulations (Fig. 5h) change the timbre or duration of a note." [65]

Analyzing those types, it's clear that miss identifying a dynamic (Fig. 5g) or an articulation (Fig. 5h) symbol will cause less damage to the final result than miss identifying rests (Fig. 5d) or notes (Fig. 5g) that can change the melody and the alignment of the entire music.

Table 1 Music Symbols [6]

Category	Example
Empty Note Head	
Black Note Head	
Augmentation Dot	
Rest Duration 4/4	
Rest Duration 2/4	
Rest Duration 1/4	
Rest Duration 1/8	
Rest Duration 1/16	
Rest Duration 1/32	
Rest Duration 1/64	
Barline Single	
Barline Double	
Barline End	
Barline Start Refrain	
Barline End Refrain	
Slur	
Sharp	
Flat	
Natural	
Double Sharp	
Double Flat	

Table 1 (continued)

Category	Example
Treble Clef	
Bass Clef	
Tenor Clef	
Hook 1 (1/8)	
Hook 2 (1/16)	
Hook 3 (1/32)	
Hook 4 (1/64)	
Beam 1 (1/8)	
Beam 2 (1/16)	
Beam 3 (1/32)	
Beam 4 (1/64)	
Accent	
Number 1	
Number 2	
Number 3	
Number 4	
Number 5	
Number 6	
Number 7	
Number 8	
Number 9	
Piano	
Forte	

Table 1 (continued)

Category	Example
Comma	
C	
Staccato	
Fermata	
Mordent	
Turn	
Grace Note	
Trill	
Tenuto	

We did some tests using sheet music from a guitar initiation book [72] and Musescore [62]. Figures 6 and 7 show some mistakes with different impacts on final results.

1.2 Metrics overview

The metrics are mainly used to measure the element loss, in the case of OMR scores, thus being able to measure the performance of the algorithms, usually using a scale from 0 to 100 [12, 35]. After a thorough analysis in this work, the main metrics used in OMR are: Accuracy, Precision, Recovery and Error Rate.

In the metrics has its focus on the predictive capacity of the model, using a confusion matrix to analyze this capability. The confusion matrix is the form of representation of the quality obtained from a classification, being expressed through the correlation of information of the reference data (understood as true) with the classified data [24].

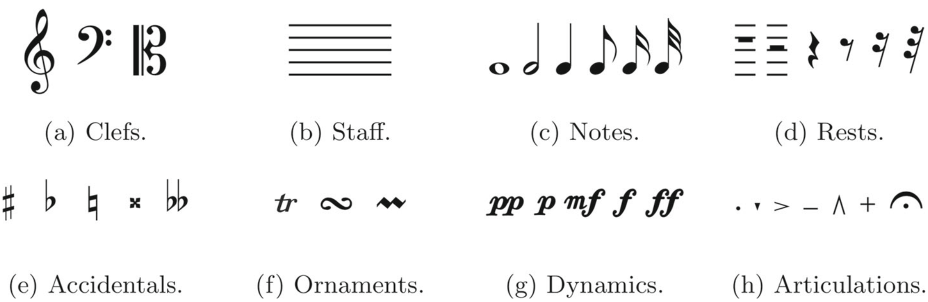

Fig. 5 Music notation symbols [65]

Prediction outcome

		p	**n**	**total**
actual value	**p′**	True Positive (tp)	False Negative (tn)	P′
	n′	False Positive (fp)	True Negative (tn)	N′
	total	P	N	

The confusion matrix is divided into the following categories: True positives (TP) corresponding labeled as positive. False positives (FPs) refer to the negative elements incorrectly labeled positive. True Negatives (TN) correspond to correctly labeled negatives. And false negatives (FN) refer to positive examples labeled incorrectly as negative [24].

Accuracy studies allow us to assess to what degree the data measures what they should measure or how much the results of a measurement correspond to the true state of the phenomenon measured [13]. The accuracy of a test or a method is commonly considered in relation to some gold standard. Accuracy can be treated as the combination of precision and recall, using a confusion matrix to perform the calculation [103]. Accuracy is represented by:

$$Accuracy = \frac{tp + tn}{tp + fp + fn + tn} \tag{1}$$

Precision is used in OMR to tell the number of properly divided detected scores by the number of detected elements) [4]. As can be seen in the formula:

$$Precision = \frac{tp}{tp + fp} \tag{2}$$

Similar to Precision, the Recall is number correctly detected scores divided by the number of score of the set that should be detected [4]. Is represented by:

$$Recall = \frac{tp}{tp + fn} \tag{3}$$

The Error rate, is also calculated through matrix confusion, and aims to compare the performs between two classifiers [23]. The formula represented by:

$$ErrorRate = \frac{fp + fn}{tp + tn + fp + fn} \tag{4}$$

Another commonly used metric is the weighted average F-measure (MP-F) consisting of the combination of recall and precision. The choice for the use of two averages occurred due to the imbalance of the classes, and could be detrimental to the performance of the classifiers [16, 17].

In the session Section 3.3, all metrics found in this article are exposed.

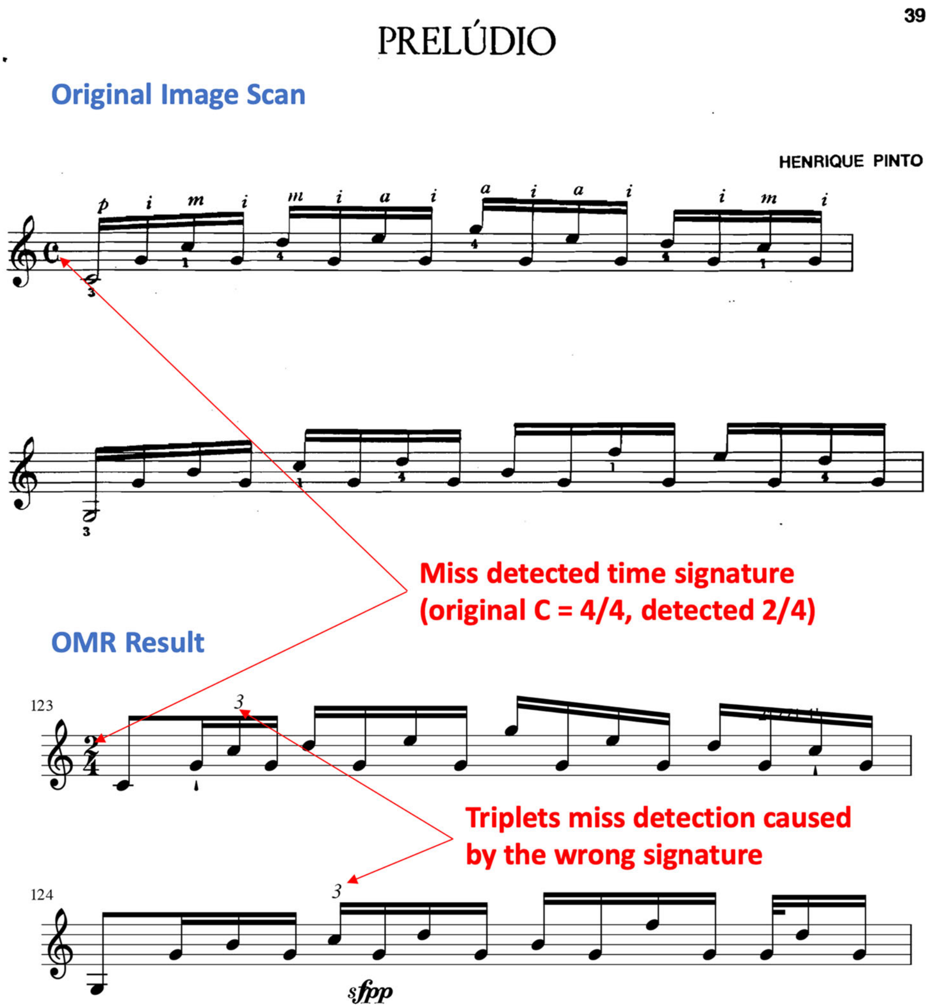

Fig. 6 OMR Mistakes Example 1

2 Research method

A systematic literature review (SLR) is a line of research with the purpose of identifying, evaluating and interpreting all relevant literature concerning a particular question, topic or interest. A review protocol is used to organize the research and define all the steps to be followed in order to avoid research bias and researcher impartiality. [45]

2.1 Justification

Besides the large number of works in OMR, there is still no standard for measuring, evaluating and testing OMR systems and algorithms. There are some databases available, but

Fig. 7 OMR Mistakes Example 2

usually authors use their own customized dataset and most of the times they are not available for future works [12]. An initial search has been conducted in order to find SLR works on OMR. It has been done using Google Scholar [40] with the following query string:

- allintitle: ("Systematic Review" OR survey OR "State of art") Optical Music Recognition

And also on Scopus [10] with the following query string:

- ("Systematic Review" OR survey OR "State of art" OR "State-of-art" OR "State of the art" OR "State-of-the-art") AND ("Optical Music Recognition" OR OMR)

After running the searches, we got 18 references, but only one was a review for OMR [82]. This result has reinforced and validated the idea of conducting a SLR for metrics and evaluation for OMR.

2.2 Research protocol

The research protocol defines the process that will be applied to perform the review aiming to achieve the proposed objective. The protocol comprises the research questions, the search strategy and the inclusion and exclusion criteria.

The chosen research question for this work is:

- How are OMR systems evaluated?

The main objective is to enumerate and compare methods of evaluation and metrics in order to allow future works to use them and establish new standards focusing the improvement of this area. Three outcomes are expected from this work:

- Identify more relevant works regarding evaluation and metrics for OMR;
- Identify new needs and possibilities in this area;
- Allow comparative studies among several existing methods;

2.2.1 Keyword definition

OMR unites two distinct areas: computing and music. It makes necessary to search for synonyms and keywords that can be found in it's works, taking into account the acronym "OMR" that can be found in several other areas. The keywords were chosen to find all available publications on the subject and are shown in Table 2:

2.2.2 Online libraries and search strings

The online libraries used in this research are shown in the Table 3:

For each online library it was needed to use a specific search string due to limitations and syntax differences:

1. **ACM:** Several searches combining the String:
   +(test* measur* metric* evaluat* assess* compar*) +("Optical Music Recognition" OMR "Music Score" "Music Scores")
   - With all the following Strings:
     +measur* +"Optical Music Recognition"
     +measur* + OMR
     +metric* + "Optical Music Recognition"

Table 2 Keywords

Test
Measurement
Metrics
Evaluation
Assessment
Comparison
Optical Music Recognition
OMR
Music Score
Music Scores

Table 3 Research bases

1. ACM [30]
2. IEEE [41]
3. Scopus [10]
4. SpringerLink [2]
5. Taylor & Francis Online [48]

+metric* + OMR
+evaluat* + "Optical Music Recognition"
+evaluat* + OMR
+assess* + "Optical Music Recognition"
+assess* + OMR
+compar* + "Optical Music Recognition"
+compar* + OMR

2. **IEEE:** Advanced search was used with the option "Metadata Only". Two strings where needed.
   (measur* OR metric* OR evaluat* OR assess* OR compar*) AND ("Optical Music Recognition" OR OMR OR "Music Score" OR "Music Scores")
   (test*) AND ("Optical Music Recognition" OR OMR OR "Music Score" OR "Music Scores")
3. **Scopus:**
   (test* OR measur* OR metric* OR evaluat* OR assess* OR compar*) AND ("Optical Music Recognition" OR OMR OR "Music Score" OR "Music Scores")
4. **SpringerLink:**
   (test* OR measur* OR metric* OR evaluat* OR assess* OR compar*) AND ("Optical Music Recognition" OR OMR OR "Music Score" OR "Music Scores")
5. **Taylor & Francis:** Search modified to search only for Title:
   (test* OR measur* OR metric* OR evaluat* OR assess* OR compar*) AND ("Optical Music Recognition" OR OMR OR "Music Score" OR "Music Scores")

2.2.3 Inclusion and exclusion criteria

The selection of articles was planned to be done in two different steps. The first step was to eliminate most unrelated articles quickly, using only the article title and abstract. The exclusion criteria were defined as:

- Exclude non-English articles;
- Exclude articles with title clearly not related to OMR;
- Exclude articles which abstract clearly indicates it is not related to OMR;
- Exclude duplicated articles;

The second step was more demanding because the entire article was considered and analyzed. The inclusion and exclusion criteria were defined as:

- Exclusion Criteria;
  - Exclude articles not related to OMR;
  - Exclude articles related to OMR but with no dataset, no comparison and no metric;

  - Exclude duplicated articles;

- Inclusion Criteria;
  - Include articles containing any metric or evaluation of OMR;
  - Include articles containing comparison among OMR systems or algorithms;
  - Include articles containing OMR datasets;

2.3 Conducting the review

In this subsection it is demonstrated how the research was conduced.

2.3.1 Selection of primary papers

After the execution of the searches on the databases, a total of 932 articles were found and organized in the Mendeley[1] platform (Reference manager and research network), which assisted in the pre-selection of the articles by excluding articles with duplicate titles, reducing them to 802 articles to be analyzed.

At this point, 5 teams were formed and each researcher was part of 2 teams. Those remaining articles were distributed to the teams in order to eliminate unrelated articles and duplicated titles only by analyzing abstract as described on the research protocol. Each team evaluated the articles applying the exclusion criteria defined for the first step.

At the end of this step, 624 articles were discarded and 178 articles were considered for the more careful analysis.

2.3.2 Detailed analysis in pairs

Each team now was responsible for 36 articles, except one group that was responsible for 34. Each team evaluated articles contents applying the inclusion and exclusion criteria defined for the second step.

2.3.3 Selected articles

The results were organized as spreadsheets and validated using the Kappa coefficient, since it's an efficient method to verify the level of agreement of the teams [99].

When the value one (1) is obtained for this coefficient, it means that the agreement between the teammates is 100%. The closer to zero is the result, the lower is the agreement. See Table 4

At the final of this step, 94 articles were selected to be part of the results as shown on Figs. 8 and 9.

3 Results

All the selected articles were analyzed and classified regarding metrics, datasets, stages of OMR, type of music score and existing comparisons. The concept was to identify all

[1] https://www.mendeley.com/

Table 4 Kappa

Member 1	Member 2	Kappa
João	Luciano	0,63
Bruno	Luciano	1,00
Maicon	William	0,88
João	William	0,66
Maicon	Bruno	0,81
Overall Kappa		**0,81**

datasets and metrics, to find the examples of comparison and provide lists of articles in an organized form. Some works could not be classified due to their particularity, such as: writer recognition for historical archives [7], layout analysis for ancient handwritten scores [18], audio and score alignment [42], handwritten score alignment [85], watermarking for music scores [89] and works on Chinese numeric music notation [105–107].

3.1 Metrics

The majority of the 94 selected articles uses accuracy as metric, many times referenced as Recognition Rate or Classification Rate. Several statistical metrics can also be found such as: Precision, Recall, Error Rate, TP (True Positives), TN (True Negatives), FP (False Positives), FN (False Negatives), Specificity, Sensitivity, F-Measure (F-Score) and Errors type I and II. On Table 5 the metrics were classified considering their scope and stage, and the following categories were selected:

- M1 - Metrics presented on general results or at the end of entire process.
- M2 - Metrics presented to evaluate the staff detection and removal.
- M3 - Metrics presented to evaluate the symbol segmentation/classification/recognition.
- M4 - Metrics presented to evaluate the result music notation, but only 3 works were presented:
  - [11] uses Note pitch error (%), Duration error (%) and Accuracy.
  - [76] uses recall and precision.
  - [21] uses F-Measure, precision and recall.

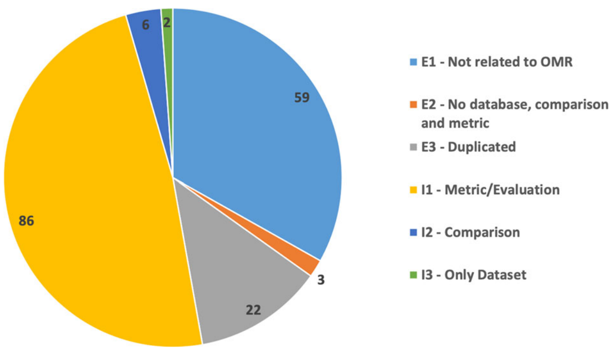

Fig. 8 Result of Detailed Analysis

Table 5 Metrics

Metric	M1-General		M2-Staff Det./Rem.		M3-Symbol Rec.	
	Q	Ref.	Q	Ref.	Q	Ref.
accuracy	20	[31, 44, 57, 74, 84, 94, 109] [20, 28, 49, 52, 68, 83, 96] [1, 13, 46, 50, 60, 103]	5	[26, 61, 70, 93, 104]	20	[3, 53, 86–88, 91, 108] [23, 37, 52, 81, 90, 95, 100] [43, 47, 56, 63, 66, 71]
precision	3	[9, 60, 103]	6	[21, 25, 27, 36, 69, 102]	9	[4, 36, 43, 47, 51, 63, 71, 76, 77]
TP/TN	2	[49, 50]	0		1	[56]
FP/FN	2	[49, 50]	0		5	[36, 51, 56, 59, 78]
error rate	4	[23, 29, 88]	6	[3, 19, 27, 32, 92, 98]	5	[15, 23, 43, 86, 97]
recall	3	[9, 60, 103]	7	[21, 25–27, 61, 69, 102]	5	[4, 59, 63, 76, 77]
specificity	0		3	[26, 61, 102]	0	
sensitivity	0		0		2	[43, 47]
F-Measure	0		8	[14, 16, 17, 21, 25, 38, 55, 102]	0	

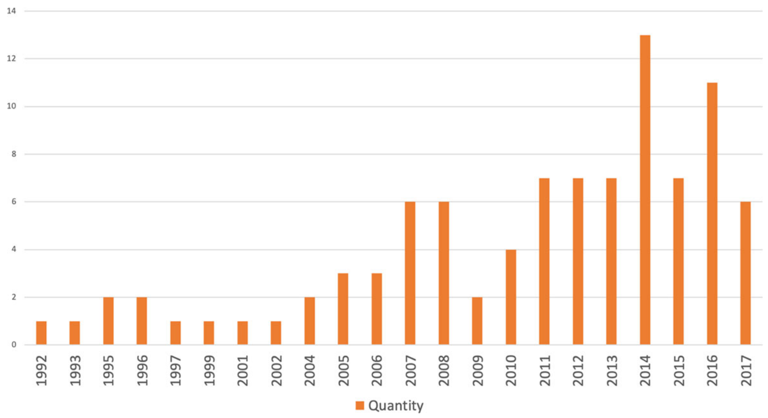

Fig. 9 Histogram by Year

3.2 Stages of OMR

Nearly 50% of the articles are focused only on the Stage 2 of OMR, therefore the recognition of musical symbols. The distribution can be checked on Table 6. Some of the collected data about OMR and its stages and methods are presented on the following sections, resulting on a taxonomy shown on Fig. 10. As the main focus of the paper is on metrics and evaluation of Optical Music Recognition systems, more detailed information can be obtained in some OMR reviews found during this work. Fornés and Sánchez [39] provide a review of OMR methods and stages, Rebelo et al. [82] present the State-of-art of OMR systems. There are also works focused only on a specific task as binarization methods [9] and comparison works regarding staff removal [22] and symbols recognition [22].

3.2.1 Stage 1: Image pre-processing

Concerning the image pre-processing, it was possible to identify different methods for each kind of task. Noise removal has examples using morphological operations [31, 37] and median filters [1, 18]. Morphological operations were also used for blurring correction. Most of the works considering image skew uses the Hough Transform Algorithm.

Binarization was the algorithm with more examples and it's important to highlight the comparative survey found [9]. This algorithm usually begins with a gray-scale conversion. Otsu's [67] threshold was the most used method [13, 18, 19, 98, 103]. However other methods were presented, such as: adaptive and heuristic techniques [37, 52, 73], Niblack [31], constant threshold [95, 96], adaptive filtering [1] and Gaussian Mixture Markov Random Field (GMMRF) [102].

3.2.2 Stage 2: Recognition of musical symbols

It is important to highlight the large number of works and competitions [32, 34, 38, 101] involving this stage, mainly discussing staff lines detection and removal. For this task, the majority of works use Histograms/Y-Projection [1, 13, 42, 52, 56, 88, 93, 96]. Other methods are also used, such as: median filters [31, 36, 37, 51], morphological operations [36, 60], shorted/stable path [25, 63, 80, 103] and artificial intelligence methods [14, 16, 17, 26, 61,

Table 6 Stages analysis

Stage	Quant.	Perc.	References
S1	4	4,7%	[9, 73, 74, 102]
S1, S2	8	9,4%	[70, 77, 96, 100] [13, 20, 25, 69]
S2	42	49,4%	[3, 57, 59, 91, 93, 104, 108] [6, 22, 36, 51, 75, 80, 87] [19, 27, 32, 37, 78, 81, 95] [26, 49, 54, 64, 92, 98, 101] [29, 38, 47, 56, 60, 63, 71] [4, 14–17, 61, 66]
S2, S3	3	3,5%	[43, 88, 90]
S3	2	2,4%	[50, 83]
S3, S4	1	1,2%	[5]
All	25	29,4%	[11, 44, 53, 84, 86, 94, 109] [8, 23, 28, 52, 76, 82, 97] [39, 46, 55, 68, 79, 105, 106] [1, 12, 21, 103]

70]. There are also many papers proposing new methods for staff detection/removal [20, 27, 89, 92, 98] and comparative studies [22, 32, 34, 38, 101]. There is also a paper proposing a method for staff detection without having to remove it [13].

After staff detection/removal, usually, comes the symbol primitive segmentation using methods such as: hough transformation [88], morphological operations [37, 54, 59], histogram/x-projection [13, 52, 96], templates [42], hierarchical decomposition [56], combined neural network [103], SVM classifier [21].

Once the primitive segmentation is ready, the next step is the symbol recognition. There are examples using: fuzzy model [86–88, 91], musical rules and pattern matching [52, 103], neural networks [53, 95, 96], decision tree [1, 20, 43], neural networks [56, 71], kNN [43,

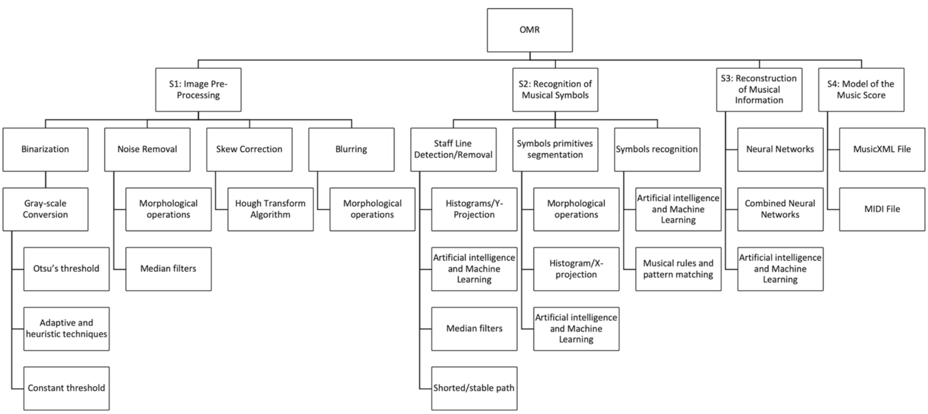

Fig. 10 Taxonomy

46], random forest [47] workflow for primitive assembly [50], SVM [29, 66] and finite state machines [15]. There is also a comparative study on symbols recognition [81].

3.2.3 Stage 3: Reconstruction of musical information

Very few papers address this stage. Neural networks [57], three Combined Neural Networks (CNN) using majority vote [83] and semantic reconstruction are some of the methods adopted for this stage.

3.2.4 Stage 4: Model of the music score

Some papers generate XML, MIDI or equivalent files at the end of the process [1, 8, 52, 68].

There are also full OMR systems and frameworks such as COMSCAN [90], Lemon [84], Gamera Gamut [76] and Aruspix [100].

3.3 Comparisons

Concerning the existing comparisons, 26 articles present some kind of comparison and they were divided in two categories:

- C1 - Comparison with commercial systems: SmartScore, MIDISCAN, NoteScan, PhotoScore, SharpEye.
- C2 - Comparison with other algorithms.

See Table 7 for more details.

See Table 8 for commercial systems used for comparisons.

It was possible to identify two events competitions: "ICDAR / GREC 2011 Competition: Writer Identification and Staff Removal" [32][2] and "ICDAR / GREC 2013 Competition on Music Scores: Staff Removal" [101][3]. Both are focused on Stage 2: staff removal and writer recognition of handwritten music scores. The writer recognition is usually made using classification algorithms after the staff removal. There was also an extension of the ICDAR / GREC 2011 Competition with some new image distortion creating tree levels of difficulty [34].

It is possible to see on Table 7 the difficulty of establishing benchmarks and rankings. Each work uses different metrics and datasets for comparing to commercial systems and with other methods.

3.4 Available datasets

In this section the datasets will be presented in summary form in Table 9 and detailed as follows.

3.4.1 The synthetic score database by Christoph Dalitz

It's a test set from Gamera Kit used to evaluate staff removal algorithms, contains historical scores, modern scores and tablatures. The images are in PNG format with 300dpi and were generated from 32 music scores using different softwares, also provides the ground truth.[4]

[2] http://www.cvc.uab.es/cvcmuscima/competition/index.htm

[3] http://www.cvc.uab.es/cvcmuscima/competition2013/

[4] http://gamera.informatik.hsnr.de/addons/musicstaves/testset-musicstaves.tar.gz

Table 7 Articles with comparison

Category	Quant.	Metrics
C1 - Comparison with commercial systems	12	accuracy [20] FP [84] visual comparison [86, 87] error rate [11, 88] error/rank [8] TP/FP/FN/accuracy [49, 50, 79] precision/recall [4] very specific and detailed metrics [6]
C2 - Comparison with other algorithms	14	accuracy [37, 61, 74, 81, 97, 100, 109] recall/precision [9, 76] visual comparison [80] error rate [22, 32, 92] F-Measure/precision/recall/specificity [102]

3.4.2 CVC-Muscima database

CVC-Muscima [33] is a database for handwritten staff removal and writer identification. With 1000 images for writer identification and 12,000 for staff removal, it contains transcriptions from 50 different musicians writing the same 20 scores. All images have been generated in 300dpi and 24bits and later converted to 8bit gray scale image. The 1000 images have been transformed in order to generate deformations and providing a database for staff removal with 12,000 images.[5]

3.4.3 IMSLP - Petrucci music library

International Music Score Library Project (IMSLP) or Petrucci Music Library is a virtual library of musical scores and works from contemporary composers. It started in 2006 with the objective of provide the complete works of Johann Sebastian Bach, although now there are works, music scores and recordings.[6]

3.4.4 HOMUS database

Handwritten Online Musical Symbols (HOMUS) is a handwritten musical score database developed by 100 musicians from Escuela de Educandos Asociación Musical el Avanç (El Campello, Spain) and the Superior Music Conservatory of Murcia "Manuel Massotti Littell" (Murcia, Spain). The musicians have drew 32 musical symbols on their own style generating a set of 15,200 samples in 36 models. The symbols were draw using a Samsung Galaxy Note 10.1 device with 149ppi (pixel per inch) and sample rate of 60fps using Stylus S-Pen.[7]

[5] http://www.cvc.uab.es/cvcmuscima/index_database.html

[6] http://imslp.org/wiki/Main_Page

[7] http://grfia.dlsi.ua.es/homus/

Table 8 Commercial systems

Commercial system	References
Capella-Scan	[8]
MIDISCAN	[84]
NoteScan	[84]
O^3MR	[6]
PhotoScore	[4, 11]
SharpEye	[6, 20, 49, 50, 79]
SmartScore	[6, 8, 11, 49, 50, 86–88]

3.4.5 SNU Dataset for online music symbol recognition

It's handwritten musical symbols database with 1,716 sample of symbols and 18 different musicians. It contains 16 musical symbols found on HOMUS Database.[8]

3.4.6 Digital scriptorium

Digital Scriptorium is an organization with Medieval and Renaissance manuscripts in digital format and offers free access for public domain works. Until September of 2017 there were 8,133 manuscripts and 75,922 image with 34 participants from distinct locations. It also provides identification of images and metadata as locations, dates, etc.[9]

3.4.7 MUSICNETWORK OMR assessment

Assessment form available online providing seven printed scores with three results of OMR systems each. The form intends to collect opinions about the results and relevances about OMR elements.[10]

3.4.8 NEUMES project

Neumed & Ekphonetic Universal Manuscript Encoding Standard provides an infrastructure for medieval chants manuscripts through a software for transcription and description in digital format.[11]

3.4.9 (OMR-ChSR6306) OMR chord separation and recognition 6306 database

It's a database of musical chords images captured from any music score using Samsung Note 2 camera. The music scores are photographed and processed in order to extract the chords, using binarization, staff line removal, forming a database with 6,306 chord images. The result images have resolution from 7x20 to 30x50 and are destined to chord recognition.[12]

[8] http://mipal.snu.ac.kr/index.php/Repository
[9] http://www.digital-scriptorium.org/
[10] http://www.disit.org/5932
[11] http://www.scribeserver.com/NEUMES/index.html
[12] https://sites.google.com/site/elyorkodirovresearch/omr-chsr6306

Table 9 Datasets

Dataset	Summary	Groudtruth
The Synthetic Score Database by Christoph Dalitz	Historical scores, modern scores and tablatures images are in PNG, 300dpi.	Yes
CVC-Muscima Database	Handwriten scores images (300dpi). 1,000 images for writer identification and 12,000 for staff removal, including deformations. Transcriptions from 50 different musicians writing the same 20 scores.	Yes
IMSLP - Petrucci Music Library	Library of musical scores, works and recordings from contemporary composers.	No
HOMUS Database	Handwritten musical score images (149ppi) drawn by 100 musicians including 32 musical symbols resulting a set of 15,200 samples in 32 models.	Yes
SNU Dataset for Online Music Symbol Recognition	Handwritten musical score images drawn by 18 musicians resulting a set of 1,716 samples.	Yes
Digital Scriptorium	Until September of 2017 there ware 8,133 manuscripts and 75,922 image with 34 participants from distinct locations.	No
MUSICNETWORK OMR Assessment	Seven printed scores images with three results of OMR systems each.	No
NEUMES Project	Medieval chants manuscripts.	Yes
(OMR-ChSR6306) OMR Chord Separation and Recognition 6306 Database	Database of 6,306 musical chords images captured from music scores using SamsungNote 2 camera. Destined to chord recognition.	Yes

4 Discussion

After all the information gathered and organized, it's possible to see that nowadays there are some databases that can be used to develop, test and compare existing OMR systems. However, a set of standards and methods has not been established yet. Most of the metrics are only statistical, but that is not the best scenario for OMR systems. Despite that there is no agreement about how to measure errors yet, some authors highlight the importance of considering errors with different weights. Bugge et al. [8] have presented a detailed work about error counting when prioritizing sound. Their work converts the MusicXML final result in a simpler version and uses a sequence alignment algorithm to compare results among several OMR systems with very specific error rules. Byrd and Simonsen [12] have proposed rules, definitions and metrics for a Standard Testbed for Optical Music Recognition, addressing also the counting of errors. It is an important step in an area that suffers from this lack for over 20 years. Their work addresses issues such as: image quality level, complexity of notation, tightness of spacing and rules for error counting.

5 Conclusion

Although there is still no standard methodology for evaluating OMR systems, it's possible to observe specific efforts and a greater concern for the last years. A good number of datasets are already available and some of the analyzed works can give good directions for future works. Some statistic metrics can still provide good results when comparing only some stages of OMR if one of the standard datasets is adopted. Researchers should always try to use all the information gathered and datasets available in order to be related and referenced to the great deal of already available works. This work together with those important works discussed above should support the OMR community in the arduous task of achieving its own standard evaluation method.

Acknowledgments We would like to thank CAPES (Coordination for the Improvement of Higher Level Personnel) and CNPq (National Council for Scientific and Technological Development) from the Brazilian government and Fundação Araucária for their financial support on the project.

References

1. Adamska J, Piecuch M, Podgórski M, Walkiewicz P, Lukasik E (2015) Mobile system for optical music recognition and music sound generation. In: IFIP International conference on computer information systems and industrial management. Springer , pp 571–582
2. AG SIP (2017) Springer link. https://link.springer.com/
3. Bainbridge D, Bell T (1997) Dealing with superimposed objects in optical music recognition. In: Proceedings of the 6th International Conference on image processing and its applications. IET
4. Baró A, Riba P, Fornés A (2016) Towards the recognition of compound music notes in handwritten music scores. In: Proceedings of the International Conference on frontiers in handwriting recognition (ICFHR). IEEE, pp 465–470
5. Baumann S (1995) A simplified attributed graph grammar for high-level music recognition. In: Proceedings of the 3rd International Conference on document analysis and recognition. IEEE, vol 2, pp 1080–1083
6. Bellini P, Bruno I, Nesi P (2007) Assessing optical music recognition tools. Comput Music J 31(1):68–93
7. Bruder I, Ignatova T, Milewski L (2004) Knowledge-based scribe recognition in historical music archives. Lecture notes in computer science, pp 304–316
8. Bugge EP, Juncher KL, Mathiasen BS, Simonsen JG (2011) Using sequence alignment and voting to improve optical music recognition from multiple recognizers. In: Proceedings of the 12th international society for music information retrieval (ISMIR) conference
9. Burgoyne JA, Pugin L, Eustace G, Fujinaga I (2007) A comparative survey of image binarisation algorithms for optical recognition on degraded musical sources. In: Proceedings of the international Society for music information retrieval (ISMIR) conference, pp 509–512
10. Scopus EBV (2017) https://www.scopus.com/
11. Byrd D, Schindele M (2006) Prospects for improving omr with multiple recognizers. In: Proceedings of the int. Society for music information retrieval (ISMIR) conference, pp 41–46
12. Byrd D, Simonsen JG (2015) Towards a standard testbed for optical music recognition Definitions, metrics, and page images. J Music Res 44(3):169–195
13. Calvo-Zaragoza J, Barbancho I, Tardón LJ, Barbancho AM (2015) Avoiding staff removal stage in optical music recognition: application to scores written in white mensural notation. Pattern Anal Appl 18(4):933–943
14. Calvo-Zaragoza J, Micó L, Oncina J (2016) Music staff removal with supervised pixel classification. Int J Doc Anal Recogn (IJDAR) 19(3):211–219
15. Calvo-Zaragoza J, Oncina J (2017) Recognition of pen-based music notation with finite-state machines. Expert Syst Appl 72:395–406
16. Calvo-Zaragoza J, Pertusa A, Oncina J (2017) Staff-line detection and removal using a convolutional neural network. Mach Vis Appl 28:1–10

17. Calvo-Zaragoza J, Vigliensoni G, Fujinaga I (2017) Staff-line detection on grayscale images with pixel classification. In: Iberian conference on pattern recognition and image analysis. Springer, pp 279–286
18. Campos VB, Calvo-Zaragoza J, Toselli AH, Ruiz EV (2016) Sheet music statistical layout analysis. In: Proceedings of the International Conference on frontiers in handwriting recognition (ICFHR). IEEE, pp 313–318
19. Cardoso J, Rebelo A (2010) Robust staffline thickness and distance estimation in binary and gray-level music scores. In: Proceedings of the International Conference on pattern recognition (ICPR). IEEE, pp 1856–1859
20. Chen Y-S, Chen F-S, Teng C-H (2013) AN optical music recognition system for skew or inverted musical scores. International Journal of Pattern Recognition and Artificial Intelligence 27(7):1353005
21. Chen L, Duan K (2016) Midi-assisted egocentric optical music recognition. In: Proceedings of the IEEE winter conference on applications of computer vision (WACV). IEEE, pp 1–9
22. Dalitz C, Droettboom M, Pranzas B, Fujinaga I (2008) A comparative study of staff removal algorithms. IEEE Trans Pattern Anal Mach Intell 30(5):753–766
23. Dalitz C, Michalakis GK, Pranzas C (2008) Optical recognition of psaltic byzantine chant notation. Int J Doc Anal Recogn (IJDAR) 11(3):143–158
24. Davis J, Goadrich M (2006) The relationship between precision-recall and roc curves. In: Proceedings of the 23rd international conference on Machine learning. ACM, pp 233–240
25. Dinh CM, Yang HJ, Lee G, Kim SH (2016) Fast lyric area extraction from images of printed korean music scores. IEICE Trans Inf Syst 99(6):1576–1584
26. dos Santos Montagner I, Hirata R, Hirata NS (2014) A machine learning based method for staff removal. In: Proceedings of the International Conference on pattern recognition (ICPR). IEEE, pp 3162–3167
27. Dutta A, Pal U, Fornes A, Llados J (2010) An efficient staff removal approach from printed musical documents. In: Proceedings of the International Conference on pattern recognition (ICPR). IEEE, pp 1965–1968
28. Fahn C-S, Lu K-J (2014) Humanoid recognizing piano scores techniques. In: Proceedings of the International Conference on information science, electronics and electrical engineering (ISEEE). IEEE, vol 3, pp 1397–1402
29. Fang Y, Gui-fa T (2015) Visual music score detection with unsupervised feature learning method based on k-means. Int J Mach Learn Cybern 6(2):277–287
30. A. for Computing Machinery. Acm digital library (2017) https://dl.acm.org/
31. Fornes A, Lladós J, Sánchez G, Bunke H (2009) On the use of textural features for writer identification in old handwritten music scores. In: Proceedings of the 10th International Conference on document analysis and recognition (ICDAR). IEEE, pp 996–1000
32. Fornes A, Dutta A, Gordo A, Llados J (2011) The icdar 2011 music scores competition: Staff removal and writer identification. In: Proceedings of the International Conference on document analysis and recognition (ICDAR). IEEE, pp 1511–1515
33. Fornés A, Dutta A, Gordo A, Lladós J (2012) Cvc-muscima: a ground truth of handwritten music score images for writer identification and staff removal. Int J Doc Anal Recognit 15:1–9
34. Fornés A, Dutta A, Gordo A, Lladós J (2013) The 2012 music scores competitions: staff removal and writer identification. In: Graphics recognition. New trends and challenges. Springer, pp 173–186
35. Fornės A, Dutta A, Gordo A, Lladȯs J (2013) The 2012 music scores competitions Staff removal and writer identification. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 7423 LNCS: 173–186
36. Fornés A, Lladós J, Sánchez G (2005) Primitive segmentation in old handwritten music scores. In: International workshop on graphics recognition. Springer, pp 279–290
37. Fornés A, Lladós J, Sánchez G (2007) Old handwritten musical symbol classification by a dynamic time warping based method. In: International workshop on graphics recognition. Springer, pp 51–60
38. Fornés A, Kieu VC, Visani M, Journet N, Dutta A (2013) The icdar/grec 2013 music scores competition: Staff removal. In: International workshop on graphics recognition. Springer, pp 207–220
39. Fornés A, Sánchez G (2014) Analysis and recognition of music scores. In: Handbook of document image processing and recognition. Springer, pp 749–774
40. Google (2017) Google scholar. https://scholar.google.com/
41. IEEE (2017) Ieee xplore digital library. http://ieeexplore.ieee.org/search/
42. İzmirli Ö, Sharma G (2012) Bridging printed music and audio through alignment using a mid-level score representation. In: Proceedings of the Int. Society for Music Information Retrieval (ISMIR) Conference, pp 61–66
43. Jastrzebska A, Lesinski W (2016) Optical music recognition as the case of imbalanced pattern recognition: a study of single classifiers. In: Knowledge, information and creativity support systems: Recent trends, advances and solutions. Springer, pp 493–505

44. Kato H, Inokuchi S (1992) A recognition system for printed piano music using musical knowledge and constraints. Structured Document Image Analysis:435–455
45. Kitchenham B (2004) Procedures for performing systematic reviews. Keele UK Keele Univ 33(2004):1–26
46. Kodirov E, Han S, Lee G-S, Kim Y (2014) Music with harmony: chord separation and recognition in printed music score images. In: Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication. ACM, pp 50
47. Lesinski W, Jastrzebska A (2015) Optical music recognition: music Standard and cost-sensitive learning with imbalanced data. In: IFIP International conference on computer information systems and industrial management. Springer, pp 601–612
48. Limited IU (2017) Taylor & francis online. http://www.tandfonline.com/
49. Liu X (2012) Note symbol recognition for music scores. Intelligent Information and Database Systems 7197:263–273
50. Liu X, Zhou M, Xu P (2015) A robust method for musical note recognition. In: Proceedings of the 14th International Conference on computer-aided design and computer graphics (CAD/graphics). IEEE, pp 212–213
51. López B et al (2005) Staff and graphical primitive segmentation in old handwritten music scores. Artif Intell Res Dev 131:83
52. Luangnapa N, Silpavarangkura T, Nukoolkit C, Mongkolnam P (2012) Optical music recognition on android platform. Adv Inf Technol 344:106–115
53. Luckner M (2006) Recognition of noised patterns using non-disruption learning set. In: Proceedings of the 6th International Conference on intelligent systems design and applications (ISDA). IEEE, vol 1, pp 557–562
54. Malik R, Roy PP, Pal U, Kimura F (2013) Handwritten musical document retrieval using music-score spotting. In: Proceedings of the 12th international conference on document analysis and recognition (ICDAR). IEEE, pp 832–836
55. Märgner V, Abed HE (2014) Tools and metrics for document analysis systems evaluation. Handbook of Document Image Processing and Recognition:1011–1036
56. Mehta AA, Bhatt MS (2015) Optical music notes recognition for printed piano music score sheet. In: Proceedings of the International Conference on computer communication and informatics (ICCCI). IEEE, pp 1–6
57. Miyao H, Nakano Y (1995) Head and stem extraction from printed music scores using a neural network approach. In: Proceedings of the 3rd International Conference on document analysis and recognition. IEEE, vol 2, pp 1074–1079
58. Modayur BR, Ramesh V, Haralick RM, Shapiro LG (1993) MUSER A prototype musical score recognition system using mathematical morphology. Mach Vis Appl 6(2-3):140–150
59. Modayur BR, Ramesh V, Haralick RM, Shapiro LG (1993) Muser: A prototype musical score recognition system using mathematical morphology. Mach Vis Appl 6(2):140–150
60. Montagner IS, Hirata R, Hirata NS (2014) Learning to remove staff lines from music score images. In: Proceedings of the IEEE international conference on image processing (ICIP). IEEE, pp 2614–2618
61. Montagner IS, Hirata NS, Hirata R (2017) Staff removal using image operator learning. Pattern Recogn 63:310–320
62. MuseScore BVBA. Musescore. https://musescore.org/
63. Na IS, Kim SH (2016) Music symbol recognition by a lag-based combination model. Multimedia Tools and Applications 76:1–17
64. Nhat VQ, Lee G (2014) Adaptive line fitting for staff detection in handwritten music score images. In: Proceedings of the 8th aInternational Conference on Ubiquitous Information Management and Communication. ACM, pp 99
65. Novotnỳ J., Pokornỳ J (2015) Introduction to optical music recognition: Overview and practical challenges. In: Proceedings of the int. Workshop on databases, texts, specifications, and objects (DATESO), pp 65–76
66. Oh J, Son SJ, Lee S, Kwon J-W, Kwak N (2017) Online recognition of handwritten music symbols. Int J Doc Anal Recogn (IJDAR) 20(2):79–89
67. Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 9(1):62–66
68. Padilla V, Marsden A, McLean A, Ng K (2014) Improving omr for digital music libraries with multiple recognisers and multiple sources. In: Proceedings of the 1st International Workshop on Digital Libraries for Musicology. ACM, pp 1–8

69. Pedersoli F, Tzanetakis G (2016) Document segmentation and classification into musical scores and text. Int J Doc Anal Recogn (IJDAR) 19(4):289–304
70. Piatkowska W, Nowak L, Pawłowski M, Ogorzałek M (2012) Stafflines pattern detection using the swarm intelligence algorithm. Computer Vision and Graphics 7594:557–564
71. Pinheiro Pereira RM, Matos CE, Braz Junior G, de Almeida JD, de Paiva AC (2016) A deep approach for handwritten musical symbols recognition. In: Proceedings of the 22nd Brazilian Symposium on Multimedia and the Web. ACM, pp 191–194
72. Pinto H (1978) Iniciaċao ao violao, vol 1. Ed. Ricordi, Sao Paulo
73. Pinto T, Rebelo A, Giraldi GA, Cardoso J (2011) Music score binarization based on domain knowledge. In: IbPRIA. Springer, pp 700–708
74. Pugin L, Burgoyne JA, Fujinaga I (2007) Goal-directed evaluation for the improvement of optical music recognition on early music prints. In: Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries. ACM, pp 303–304
75. Pugin L, Burgoyne JA, Fujinaga I (2007) Reducing costs for digitising early music with dynamic adaptation. In: International conference on theory and practice of digital libraries. Springer, pp 471–474
76. Pugin L, Hockman J, Burgoyne JA, Fujinaga I (2008) Gamera versus aruspix–two optical music recognition approaches. In: Proceedings of the int. Society for music information retrieval (ISMIR) conference. Citeseer
77. Ramirez C, Ohya J (2010) Symbol classification approach for omr of square notation manuscripts. In: Proceedings of the int. Society for music information retrieval (ISMIR) conference, pp 549–554
78. Raphael C, Wang J (2011) New approaches to optical music recognition. In: Proceedings of the int. Society for music information retrieval (ISMIR) conference, pp 305–310
79. Raphael C, Jin R (2014) Optical music recognition on the international music score library project. In: Proceedings of the documento recognition and treieval (DDR) conference, pp 90210f–90210f
80. Rebelo A, Capela A, Da Costa JFP, Guedes C, Carrapatoso E, Cardoso J (2007) A shortest path approach for staff line detection. In: Proceedings of the 3rd int. Cont. on automated production of cross media content for multi-channel distribution (AXMEDIS). IEEE, pp 79–85
81. Rebelo A, Capela G, Cardoso J (2010) Optical recognition of music symbols. Int J Doc Anal Recognit 13(1):19–31
82. Rebelo A, Fujinaga I, Paszkiewicz F, Marcal ARS, Guedes C, Cardoso J (2012) Optical music recognition: state-of-the-art and open issues. Int J Multimed Inf Retrieval 1(3):173–190
83. Rebelo A, Marċal AR, Cardoso J (2013) Global constraints for syntactic consistency in omr: an ongoing approach. In: International conference image analysis and recognition. Springer, pp 734–741
84. Reed KT, Parker JR (1996) Automatic computer recognition of printed music. In: Proceedings of the 13th International Conference on Pattern Recognition. IEEE, vol 3, pp 803–807
85. Riba P, Fornés A, Lladós J (2015) Towards the alignment of handwritten music scores. In: International workshop on graphics recognition. Springer, pp 103–116
86. Rossant F, Bloch I (2004) A fuzzy model for optical recognition of musical scores. Fuzzy Sets Syst 141(2):165–201
87. Rossant F, Bloch I (2005) Optical music recognition based on a fuzzy modeling of symbol classes and music writing rules. In: Proceedings of the IEEE international conference on image processing (ICIP). IEEE, vol 2, pp II–538
88. Rossant F, Bloch I (2007) Robust and adaptive omr system including fuzzy modeling, fusion of musical rules, and possible error detection. EURASIP J Appl Signal Process 2007(1):160–160
89. Schmucker M (2002) Staff line features as information carrier. In: Proceedings of the First International Symposium on Cyber Worlds. IEEE, pp 168–175
90. Sharif M, Arshad Q-A, Raza M, Khan WZ (2009) [comscan]: an optical music recognition system. In: Proceedings of the 7th International Conference on Frontiers of Information Technology. ACM, pp 34
91. Su M-C, Tew C-Y, Chen H-H (2001) Musical symbol recognition using som-based fuzzy systems. In: Proceedings of the joint 9th IFSA world congress and 20th NAFIPS international conference. IEEE, vol 4, pp 2150–2153
92. Su B, Lu S, Pal U, Tan CL (2012) An effective staff detection and removal technique for musical documents. In: Proceedings of the 10th IAPR international workshop on document analysis systems (DAS). IEEE, pp 160–164
93. Szwoch M (2005) A robust detector for distorted music staves. Lect Notes Comput Sci 3691:701–708
94. Szwoch M (2008) Using musicxml to evaluate accuracy of omr systems. Diagrammatic Representation and Inference 5223:419–422

95. Tambouratzis T (2011) Identification of key music symbols for optical music recognition and on-screen presentation. In: Proceedings of the international joint conference on neural networks (IJCNN). IEEE, pp 1935–1942
96. Tambouratzis T (2013) The digital music stand as a minimal processing custom-made optical music recognition system, part 1: Key music symbol recognition. Int J Intell Syst 28(5):474–504
97. Thomas V, Wagner C, Clausen M (2011) Ocr based post processing of omr for the recovery of transposing instruments in complex orchestral scores. In: Proceedings of the int. Society for music information retrieval (ISMIR) conference, pp 411–416
98. Timofte R, Van Gool L (2012) Automatic stave discovery for musical facsimiles. In: Asian conference on computer vision. Springer, pp 510–523
99. Viera AJ, Garrett JM et al (2005) Understanding interobserver agreement: the kappa statistic. Fam Med 37(5):360–363
100. Vigliensoni G, Burgoyne JA, Hankinson A, Fujinaga I (2011) Automatic pitch recognition in printed square-note notation. In: Proceedings of the int. Society for music information retrieval (ISMIR) conference
101. Visaniy M, Kieu VC, Fornés A, Journet N (2013) Icdar 2013 music scores competition: Staff removal. In: Proceedings of the 12th international conference on document analysis and recognition (ICDAR). IEEE, pp 1407–1411
102. Vo QN, Kim SH, Yang HJ, Lee G (2016) An mrf model for binarization of music scores with complex background. Pattern Recogn Lett 69:88–95
103. Wen C, Rebelo A, Zhang J, Cardoso J (2015) A new optical music recognition system based on combined neural network. Pattern Recogn Lett 58:1–7
104. Wijaya K, Bainbridge D (1999) Staff line restoration. Proceedings of the 7th International Conference on Image Processing and Its Applications
105. Wu F-HF, Jang J-SR (2014) An architecture for optical music recognition of numbered music notation. In: Proceedings of International Conference on Internet Multimedia Computing and Service. ACM, pp 241
106. Wu F-HF, Jang J-SR (2014) On the way to ambient media for sheet music by techniques of information retrieval. In: Proceedings of the IEEE international conference on multimedia and expo workshops (ICMEW). IEEE, pp 1–6
107. Wu F-HF (2016) An evaluation framework of optical music recognition in numbered music notation. In: Proceedings of the IEEE international symposium on multimedia (ISM). IEEE, pp 626–631
108. Yadid-Pecht O, Gerner M, Dvir L, Brutman E, Shimony U (1996) Recognition of handwritten musical notes by a modified neocognitron. Mach Vis Appl 9(2):65–72
109. Yoo J, Kim G, Lee G (2008) Mask matching for low resolution musical note recognition. In: Proceedings of the IEEE international symposium on signal processing and information technology. IEEE, pp 223–226