



# Machine learning and non-machine learning methods in mathematical recognition systems: Two decades' systematic literature review

Sakshi<sup>1</sup> · Vinay Kukreja<sup>2</sup>

Received: 30 January 2023 / Revised: 8 June 2023 / Accepted: 16 July 2023 /

Published online: 23 August 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

Tools based on machine learning (ML) have seen widespread application in the prediction and categorization of mathematical symbols and phrases. The purpose of this work is to conduct a comprehensive analysis of the machine learning and non-machine learning strategies that are currently in use for the recognition of mathematical expressions (MEs). The authors collected and analyzed research studies on the recognition of MEs (and closely related models and issues as well), which are published from January 2000 to December 2022 in the SLR. The review has nominated 98 primary studies out of the extracted 202 studies after heedful filtering using inclusion/exclusion criteria and quality assessment. The pertinent data is derived from IEEE explore, Science Direct, Wiley, Scopus, ACM Digital Library, etc. For assiduously reviewing and synthesizing the data, the authors used grounded theory and other qualitative and quantitative techniques. The analysis reveals that the support vector machine as an ML model with CROHME as the dataset and expression recognition rate as an accuracy metric is frequently used in the chosen studies. Recognition is typically fragmented down into three stages—segmenting symbols, recognizing symbols, and analyzing structures—in non-ML studies. In conclusion, this work aims to synthesize the results of existing research to provide a summary of the state-of-the-art in recognizing handwritten MEs.

**Keywords** Mathematical expression recognition · Handwritten mathematical expressions · Classification · Machine learning · Math expressions · SVM · CNN

---

✉ Vinay Kukreja  
onlyvinaykukreja@gmail.com  
Sakshi  
asakshi541@gmail.com

<sup>1</sup> Department of Computer Science and Applications, Sharda University, Greater Noida, Uttar Pradesh, India

<sup>2</sup> Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

# 1 Introduction

An idea is the origin point of innovation and research. The key initiative towards realizing these ideas is to build and enhance cost-effective and efficacious methods for writing down the documented knowledge into the corresponding electronic format that can be processed by computers and distributed with the help of the Internet. With the numerous expansion in internet users in recent years, the escalating drift of disseminating and exchanging information is consistently done via digital mediums [38]. The categorization and recognition of mathematical expressions (MEs) have become a fascinating and stimulating study area of pattern recognition with endless real-world ramifications due to the fact that MEs represent an essential part of the engineering and scientific literature. [1]. In addition, recognizing handwritten mathematical expressions (also known as HMEs) is a difficult classification task that requires real-time identification of all the symbols that include input as well as the intricate 2D relationships that exist between subexpressions and symbols [119]. These subexpressions can be nested containing Greek and Latin letters, special symbols, and characters. This task of HME recognition also becomes more onerous as it has complicated tasks like structure analysis, symbol recognition, context of MEs unequivocally. This method of verification is complicated enough that it requires additional time to be spent on computational work. Therefore, recognizing them is unquestionably a difficult and laborious task, particularly when attempting to recognize them from a handwritten source of information.

## 1.1 Review gaps and importance

- a. Although researchers carried out plenty of work in the arena of recognizing MEs and symbols, there is a need to systematically collect, compile, and consolidate the recent works in this field. Other reviews [98] and surveys [221] finely presenting the works of HME. Yet none of the literature works systematically reports the studies and judiciously covers all significant empirical instances of literature available on MEs
- b. Unlike the traditional review methods that attempted to present the past works by producing a summarized result form of several studies of the concerned area, the objective of this SLR is to provide a complete possible list of all studies related to the subject area from different research aspects.
- c. To the best of our knowledge, no systematic review focuses on the extensive identification and classification of techniques used for MEs recognition. This is the **first-ever SLR** that aims to be as fair as possible by being auditable [79] and providing transparency to the researchers in this mine to dig and extract for what is left unexplored and unmined.
- d. The entire research methodology used for drafting this study has been presented in detail. Every step involved in the review process has been kept transparent to the readers. For instance, data collection, data selection, answer extraction design; each task involved in research has been vividly depicted.
- e. The uniqueness of the study is the research methodology involved. Apart from strictly adhering to the SLR guidelines, the authors have endeavored to experiment creatively with inter-disciplinary concepts during data synthesis and answer design.
- f. The intended effect of the study is to establish a synthesis of research questions through the use of ground theory, which is a qualitative research methodology. This method has been deployed for the inaugural time in conjunction with systematic literature analysis.

- g. In addition, the sub-processes that make up the recognition process have been dissected in great detail.

A better approach for summarizing the studies and research performed over a period of time for directing the researchers and future aspirants interested in the topic is a real need of the hour. So, this review is planned, conducted, and reported to broadly specify what techniques outperform the rest and where there is a genuine requirement for implementing a new method for recognizing MEs.

## 1.2 Research objectives

This paper conscientiously reviews all the studies between the period between January 2000 and June 2021. To perform the review study, the authors have collated several techniques that belong to multiple computers science domains, like computer vision, digital image processing, and artificial intelligence. This SLR endeavours and attempts to summarise, analyse and methodically assess the empirical evidence regarding

- a) Identification of methods and techniques used in recognition of HMEs.
  - b) Extract out the kind of HMEs used for recognition study.
  - c) Listing out the most frequently used dataset in the research and for empirical analysis.
  - d) Analyzing the accuracy measures and evaluating the accuracy values of the several used methods.
  - e) Focusing on the pure or hybrid techniques implemented in the subprocess analyzing the performance and capability of the applied techniques for recognizing HMEs.
- 
- g) Comparing performance accuracy of contrasting ML techniques to ensure which method outperforms the other techniques belonging to the same header.
  - h) Analyzing the actualization and capability of the applied techniques for recognizing HMEs, i.e., ML versus non-ML or conventional statistical methods.
  - i) Summarizing the set of journals publishing the research on this stimulating research area.

## 1.3 Motivation

The several motivating factors for carrying research in this domain are listed in Table 1.

Apart from the listed factors, the primary factor that motivated and led to the writing of the research review in this domain is the lack of a systematic survey that could compile and extract all the pivotal attributes and determinants like recognition model, datasets, sub-processes, performance metrics, and other meta-data analysis, keeping the selection and synthesis process all crystalline and pellucid at every stage. This facilitates a better readership and provides more crisp insights from bulk literature present on MEs.

## 1.4 Focus of the study

The focus of the study is on the following points:

**Table 1** Motivation factors

Motivation	Description
<i>Practical Motivation</i>	For practical motivation, recognition of HMEs is significant for making an interface between humans and machines. When enormous quantities of resources need to be digitized, it is absolutely necessary to search for automatic document digitization solutions in order to scan the contained equations. These scanned expressions are recognizable in their entirety. In addition, it is beneficial to enter mathematical information into computers by either using HME or taking a picture of an equation. Both of these methods have their advantages. After an equation has been entered into a system, there are many different programs that can handle and run the equation after it has been entered. This allows the equation to be processed effectively.
<i>Scientific Motivation</i>	Recognizing HMEs becomes crucial for digitizing a large quantity of scientific materials for scientific motivation. This number is astounding, and the community of eminent experts is always producing new knowledge to supplement the existing data that has to be represented. It is essential to digitize the data in order to make it possible to search for and access a collection of relevant documents. Additionally, the top professions such as science, research, and practice use computers to access math information. As a result, the digitization of numerical data using the ME recognition approach makes search, access, and retrieval even easier than previously.
<i>Education Motivation</i>	For educational purposes, the ability to recognize handwritten arithmetic notations can be beneficial for teachers and students in information retrieval since it enables the use of a math expression as a search query to obtain a set of pertinent documents.
<i>Social Motivation</i>	In social motivation, these recognition techniques can be competently used for providing brilliant accessibility and approachability to disabled people. It can be noted that an HME or printed ME can be transformed into notations of Braille script or can be recited with the help of text-to-speech tools. So, the recognition of HMEs has commensurable scope in practical, scientific, educational, and social motivation.

- To acquire deeper understanding of the evolving domain of MEs recognition.
- To provide a framework to refer for future research projects by identifying gaps in the recognition of mathematical expression domain.
- To conclude and formulate facts from the meta data and evidences present in the literature.

Thus, the primary focus of this SLR is to provide a comprehensive and unbiased analysis based on the evidences captured from the literature and all the meta data related information that will be extracted in this process. The facts and findings of the study will direct the future research in the domain of MEs recognition.

## 1.5 Research questions

The research questions addressed by this study are tabulated in Table 2.

## 1.6 Paper organization

The organization of the rest of this paper is as follows; the extended section 1.7 briefs about all the critical questions that a beginner needs to know for the understanding of this

**Table 2** Research Questions

Research Questions	Motivation
RQ1. Which ML/non-ML techniques are used in the studies for recognition?	To identify the trends and dominance of techniques used in the recognition process of HMEs.
RQ2. Which datasets are frequently used in the studies?	The availability of a reliable dataset with sufficient data is a fundamental requirement while training a model for recognition.
RQ3. What type of handwritten mathematical expressions are used?	To observe the kind of handwritten math lexes (online/offline) used in the research.
RQ4.a) What metric is used to evaluate the accuracy?	To identify and categorize what metrics have been widely used for evaluating accuracy.
RQ4.b) How much accuracy has been achieved?	To extract the empirical details of accuracy for evaluation of the recognition method used.
RQ5. What approach is followed in the study, or what is the proposed system?	To spot varied approaches used in the recognition process of handwritten mathematical expressions.
RQ6. What kind of techniques is/are used in the sub-process?	For interpreting the optimization procedures carried in the sub-processes of the recognition model.
RQ7. Are there any ML Techniques outperforming other ML techniques?	To identify which ML technique is best among all ML techniques proposed and evaluated to date.
RQ8. Are there any ML techniques that distinctly outperform other non-ML techniques?	To highlight the evaluation details while comparing the performance of ML techniques with non-ML techniques
RQ9 Which are the dominant journals/conference proceedings for papers analyzing handwritten mathematical expressions recognition?	To guide new researchers for research projects.

domain of HMER. In Section 2, the entire research methodology is presented. Section 3 presents the statistical analysis of the HMER related studies. Section 4 contains the results and discussions around the formulated research questions. In Section 5, a variety of results and highlights the extracted facts from the findings are discussed, termed under the heading named Summary and Findings. Section 6 has limitations, and Section 7 holds the calculated conclusions from this study—all references, bibliometric contents, and appendixes that have been quoted at the end. To enhance the readability of the study, the entire road-map to paper organization has been presented in Fig. 1.

## 2 Background

The earliest research work on MEs was in 1965 [15], where Anderson worked on syntax-directed recognition of printed 2D MEs. With the advancement of technology, the interest of researchers shifted from MEs to handwritten mathematical symbols and expressions. The idea is to recognize HMEs and handwritten equations from scanned images or pen-based computing technologies like electronic tablets [38], digital pens, and other gadgets. Before they embark on the review process, the authors have tried answering all possible prerequisite questions briefly to understand the objective of undertaking the research topic. Like HMEs, types, why recognition is necessary, types of recognition, the stepwise procedure in recognition, challenges involved, types of inputs to a mathematical expression recognition system, and problems related to the math recognition system.

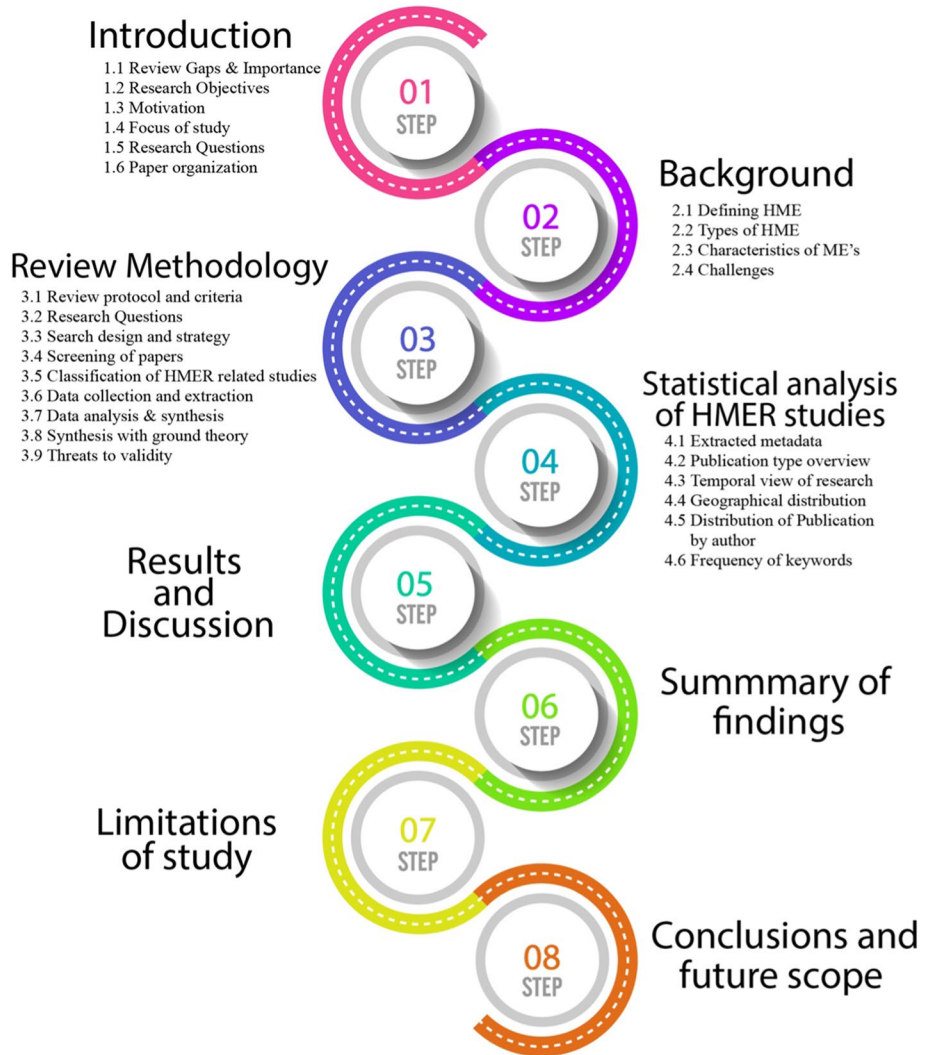


Fig. 1 Roadmap for Paper Organization

## 2.1 Defining handwritten mathematical expressions

The term “mathematical expressions” (MEs) refers to a finite set of symbols that have been ordered in accordance with a rule or experimental study that is connected to some context, most commonly science and research. These MEs consist of symbols like as numerals, operators, constants, variables, functions, parentheses, special characters, and letters (Greek and Latin) arranged in a well-formed order in accordance with some formal propositional norms. As a result, mathematical expressions (MEs) make use of symbols, letters,

and notations in order to deliberately depict various mathematical, scientific, and technological laws or formulas. In addition to this, it is not simply a collection of symbols that have been arranged in a random fashion; rather, it possesses a well-organized structure that is subject to the regulations of the system of mathematical notation [188]. These MEs, when considered in handwritten form, constitute the HMEs.

## 2.2 Types of handwritten mathematical expressions

When it comes to handwritten, MEs can be categorized as offline and online based on inputs mode. According to [145], handwritten data must be converted to digital form with the help of different modes like scanning or writing through special pens on the electronic surface, such as a digitizer combined with a liquid crystal display. The former method we examine the writing on paper constitutes *offline handwriting*; writing produced by finger or digital pen on some electronic device like tablet comprises the *online handwriting*. Similarly, the HMEs are divided into the set of offline/online HMEs. The difference in both online and offline HMEs is tabulated in Table 3.

## 2.3 Characteristics of mathematical expressions

MEs integrate writing letters with drawing a variety of signs and related symbols. The following are the traits of the MEs, according to [172]:

- **Two-Dimensional in nature:** Although they frequently convey implicit meanings, the two-dimensional interactions between symbols are crucial. The two-dimensional relationships of mathematical expressions are examined using TEX's math environment commands and the current MathML standard.
- Inline (Example:  $2x$ )
- Subscript (Example:  $x_2$ )
- Superscript (Example:  $x^2$ )
- Prescript (Example:  ${}^yF$ )
- Enclosed (Example:  $\sqrt{x}$ )

The entirety of the MEs is comprised of symbols that exhibit two-dimensional relationships. A fundamental understanding of mathematics is essential in identifying and comprehending two-dimensional connections.

- **Implicit Semantics:** Certain symbols and letters possess inherent semantic meaning. The determination of interpretations of expressions with implicit operators heavily relies on the identification of symbols involved. Two examples of mathematical expressions are  $f(y-1)$  and  $x(y-1)$ . Concerning the  $f(y-1)$  scenario, it is generally accepted that  $f$  denotes a mathematical function, while  $(y-1)$  serves as the input or argument of the declared function. In the scenario of  $x(y-1)$ , it is important to note that  $x$  is not being used as a function, but rather an implicit multiplication operator is assumed to exist between  $x$  and  $(y-1)$ . Thus, there the symbols used in the equations carry different semantics in varied contexts of usability.

**Table 3** Comparative Analysis of types of Handwritten Mathematical Expression

Offline HMEs	Online HMEs
It primarily focuses on the MEs that are handwritten or in printed form and inputted in scanned data.	It is concerned with MEs that are drawn or written directly on a contact-sensitive surface device such as a data tablet, touch screen, or other similar device.
Offline ME recognition adduces the conversion of handwritten or printed MEs expressed regarding images into editable symbols [12].	Online MEs are figured out in space by arranging points, and they include information about the passage of time as stated by the researchers.
It is a fixed representation of data in the form of material that has been printed out. Since the data, in this case ME, was initially written down on paper before being scanned for input.	It is a simulation of simultaneous contact with the connect system, and it displays the results on monitors, therefore it may be considered a dynamic representation of the data.
Offline recognition is required for this process, and the MEs themselves are represented as bitmaps or pictures.	
Every symbol that originates from ME and is derived from offline sources is converted into a collection of black pixels. Pixels have to be grouped together in order to recognize the overall expression of the image.	It then goes through a process of online recognition, in which MEs are drawn or written on a contact-sensitive device, which then saves the characters as digital inked versions of themselves.
As an offline expression only carries a limited amount of information, the accuracy rate of their recognition is significantly lower when compared to the accuracy rate of online handwritten expressions.	As additional information is equipped in the online handwritten expressions, it can yield better accuracy results than offline handwritten expressions.
In general, this model operates on grayscale images obtained from the camera or any other scanned devices.	The online mode generally stores the data in a sequential form, and it is collected through an interactive interface in a specific manner [57].
When working in offline mode, the image being input does not have any of this information, either in the form of geometric or temporal information. This mode of input is consequently made more difficult to access and is utilized significantly less frequently during the recognition process.	This mode generally takes input in points sequence, and these points have temporal and geometric information. For recognition purposes, temporal information advantage is used that is present in an online input image.
Spatio-luminance analysis is performed for an input image.	Spatio-temporal representation of the input is used.
Complete writing is available only in the form of an image [145].	The availability of the order of two-dimensional coordinates or strokes is there, which is made by the writer.

- Arbitrary Associations:** Considering the fact that there are numerous potential two-dimensional links between symbols and expressions, only a few interactions are permitted depending on the nature of the symbols and the expressions themselves. Mathematicians who are well-versed in the subject are aware of the proper associations and finally put an interpretation on the terms. Whether implicit or not, operators have the ability to link symbols and subexpressions. Linear Prefix ( $x$ ), Infix ( $x + y$ ), Postfix ( $x!$ ), Bounding ( $[, ]$ ), Vertical ( $x y$ ), Implicit ( $\times 2, 2x$ ), Tabular (matrices), and Enclosing ( $x$ ) are examples of several types of operators. The different types of operators can be further divided into unary ( $x!$ ), binary ( $\times 2$ ), and N-ary ( $x + y + z$ ) operators since some operators can only accept a particular amount of parameters. The rules set a limit on the number of arguments and the positions of



those arguments. Prescripts, for instance, are so uncommon in general that only certain symbols might be associated with them.

- **Conventional Dependency:** Conventions regulate how to use mathematical symbols correctly. The conventions to be followed depend on the mathematical specialties and the text's point of origin. The conventions of several branches of mathematics may be followed. As a result, different fields may have different ways of expressing the same mathematical concept. For instance, the imaginary number  $\sqrt{-1}$  is denoted by the letters “i” in calculus texts and “j” in engineering. Their conventions may be adopted by other nations. These conventions and 2D relationships between the symbols are the primary cause of the inherent ambiguities caused by math expressions.
- **Variant Scales:** The symbols and unique operators employed in mathematical calculations are referred to as the different scales of mathematical symbols. One of the fundamental qualities of mathematical symbols that also contributes to ambiguity and issues with expression recognition are the different scales of handwritten math symbols utilized in the creation of HMEs. The scale of the symbol alters its semantics and ultimately differentiates the meaning of the expression.

## 2.4 Challenges caused by the inherent properties of ME

Several challenges are encountered while recognizing mathematical symbols and expressions. These challenges are enough, and sufficient listing of these can be maintained. But the reason for such an ample number of challenges is the inherent properties of the symbols and expressions. Also, HME recognition is challenging due to various writing styles and MEs formats [198]. The authors have tried to list out several challenges belonging to the recognition process of HME, along with the properties that are the cause of these challenges occurring. One of the primary reasons for these challenges is the two-dimensional structure and spatial relationship among the symbols used in the MEs. Table 4 accrues to the clarity about the challenge residing in recognizing handwritten mathematical symbols and expressions due to their inherent properties, which are majorly responsible for ambiguity while recognizing.

### 2.4.1 Other challenges and CROHME

One of the particular challenges involved in the recognition process of HMEs is the lack of sufficient training and testing, especially for academic recognition systems [111]. Due to the lack of a sizeable common dataset of online or offline HMEs which can be used in the recognition process, there has been significantly no central benchmark available to compare the different research works done by various researchers. This non-availability of an open dataset of HMEs is also one of the challenges which forced the researchers to develop their dataset of MEs, which is either the images collected from handwritten expressions by several writers or volunteers or sometimes be the corpus of expressions that are recuperated from prior work of Raman [152], as well as through the mathematical expression base of Aster. So, these datasets collected had a limitation as they tended to contain and cover a subset of math expressions or expressions limited to a specific domain of math. Thus, it is impossible to compare the performance accuracy of different systems as no standard evaluation measure is available until the year 2010. The year 2010 proves to be a remarkable year in the research history of handwritten mathematical symbols and expressions as CROHME, series of competitions was first conducted in 2011.

**Table 4** List of challenges caused due to properties of HME

Properties of HME	Challenges caused
<p><b>Wide variety of symbols, notations, and signs:</b> MEs are not as simple as plain text. It includes a wide range of signs, unique symbols, notations, and special characters. These symbols are not easily drawn as the letters and alphabets of plain text.</p> <p><b>The distinct writing style by different authors and writers:</b> Every person has their own writing style and variation, which makes it more challenging to read and identify the expressions. The expressions employ unique symbolic notations that are even challenging for the human eye to comprehend.</p> <p><b>Varied spatial relationships among the symbols:</b> Spatial relation is how the symbols in the notation are located w.r.t. other symbols in the entire expressions. In the case of MEs, there are several spatial relationships involved, like subscript, superscript, inside, below, and above, etc. These different spatial relations among the symbols and operators lead to confusion and make the process of recognition more challenging and complex.</p> <p><b>Multidisciplinary nature of the expressions and symbols used:</b> Mathematical, symbolic notations and expressions are not just confined to solving of classical mathematical problems, but it holds its roots in the form of different and multidisciplinary mathematical subjects like discrete mathematics, engineering mathematics, and many other types of pure and applied mathematical subjects.</p> <p><b>Notations hold many meanings:</b> The same letter or symbols used in various MEs may hold different meanings, especially when it comes to pure mathematics or engineering maths. In the p-adic theory of representation, German <i>g</i> signifies Lie algebra, plain italic <i>G</i> represents rational points over a field <i>k</i>, bold <i>G</i> suggests a group over an algebraically closed field, and sans-serif <i>G</i> indicates a reductive quotient over the residual field <i>k</i>. [122].</p> <p><b>Contains symbols and characters from several different kinds of literature and languages:</b> MEs consist of special symbols, operators, and letters taken from other literature and languages. Example: <math>\beth, \aleph, \delta, \Delta, \pi, \Pi, \sigma, \Sigma, \Phi</math> these symbols are based on Greek letters whereas symbols resembling <math>\Lambda</math> are grouped with <i>V</i> under Latin letters.</p>	<p><b>Inputting the MEs in the system:</b> The problem is encountered while inputting the mathematical symbols in the system as the special symbols and notation involved in the expressions needs distinct keys for typing the symbols, forming the entire MEs</p> <p><b>A hurdle for recognizing MEs both in the offline and offline form:</b> Due to varied writing styles from one writer to another, it becomes challenging to recognize mathematical symbols used in the expressions.</p> <p><b>More efforts are required to train recognizers:</b> Apart from the spatial features mentioned before, many other spatial features must be recognized well for the proper prediction of the mathematical expression. This task acquires extensive training of the algorithm/classifiers used in the recognition process. Numerous comparable symbols with minimal changes are used in mathematics, and various notations are used to express meaning through oblique spatial relationships between the symbols. To fully capture all such relationships is quite difficult. [204].</p> <p><b>Difficulty in classification and interpretation of MEs according to the different branches of maths:</b> As the math expressions and symbols can belong to other math branches, this can limit the understanding of notations used for the variables in various subjects.</p> <p><b>Leads to confusion in the process of recognition:</b> As the same symbol and notation can be used differently in different mathematical subjects, thus it becomes a challenging task to comprehend their interpretation and meaning. It leads to confusion and difficulty while predicting and recognizing HMEs and symbols. For example, “dx” has different meanings in “<math>\int x^2 dx</math>” and in “<math>cy + dx dx</math>”. Check out how “dx” features in the integral in the first formula but multiplies two variables in the second. In the expressions as mentioned earlier, it can be seen that the same characters have a different meaning under a different context [39].</p> <p><b>Causing ambiguity and challenge in dictionary definition:</b> Greek <math>\nu</math> and Latin <i>V</i> have a close resemblance [122]. This close resemblance in some letters and notations often causes ambiguity, which becomes a problematic concern while recognizing the symbols and expressions of maths. Moreover, these similar symbols also become a challenge while defining the dictionary of symbols used in an expression.</p>

**Table 4** (continued)

Properties of HME	Challenges caused
<p><b>No formal definition of mathematical notations and dialects used in the expression writing:</b> There is no precise and formal math notation, and several dialects are used ([200]; [208]). The users can invent or redefine the maths symbols as when needed, just like natural languages.</p>	<p><b>Overhead while defining the production rules of grammar for the expressions:</b> As mathematics have its language, with formal definitions of notations and dialects used in the expression, it becomes difficult to define the production rules of grammar, which can identify whether the expression is in the correct format or not.</p>
<p><b>The two-dimensional structure of the symbols used in the ME:</b> The symbols used in the ME have spatial relations (superscript, subscript, right, above, below) because of the 2D structures between pairs of symbols used. Sometimes, a mathematical expression's relationships between symbols are based on their relative placements.</p>	<p><b>Causing ambiguity due to distinct positions of symbols in the expression:</b> One of the challenging tasks is identifying the spatial relationships between symbols (the symbol layout), predominantly for handwritten notation [208]. For example, in the expression 'a2a', it looks like that 'a' is multiplied by 2, but 2 is the superscript, i.e., it represents the square of 'a' [38].</p>

Moreover, CROHME is the dataset used by the maximum researchers in the field of HME recognition. The studies of different researchers witness that significantly less work has been done in math recognition with standard encodings, benchmark datasets, or evaluation tools [131]. To start working more easily on handwritten math recognition, the CROHME competition was organized in 2011 and meaningfully compared systems using the common publically available dataset provided by the competition [128].

### 3 Review methodology

This study will comprise the planning, the execution, and the description of the result analysis of the area of research, i.e., recognition of HMEs. According to the [96] guidelines, the authors designed the planning phase, where they apply the review protocol including the seven stages: (1) defining and framing research questions, (2) planning search methodology, (3) search process and criteria, (4) selection of relevant studies, (5) data collection and extraction, (6) data analysis, and (7) results and conclusion evaluation.

#### 3.1 Review protocol and Criteria

In SLR, critical importance is given to a review protocol [196]. The protocol has been developed according to the guidelines set up by [96] such that there is a reduction in researcher bias and there is rigorousness in SLR. After considering the principles, philosophy, and measures of SLR, a comprehensive review protocol has been developed. It mainly focuses on review background, research questions, search strategy, data extraction, quality assessment criteria for the research studies, and data analysis [126]. To differentiate between traditional or narrative literature review and SLR, review protocol plays an important role. It enhances the evaluation consistency and reduces researchers' biases since researchers have to present the search strategy and the criteria for the inclusion or exclusion of any study in the review [96, 126]. All connected series of steps performed in this process have been illustrated by Fig. 2.

### 3.2 Research questions

According to our review protocol, in the first phase, the authors set up a few research questions related to the objectives of our study. These research questions are fundamental tools for digging out the valuable information from the literature already available. So the definition and framing of these research questions are supposed to be done very carefully and critically. The entire direction of this review is entirely based on the framework designed through these research questions. While formulating the research questions, our goal is to assess the empirical evidence resulting from various studies on recognizing MEs using multiple techniques and methods. The authors have selected the research questions to cover all aspects investigated by [39] in their previous survey performed to investigate further issues and perform a better meta-analysis on the chosen studies.

To our knowledge, no SLR has ever been performed in this ME recognition to date, which could give a perspicuous idea to future researchers about the complete analysis of this challenging research area. However, the surveys performed by Tapia ([178]), Zhang [213], and even the recent survey by [222] focused only on online HMEs. Another recent short review by [98] tends to compile the studies comprehensively and not purely evidential of a systematic analysis. At the same time, the authors could find no complete review after 2000 [38] that targeted the systematic research analysis in this recognition survey process. That's why the time range selected for this review is from 2000 to the present date.

#### 3.2.1 Answer extraction design

Figure 3 explains the outline of our search strategy, selection procedure. It focuses majorly on research design for extracting data from some specific research questions to answer the other research questions of this study. According to the mode and kind of extraction used, the authors divided the extraction criterion into three categories: direct extraction, indirect interpretation, and synthesis-based indirect interpretation.

- In direct extraction, the authors retrieved the raw data from the study and used processed information to answer our questions. This criterion is applied to answer the basic research questions, to which answers could be steadily extracted directly from the chosen study. Thus, justifying the name of this criterion.

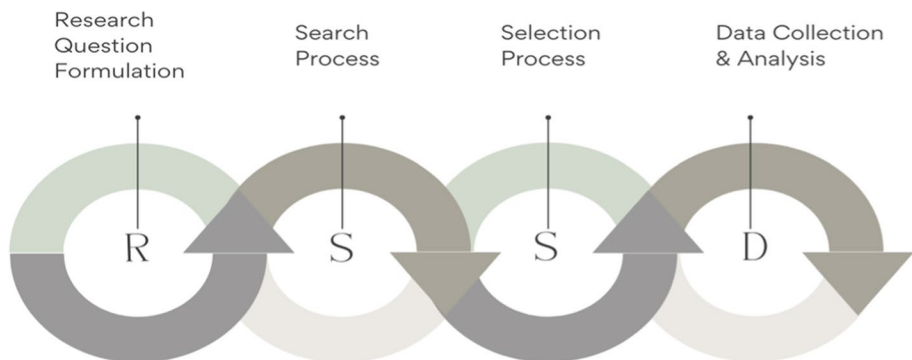
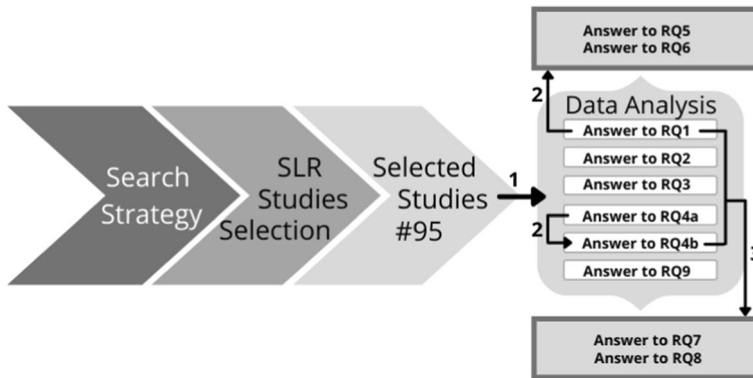


Fig. 2 Systematic review methodology



**Fig. 3** Research Design for extracting answers to Research questions. 1 denotes Direct Extraction, 2 indicates Indirect Interpretation, 3 denotes synthesis based indirect interpretation.

- Indirect interpretation works by using the information retrieved by previous questions. It makes use of this information, draws an interpretation, and answers the formulated research questions. The name given to these criteria is since we are trying to make the information pre-fetched more meaningful and valuable in our study. It uses the answers to a research question to formulate and answer the other research questions.
- Synthesis-based indirect interpretation is an extended version of the above criteria of indirect interpretation. There exists a slight difference between them. It differs from indirect interpretation as it uses the answers or information extracted from two or more questions to answer the other research question of this review study.

Co-relating the criteria to our research questions, we can thoroughly read the study and directly extract the answer to our RQ1 (identifying ML/non-ML technique). It is the instance of direct extraction. Similarly, the extracted information of RQ1 can be used to answer RQ5 and RQ6 (concentrating on the approach used in the processes and sub-processes involved). The answer to these questions comes with the interpretation of the response of RQ1, where the technique used in the recognition process is well analyzed. The answer to RQ7 and RQ8 can be fetched by indirect synthesis-based interpretation, where results of RQ1 and RQ4 can be combined to reach conclusions for both these mentioned RQs. Here, we solely aim to figure the best possible technique, outputting sufficient accuracy. So, the method and the corresponding accuracy values can answer RQ7 and RQ8 (ML outperforms other methods or not).

### 3.3 Search design and strategy

The design of the search strategy encompasses search terms, sources of literature, and the search process. The authors have extracted the significant keywords and key terms essential for formulating the search strings in search terms. The core idea behind making this search string is to extract out all candidate papers of our interest in one go. After considering these research questions, in the second phase of the SLR, a search string is designed to fetch the best of all the relevant literature that is good enough to answer the research questions.

To search for the papers, this step of defining the search terms and constructing a search string is crucial. This creation of a search string to filter the relevant literature of

our interest works iteratively and has been an unbiased strategy for searching appropriate studies. The second step is selecting digital libraries and using the data retrieval settings on these libraries to extract the papers. In the third step, inclusion and exclusion criteria are defined and applied to the research results [32].

### 3.3.1 Search terms

The following steps are used to construct the search term [29]: (a) By analyzing research questions, take out key terms. (b) Figure out the synonyms as well as alternate spellings of the acknowledged key terms. (c) Lookup for keywords in the research literature. (d) Boolean OR will be used in the search string to incorporate synonyms and alternative spellings. (e) Boolean AND will be used to integrate the essential terms. The functions of the significant key include the boolean operators used in the search string with their specific roles in the searching process. The primary critical terms from the questions are *Handwritten, MEs, techniques, ML, recognition, process, online, offline, non-ML*. Then identifying synonyms and spellings, the authors accrued another key search word in the list, i.e., MEs and prediction, used for recognition. Hence, after mustering the essential words in the pool of necessary keywords, the authors embarked on the process of search string formulation using boolean OR and AND. The final resultant search string after this procedure is (*handwritten*) AND (*“mathematical expression” OR “mathematical expressions”*) AND (*classification OR prediction OR recognition*)

### 3.3.2 Data retrieval and literature sources

For performing an automatic search, data retrieval has been performed on digital databases: 1) IEEE Explore 2) ACM Digital Library 3) Science Direct 4) Springer 5) Wiley Online Library 6) Scopus. The search execution using the appropriate search string on each of the mentioned digital libraries and the proper retrieval of the relevant studies and papers constituted the entire data retrieval process. Though an adequate amount of data is well retrieved from the above digital sources, the authors did not choose the other useful literature in magazines, books, and articles for the literature review. The concepts discoursed in these sources are not subject to review. Thus their quality can't be reliably corroborated.

After proper execution of automatic search for data retrieval, a manual search, our secondary search phase, is accomplished to ensure that we don't miss out on anything worthwhile. To fulfill this motive, the authors performed the manual search by forward and backward referencing. The authors iteratively referred to references from the retrieved studies to extract more relevant reviews from the past. This iterative process applied in the secondary search phase is called Snowballing. After snowballing, the extracted studies are added to the Mendeley library, which further helped to make suitable references in this study. The studies that are focused on are ranging from the time 2000 to 2021.

## 3.4 Screening of papers and the process of filtration of studies

When the constructed search string is run on the digital libraries, the search string fetches the different number of studies from various databases, respectively.

In screening papers, appropriate studies are selected, and this selection is based on well-defined criteria and research themes. The well-defined criteria of selection include four stages, as per the headings of Table 5. The first filter applied here is the year-wise

**Table 5** Count of studies at different stages of the refining process

Digital Databases	Total no. of studies retrieved	After year-wise filtration	After removing duplicate studies	After reading the title relevance to the selected topic	After reading the abstract
IEEE Explore	118	109	92	86	70
ACM Digital Library	65	55	42	6	4
Science Direct	51	40	33	18	15
Springer	82	70	57	38	21
Wiley Online Library	35	24	19	3	1
Scopus	159	148	141	120	91
Total studies selected	510	446	384	254	202

filtration that has a constrain of considering the studies from the year 2000. The second filter focuses on the removal of duplicate studies. The chances of redundancy of papers arouse owing to the fact that articles have been extracted from major digital databases as well as Scopus. The Scopus includes almost all articles from IEEE, ACM, Springer and Wiley. The reason why Scopus has been exclusively searched for the inclusion of articles is that the authors don't want to miss any quality publication which is published other than the mentioned digital databases. The criteria used for removing duplication are '*Exact Match*' (where the titles of all the gathered studies were compared and removed in case of any exact duplicates) and '*Cross Checking*' (where the authors, publication dates, and other bibliographic information has been cross-verified to identify potential duplicates from different sources or variations in title wording). The third filter is removal after reading the title. The search string will target to fetch the studies with the mentioned keywords in the search string.

One such example is the study titled "Strategy and Tools for Collecting and Annotating Handwritten Descriptive Answers for Developing Automatic and Semi-Automatic Marking - An Initial Effort to Math".

This study is extracted as it contains the keyword 'handwritten' and 'math'. The paper is published in the 2019 International Conference on Document Analysis and Recognition Workshops. So, because of the keyword 'recognition' also, the study has been extracted. On reading the title itself, the study is realized to not be relevant to the topic of research. Hence, removed based on the title filter. Thus, the criteria for removal-based titles include: '*Relevance*' (where titles that appeared to be directly related to the topic or research question were considered) and '*Focus*' (where titles that indicated a clear focus on the specific aspects or associated research variables). This helped to narrow down the studies to those most aligned with the research objectives. Similarly, in the case of the fourth filter, the study has been removed after analysing the relevance of the abstract. The criteria for removal of abstracts include: "*Study Objective*" (where abstracts that clearly outlined the objective or purpose of the research conducted in the study with direct mapping to our research concerns were considered, "*Methodology*" (where abstracts that briefly described the methodology employed in the study, ensuring that it aligned with the research interest of this study were prioritized). And the third criteria is "*Findings*" (where abstracts that summarized the main findings or results obtained from the study were analysed as they assess the potential relevance and contribution of their research to this study).

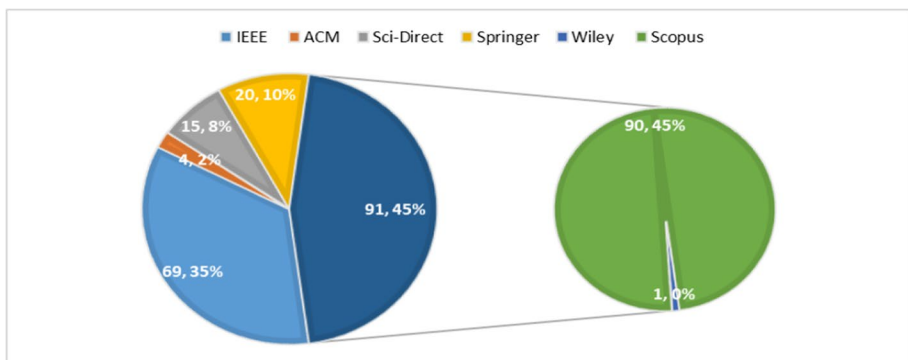
On performing analysis of the contribution of each digital library in all final count of studies selected, the authors noticed that more of the quality papers are majorly from the two digital databases involved, i.e., IEEE explore and Scopus. And Scopus ranked highest in terms of its contribution to the final set of selected papers. Below, Fig. 4 shows the analysis of the contribution of each digital library in the definitive collection of selected studies.

This aspect of the analysis of selected studies based on the relevance ratio and relevance percentage will allow us to analyze and identify what sources have retrieved the best of relevant results on the execution of the search string. The estimated values of relevance percentage show that IEEE explore and Scopus digital databases have higher relevance percentages than others. The conclusion drawn is that Scopus had the highest relevance among all digital databases compared. In contrast, Wiley had the least relevance as per the calculated values mentioned in Table 6.

### 3.5 Classification of HMER related studies

This section figures out the review findings and presents conclusions, consisting of results obtained by analyzing the selected studies for each research question formulated. The authors have conclusions discussed and answered the research questions in separate subsections by considering sufficient empirical shreds of evidence collected from the data. In the discussion, the findings are placed in a broader perspective that is nonetheless closely tied to the study issues.

At the end of this entire screening process, 28.94% of the total initial papers are identified, which, when summed up with studies selected from manual search, resulted in 202 articles. These 202 studies are collected, and their abstracts are read; the authors could classify the entire cluster assembled and allocate them in empirical categories according to the research approach. The nominated headers of the classification are presented in Table 7 as (a) generic (b) technique based (c) application-specific (d) sub-process concentrated (e) survey-based study (f) particular script oriented (g) CROHME winning studies, and (h) other evaluation and problem addressing papers. The studies of interest will include more



**Fig. 4** Analysis of the contribution of each digital library to the selected database  
**RELEVANCE RATIO** =  $\frac{\text{number of selected studies}}{\text{total number of retrieved studies}} = \text{NS} : \text{NR}$



**Table 6** Representation of Relevance Ratio and Relevance Percentage of each study

Digital Library	Relevance Ratio	Relevance percent
IEEE explore	70:118	59.32%
ACM	4:65	6.15%
Science Direct	15:51	29.41%
Springer	21:82	25.60%
Wiley	1:35	2.85%
Scopus	92:159	57.86%

articles selected from (a), (b), and (g) categories. The count of studies under each classified category is vividly depicted through Fig. 5.

### 3.5.1 Inclusion and exclusion criteria: (Primary Selection Phase)

Inclusion-Exclusion principles are the significant filters for selecting potential studies from the candidate studies retrieved after the screening process of the papers. Subsequently, many of the candidate papers lack support to address the research questions that have been raised by this current study; further filtration is needed to boost the relevance of studies collected for our review study. That's what our selection process precisely aims to do. The inclusion-exclusion parameters used in the primary selection phase are tabulated in Table 8.

These inclusion/exclusion criteria parameters are selected after numerous meetings between the authors, finalized with mutual consent.

### 3.5.2 Quality assessment criterion: (Secondary Selection Phase)

This is a pure quality check measure on the selected studies, aiming to satisfy the defined standard of the quality of the chosen studies. Some quality assessment questions are formulated to weigh the candidate studies so that the final selection of studies can be made, as mentioned in Appendix 2 Table 23. Note that the studies with a low quality imply that those with quality measures that weigh less than satisfied thresholds are supposed to be excluded from the cluster of the selected studies. A fuzzy linguistics idea is used from an SLR study by [4]. These fuzzy linguistics factors were utilized by the researchers to rank various quality assessment queries. Rather than just assigning scores on a binary scale of 0 and 1, these fuzzy variables are used to appropriately measure the score (relevance) of the survey in regard to the aforementioned assessment questions. That's why they wisely employed the fuzzy linguistics variables, avoiding the binary scale.

The score chart for assessment of the quality of the studies is shown in Table 9. The above score chart depicts the score allocation to the studies during the quality assessment criteria. In particular, the authors have organized the scores for checking and validating the relevance of each research according to the formulated question list for this review. Rating 0 depicts the study has no significance according to the objectives of this study. The score between the values from 0.1 to 0.9 scales from rarely, i.e., of less relevance, partially, i.e., mediocre relevancy, mainly implies the study is highly significant. The highest score of 1 depicts that the study is answerable to all the quality check queries designed for accessing the quality of the research individually. Since we have formulated seven questions

**Table 7** Defining the nominated headers of classification of studies

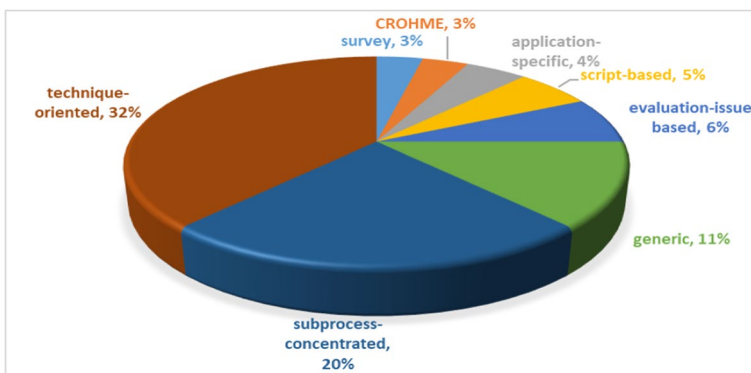
Classification	Description
(a) Generic	This category of classification included all general studies, which included not such a specific technique or sub-process. The studies selected for this category included studies related to general analysis ([1]; [18]), addressing new architectures (B.-Q. Q. [190]), general recognition methods, and tools. According to our classification, these studies constituted 11.5% of the collected studies, as shown in Fig. 5.
(b) Technique-based	The technique-based header included the papers and studies that primarily focused on proposing a new technology or model to recognize MEs. The authors reviewed the selected studies and found that many studies focused on different techniques classified in either ML or non-ML-based study as per the first research question (whether the review uses ML or non-ML recognition process). According to our classification, these studies constituted the highest portion in the graph by occupying 32% approximately a collected chunk of selected studies, as shown in Fig. 5. Classifying the technique-oriented studies from the selected cluster is necessary, and most of the answers to our research questions would be extracted from these studies.
(c) Application-specific:	This header of the subdivision of the selected studies aims to identify those whose works are dedicatedly concentrated on some specific applications or tools of recognition. This subcategory constituted about 20% of the total considered studies.
(d) Subprocess concentrated	This particular category of classification study targets identifying whether the study is focused on some specific subprocess in the recognition process. As it is found, many papers worked on improving the subprocesses involved in the recognition phenomenon by using some of the other techniques and mostly gravitated towards the subprocesses involved. Our review studies also included such papers, which are subprocesses centered and worked on enhancement of recognition system by explicitly working well on the implicit subprocesses of the system. These studies also constituted around 21% of the total selected studies for screening
(e) Survey-based	As the name of the category is defined, it aims to classify those survey-based studies or a review paper like ours. Extracting such studies is also crucial as these studies prove significant studies for directing beginners in the field. These works are of central concern as they try to objectify well what dimensions in the area are already being worked upon and what is yet to be explored. According to our findings in the classification, these studies are almost 3% of the entire retrieved studies.
(f) Script oriented	It has been observed while extracting the papers during the search process that a number of the studies are focusing on expressions from different languages and scripts. Though these studies are less of our concern, yet these studies are segregated into this category. These studies dealt with expressions from Arabic [17], Chinese [44], Bangladesh [25], Devanagri [16], and Gurmukhi scripts.

**Table 7** (continued)

Classification	Description
(g) CROHME	<p>Since 2011, a series of competitions for online recognition of HMEs has been conducted to bring all the researchers in this field to a common platform that can compare the systems using varied techniques for the same recognition process. The papers that are on the winning and excellent methods proposed in this CROHME competition have been considered.</p> <p>CROHME provides the platform to compete and evaluate the different proposed systems and has contributed well to the standardization of datasets used in the system. As earlier, every researcher developed their dataset for the recognition process, making it hard to compare different techniques as no common public dataset was available before. The CROHME series has equipped the practitioners in the arena, which is a standard common public dataset.</p>
(h) Evaluation and Problem centric	The studies included in this category are different. Many revolved around fixing or improving existing problems in the recognition system, whereas some focused on comparing a few systems and methods used in the process. These studies are only considered after reading a few of the pages of the entire case.

for assessing the quality, the summed score for a single study can be shown in Table 10. Table 10 is sculpted after numerous rounds of discussions and attentive observations made from the previous studies and works connected to quality evaluation criteria; both writers agreed on the notion of the assessment score based exclusively on the fuzzy linguistics variable.

The whole concept is concentrated on accessing the quality of the different selected studies. The studies with very extreme low scores (mainly with relevance ‘NO’ and ‘Rarely’) are excluded in this secondary selection phase of the review methodology. The apparent reason for this exclusion is the low scores against the quality assessment criteria questions. For instance, if a study receives the score for the quality assessment questions as Q1 marks 0.2, Q2 marks 1.4, Q3 marks 0.2, Q4 marks 0.3, Q5 marks 0.6, Q6 marks 0.4, Q7 marks 0.1, Q8 marks 0.1, and the total score is 3.3 which lies in the range of ‘rarely’ (refer Table 10). It implies that the study doesn’t qualify for the quality assessment procedure, so it cannot be included in selected studies.

**Fig. 5** Classification of papers based on the defined meta-data categories

**Table 8** Inclusion Criteria (Primary Selection Phase)

Inclusion Principles	Exclusion Principles
Studies ranging from 2000 to 2020 are included	Studies that focus on HME recognition but don't specify any ML or non-ML techniques are excluded.
Studies that are technique centric, generic, subprocess oriented, and some specific issues addressing studies are included	Studies based on recognizing mathematical symbols, digits, or characters, in particular, are excluded.
Studies that hold more experimentation and validation of results are included.	Studies that target some specific hardware application or so are excluded
Only the journal version is included if studies are published both in journals and conferences	Studies that stress only on the subprocesses used in the recognition are excluded
Studies that used ML techniques are prioritized over others.	Studies that addressed MEs from the specific script (e.g., Arabic or Chinese) or language are excluded Review studies and survey papers are just used for reference but are excluded from the review pool of candidate papers. Studies that don't address the raised questions of this study are excluded.

**Table 9** Score allocation according to relevance

Score	Relevance measure
0.0–0.2	No
0.3–0.5	Rarely
0.6–0.8	Partly
0.9–1.0	Yes

**Table 10** Total score for each study with relevance scale

Total score	Relevance
0.0–1.7	No
1.8–3.5	Rarely
3.6–5.3	Partly
5.4–8.0	Yes

On the contrary, if a study qualifies the defined quality criteria, satisfactorily scores more than 3.5, and reaches the scale interval named ‘Partly’ according to the relevance measure, the candidate study is confirmed to join the selected list of the studies for consideration and review. After extensive quality assessing procedures, 98 studies were chosen, and the rest were rejected as per the selection protocol. It should be noted that the studies ranking with relevance measures as ‘partly’ and ‘yes’ are selected for further review. These scores determine the studies to be of high relevance and quality for assessing the condition.

While analyzing recent studies, one of the matters of grave concern is that fewer citations are discovered, highlighting that the most recent studies carry fewer citations. There is a decent possibility that the research study can be relevant as per other questions of assessment criteria. So, considering this exception, the publications ranging from the year 2017–2021 are excluded as per Q8, i.e., inspecting the citation count. In this case, the study based on other relevant measures is included and according to the journal’s h-index and impact factor in the journal. Appendix 1 Table 22 shows the studies being grouped according to quality assessment labels.

The quality assessment criterion is applied to 198 studies out of 209 studies selected from the previous selection phase. Eleven of the considered studies are redundant when the entire database of chosen studies is compared. These 198 studies are considered for the quality assessment criterion, where each study is selected for further complete review by evaluating the studies according to the designed quality assessment questions. The qualifying range in quality assessment lies between 3.6–8.0. Any study scoring relevance scores less than 3.5 is not considered in the final database of selected studies. Thus, in this way, the potential studies are chosen for the complete review by applying quality assessment criteria. Additionally, the authors have decided to analyze the count of the studies for the particular relevance measure and count and segregate the studies as per the citation rate observed during the selection process.

Approximately 49% of the studies are selected, constituting 98 studies out of 202 studies. Overall, 97% of studies had at least one citation (note that studies ranging from 2017 to 2021 are not considered for citation check). Overall, as per the analysis, 32.32% of studies are utterly rejected and counted 64 (No), a good percentage of 19.6% (39 studies) are rarely relevant, and 14.64% of studies are of average relevance. Almost 33.33% of the studies are highly relevant and constituted the count of 66 studies for relevance measure ‘Yes.’ In this way, the entire criterion of the secondary phase is accomplished while leveraging the studies as per their relevance.

### 3.6 Data collection and extraction

Extracting the quality and quantity of information from the selected database is one of the prime objectives of this systematic study. After compiling the studies chosen from the well-defined selection phase, the authors have decided to extract metadata from the database of the final selected studies of this review. Table 11 shows the chosen studies, review analysis. The authors have tried to classify the investigations on several grounds to avoid missing any critical parameter or any vital perspective of the review that could help understand the present status of the research literature yielded to today.

**Table 11** Framework of review analysis (investigating each study against the formulated research questions)

Paper Title	References	RQ1	RQ2	RQ3	RQ4a	RQ4b	RQ5	RQ6	RQ7	RQ8	RQ9
S1	[57]	yes	yes	yes	yes	yes	yes	–	yes	–	yes
S2	[198]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
S3	[135]	yes	yes	yes	yes	yes	yes	–	yes	–	yes
S4	[191]	yes	yes	yes	yes	yes	–	yes	yes	–	yes
S5	([219]	yes	yes	–	yes	yes	–	yes	yes	–	yes
S6	[112]	yes	yes	yes	yes	yes	–	yes	–	–	yes
S7	[52]	yes	yes	yes	yes	yes	yes	–	yes	yes	yes
S8	[217],	yes	yes	yes	yes	yes	yes	–	yes	yes	yes
S9	[218]	yes	yes	yes	yes	yes	yes	–	yes	yes	yes
S10	[54]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
S11	[111]	yes	yes	yes	–	–	–	–	yes	yes	yes
S12	[187]	yes	yes	yes	yes	yes	–	–	–	–	yes
S13	[218]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S14	[142]	yes	yes	yes	yes	yes	–	yes	yes	–	yes
S15	[216]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
S16	[109]	yes	yes	yes	yes	yes	–	yes	yes	–	yes
S17	[167]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S18	[216]	yes	yes	yes	yes	yes	–	yes	yes	–	yes
S19	[216]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S20	[12]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S21	[51]	yes	yes	yes	yes	yes	–	yes	yes	yes	yes
S22	[108]	yes	yes	yes	yes	yes	yes	–	yes	yes	yes
S23	[215]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S24	[36]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S25	[171]	yes	yes	yes	yes	yes	yes	–	yes	–	yes
S26	[119]	yes	yes	–	yes	–	–	yes	yes	–	yes
S27	[89]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S28	[27]	yes	yes	yes	yes	yes	–	–	–	–	yes
S29	[140],	yes	yes	yes	yes	yes	–	–	yes	–	yes
S30	[128]	yes	yes	–	yes	yes	–	–	yes	yes	yes
S31	[10]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S32	[22]	yes	yes	yes	yes	yes	yes	–	yes	yes	–
S33	[110]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
S34	[107]	yes	yes	yes	yes	yes	yes	–	yes	yes	yes
S35	[108]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S36	[83]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S37	(Davila et al., 2014)	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S38	[11]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S39	[170],	yes	yes	yes	yes	yes	yes	–	yes	–	yes
S40	,[144]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S41	[173],	yes	yes	yes	yes	yes	–	–	yes	–	yes
S42	[46]	yes	yes	–	yes	yes	–	–	yes	–	yes
S43	[72],	yes	yes	yes	yes	yes	–	–	yes	–	yes
S44	[86]	yes	yes	yes	yes	yes	–	–	–	–	yes

**Table 11** (continued)

Paper Title	References	RQ1	RQ2	RQ3	RQ4a	RQ4b	RQ5	RQ6	RQ7	RQ8	RQ9
S45	[23]	yes	yes	yes	–	–	–	–	yes	–	yes
S46	[127]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S47	[102]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S48	[114],	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
S49	[82]	yes	yes	yes	yes	yes	–	–	–	yes	yes
S50	[169]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S51	[154]	yes	yes	yes	yes	yes	–	–	–	–	yes
S52	[7]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S53	[8]	yes	yes	yes	yes	yes	–	–	–	–	yes
S54	[35]	yes	–	yes	–	–	yes	yes	–	–	yes
S55	[21],	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S56	[20]	yes	yes	yes	yes	yes	–	–	yes	–	–
S57	[24]	yes	yes	yes	–	–	–	–	–	–	yes
S58	[118]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S59	[18]	yes	yes	yes	yes	yes	–	–	yes	–	–
S60	[157]	yes	yes	yes	yes	yes	–	–	–	–	yes
S61	[64]	yes	yes	yes	yes	yes	yes	–	yes	yes	yes
S62	[189]	yes	yes	yes	yes	yes	–	–	–	–	yes
S63	[148]	yes	yes	yes	yes	yes	–	–	–	–	yes
S64	[149]	yes	yes	yes	–	yes	yes	yes	–	–	yes
S65	[14]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S66	[122]	yes	yes	yes	yes	yes	yes	–	–	–	yes
S67	[164]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S68	[13]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S69	[104]	yes	yes	yes	yes	yes	–	–	–	–	yes
S70	[85]	yes	yes	–	yes	yes	–	–	yes	yes	yes
S71	[201]	yes	yes	yes	–	yes	–	–	–	–	yes
S72	[153]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S73	[61]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S74	[84]	yes	yes	yes	–	yes	–	–	–	–	yes
S75	[184]	yes	yes	yes	yes	yes	–	–	yes	yes	yes
S76	[69]	yes	yes	yes	yes	yes	–	–	–	–	yes
S77	[177]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S78	[60]	yes	–	yes	–	–	yes	yes	–	–	yes
S79	[95]	yes	–	yes	yes	–	yes	yes	–	–	yes
S80	[70]	yes	yes	yes	yes	yes	–	–	–	–	yes
S81	[180]	yes	–	yes	–	–	yes	yes	–	–	yes
S82	[179]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S83	[185]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S84	[3]	yes	yes	yes	–	–	–	–	yes	–	yes
S85	[179]	yes	–	yes	–	–	–	–	yes	–	yes
S86	[208]	yes	yes	yes	–	–	–	–	yes	–	yes
S87	[39]	yes	yes	yes	–	–	yes	yes	yes	–	yes
S88	[55]	yes	yes	yes	–	yes	–	–	–	–	yes

**Table 11** (continued)

Paper Title	References	RQ1	RQ2	RQ3	RQ4a	RQ4b	RQ5	RQ6	RQ7	RQ8	RQ9
S89	[90],	yes	yes	yes	yes	–	–	–	–	–	yes
S90	[38]	yes	yes	yes	yes	yes	yes	yes	yes	–	yes
S91	[206]	yes	yes	yes	–	–	–	–	yes	–	yes
S92	[43]	yes	yes	yes	–	–	–	–	yes	–	yes
S93	[91]	yes	yes	yes	–	–	yes	yes	yes	–	yes
S94	(T. [174])	yes	yes	yes	–	yes	yes	–	–	–	yes
S95	[168]	yes	yes	yes	yes	yes	yes	–	–	–	yes
S96	[94]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S97	[199]	yes	yes	yes	yes	yes	–	–	yes	–	yes
S98	[74]	yes	yes	yes	yes	yes	–	–	yes	–	–

### 3.7 Data analysis and synthesis

In this section, data analysis and synthesis are to be performed. The principle goal of this phase of data analysis and synthesis is to identify, select, aggregate, and analyze the shreds of evidence collocated from the chosen papers for answering the formulated research questions. A single piece of evidence might have a small evidence force, but the aggregation of many of them can make a point stronger [139]. Thus, evidence collection and synthesis become truly indispensable part as it helps draw conclusions, which are the tangible outcome of a good review paper. Table 12 shows the review analysis framework.

To extract answers to the research questions, the selected studies are analyzed and synthesized quantitatively (e.g., estimated accuracy of different recognition techniques) and qualitatively (e.g., focused on differentiating separate ML and non-ML methods, recording their accuracy metrics, evaluating their strengths and weaknesses). The authors employed various strategies to synthesize the extracted data associated with different kinds of research questions.

The different synthesis strategies conforming to the research questions are depicted in Fig. 6.

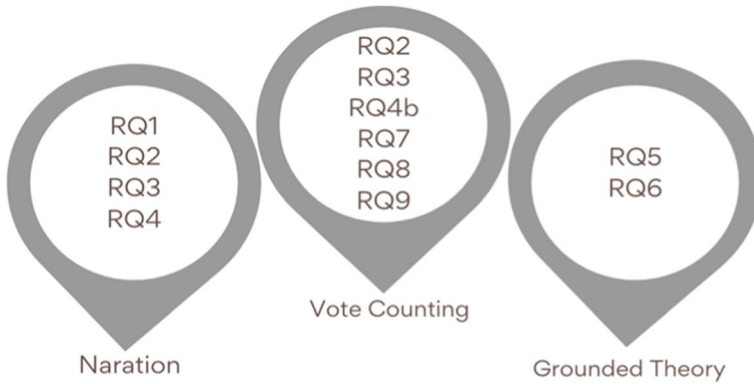
The narrative synthesis method is applied for the data pertaining to RQ1, RQ2, RQ3, RQ4a. That is, the data are tabulated in a regular and organized way consistent with the formulated research questions. To represent the extracted information, some visualization tools, including bar charts, pie charts, and other graphs, are used. These graphical representations enriched and glorified the display and presentation of the distribution of ML techniques, work on offline/online HMEs, the frequently used dataset, and their estimation accuracy data. Also, the survey studies done to date are analyzed and synthesized using this strategy.

For the data pertinent to RQ2, RQ3, RQ4a, and RQ9, which focus on comparing different techniques, their estimation accuracy, the other datasets used, kind of expressions used, the vote-counting method is used. Suppose we want to synthesize the comparisons between what types of math expressions are primarily used by researchers for recognition purposes. We can compare the number of studies using offline and the number of studies voting online. In this way, we can obtain a brief idea regarding which kind of math glyphs are frequently used. Also, we can use this voting synthesis



**Table 12** The framework of Review Analysis

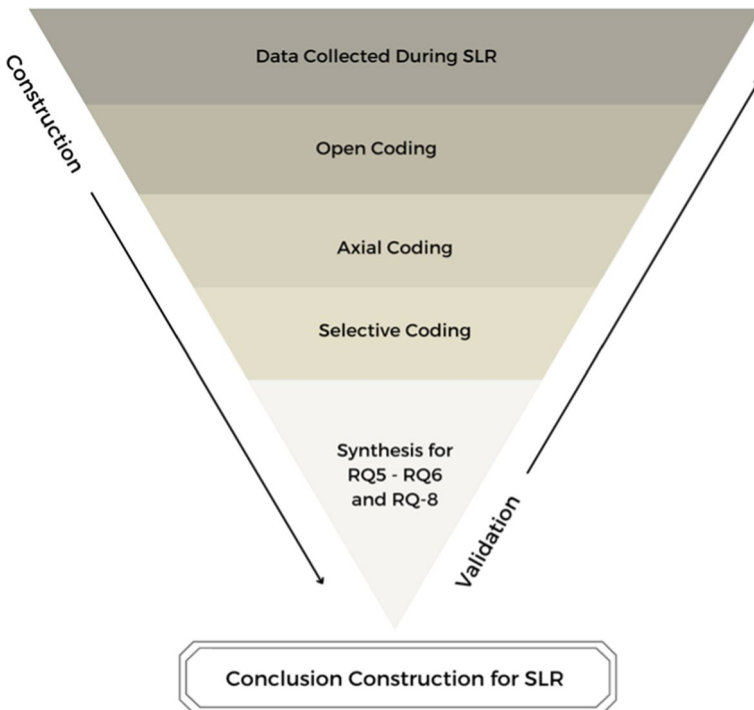
The framework of the Review Analysis	
Parameter	Requirement
Based on the approach used	Distinct approaches used in different papers are analyzed, resulting in the formation of different categories based on the theme of the study.
Based on the year of publication	Studies of literature are divided into chunks according to the period based on the year of publication.
Based on the citations	Dividing the content collected based on the number of times, the study is cited by other researchers.
Based on the kind of technique ML/non-ML	Objectifies to know the kind of technique used in the recognition process.
Based on the type of HMEs	Distinguishing whether the work is done on online handwritten or offline expressions
Based on the empirical or theoretical aspects	Where the topic has been studied theoretically, or some experimentation is involved.
Based on the publication source	Whether the study is from a journal, conference, workshop, symposium, or a magazine article



**Fig. 6** Different synthesis strategies

to estimate what accuracy metrics are more regularly used for recognition by different techniques and compare them. This voting synthesis will also help analyze and compare the ML models used for the recognition task.

For the data pertaining to RQ5 and RQ6, the concept of grounded theory is used, one of the systematic methodologies in the social sciences involving the construction of theories through methodological gathering and data analysis. But we have tried to



**Fig. 7** Conclusion construction of SLR using Grounded Theory

use this in the engineering science review to analyze the qualitative data of the different approaches used to recognize the math expressions. The basic idea of the GT approach is to read, revise and review the textual data (such as the approach evidence used by different techniques and label the variables (called group, concepts, and properties) and their interrelationships. This grounded theory has three phases, explained and illustrated in Fig. 7.

### 3.8 Synthesis with ground theory

**Open coding** The essential part of the data analysis phase is open coding. It uses techniques like identifying, categorizing, naming, and describing the approach found in the text, i.e., from the initial data collection, the researchers' categorize the information about the incidents [30, 48, 50]. The authors have implemented open coding to classify the different methods, approaches used in the recognition process, and sub-process to consolidate all the techniques used. In review, overall, 98 open codes have been found from the SLR, and those are related to several approaches used for the task of recognizing HMEs.

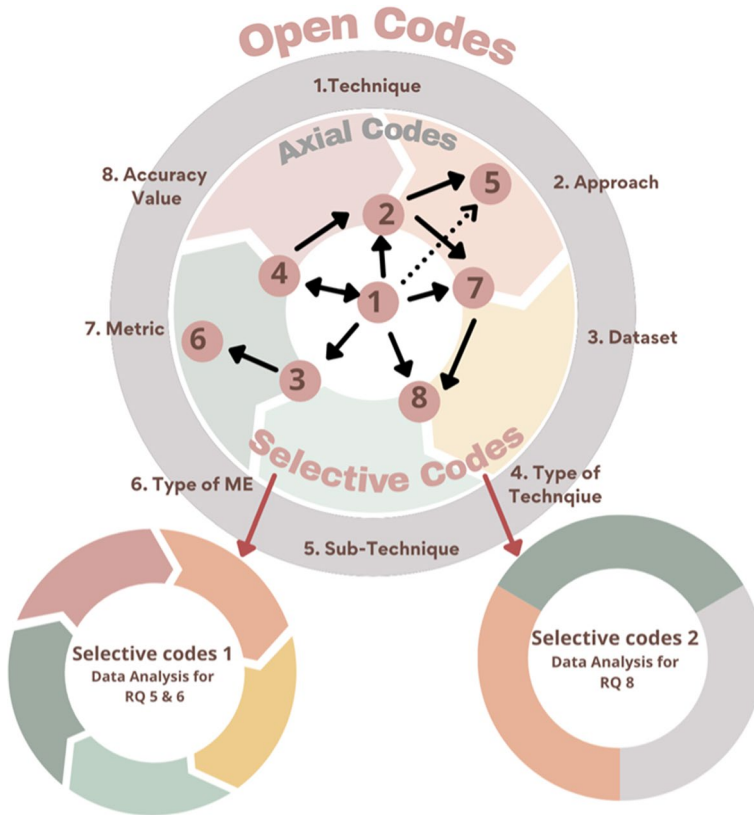
**Axial coding** It is the procedure of developing inter-relationships between the open codes (properties and groups) via a combination of deductive and inductive thinking. It includes collecting open codes together and the similar ones are confined into distinct axial coding groups. This mechanism is not just time-saving but also lessens the overhead of searching and establishing relations for entire relations in the set [30, 48, 50, 73]. The reviewers have categorized all open codes into eight principal axial codes, also known as concepts. And then, we build the interconnections among the collected open codes, and In axial coding, we defined the inter-group relationships between the chosen open codes. The solid line arrows in Fig. 8 depict the direct connections between the categorized open codes, whereas the broken line arrow differentiates indirect relationships from direct ones. And the double-headed arrows represent the attributes identified through open coding that can be derivable from each other.

**Selective coding** It is the process of selecting one core category and then relates all the further groups to that core category [30, 48]. The initial idea is to develop a single action around which everything else is covered. The authors segregated the appropriate axial chain coding from the prefabricated axial code chains to synthesize the research questions. The selective coding thus refines the synthesis procedure and assists in the consistent framing of relevant code chains that helps for qualitatively analyzing the research questions and extracting satisfying research answers. Figure 8 illustrates the refining process involved in selective coding.

### 3.9 Threat to the validity

#### 3.9.1 Limitations of search string

Though sincere efforts have been made to formulate the search string, creating a search string also holds certain restrictions owing to the count of keywords allowed in the search string. Most of the digital libraries don't support lengthier search strings containing too many keywords. In other words, it can be said that there is a limitation or restriction of



**Fig. 8** Implementing phases of Grounded theory on the attributes of the study

picking the keywords to be used in the search string. For example, a string like [(offline OR handwritten) AND (“mathematical expression” OR “mathematical expressions”) AND (regression OR “machine learning” OR classification OR “Bayesian network” OR “neural network” OR “decision tree” OR “support vector machine” OR “genetic algorithms” OR “random forest” OR “deep learning”)], won’t execute well.

### 3.9.2 Selection bias

The search string is created and used by the authors for selecting the studies for this SLR. They had tried their best to choose entirely appropriate phrases in their search string. It must be noted that the selected keywords for the formulation of the search string are picked according to the research questions initially formulated. It may be possible that authors might have missed some relevant studies as there are chances that some studies may have omitted the primary keywords that the authors chose. These essential keywords might not be mentioned in the frequently overlooked sections like title, abstract, and keywords. Though rigorous trials are taken to avoid any such setback, including the manual search

**Table 13** Metadata themes with description

Metadata theme	
Publications by publication type	It sorts the chosen articles by their type, i.e., journal, conference, or others
Publication by publishing year	It depicts the number of research articles published in a year.
Publication by geography and year	It depicts the number of research studies by each country per year.
Publication by author	It segregates the selected studies according to their bibliographic references of authors.
Publications by keyword	It extracts the keywords from the chosen articles and measures the count of common keywords per publication.

criteria is decided. Referring to the bibliographies of the various studies allows for the selection of some research that may be of interest to the reader. As a result, the initiative is still being pursued; there is a possibility that we may have missed some substantial studies that could be interpreted as posing a risk.

## 4 Statistical analysis of the selected HMER related studies

In this section, the statistical results of the selected studies will be presented concerning their publication type, publication year, geographical distribution over years, authors, and keyword count status.

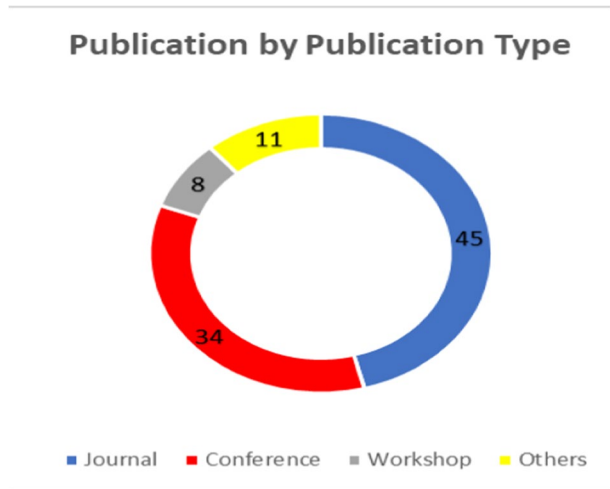
### 4.1 Extracted metadata fields from selected studies

The list of the extracted fields/ heads for metadata analysis of the selected studies is mentioned in Table 13.

### 4.2 Publication type overview

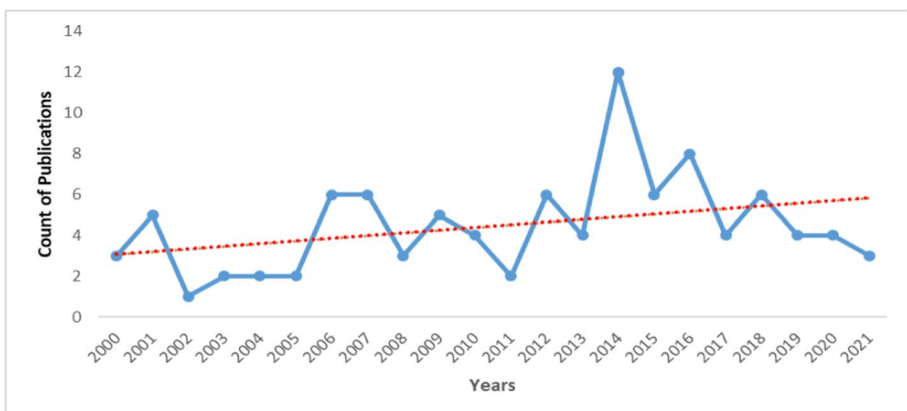
The authors have tried to identify the sources which have contributed significantly high to the collection of studies. The purpose of conducting this analysis from this vantage point is to identify the foundation of the pertinent studies. This will allow these high-ranking names of sources to be shortlisted for the purpose of preferring them as secondary sources of literature for the subject matter of this review article and any further implementations that may be carried out. Additionally, this study offers a new dimension of priority for future researchers to identify the sources for reference and an outstanding supply of literature by functioning as a secondary database for the purpose of comprehending this research subject. Figure 9 shows the publications concerning publication type. It is observed that the majority of the selected studies belong to journals and conferences, ranging from about a total of 77 studies out of a total of 95 selected studies. Other studies included some papers in the proceedings of some workshops, a well-framed thesis, and organized symposiums.

**Fig. 9** Distribution of selected studies according to publication type



### 4.3 Temporal view of research over the years

The count of selected studies is analyzed well on the scale of years, depicting the number of publications produced per year. It is noticed that the average count of papers published per year before the year 2011 range about to 3 (average of three publications per year) and in the latter half of the period under review i.e., from 2011 to 2019 (excluding few studies taken from the year 2020) range about to 6 (average six publications per year) as shown in Fig. 10. This estimated average signifies the role of CROHME in doubling the useful publication per year. CROHME, a well-established annual competition conducted to accelerate the growth of research in this challenging domain. After the start of the competition in this domain, there has been an escalating interest in the handwritten mathematical symbols and expressions recognition average of double the studies per year are produced after 2011. It is credibly not amazing since the concept of handwritten mathematical symbol and expression recognition is grasping more researchers' attention because of the advancement of the



**Fig. 10** Distribution of selected studies according to the year of publication

research work in the fields of ML, deep learning, and computer vision. Moreover, the need to input the MEs directly in the system is not indispensable; thus, research for high accuracy recognition processes and systems is expected to be exponentially rising, opening new ways and methods to recognize and retrieve mathematical character expressions.

#### 4.4 Geographical distribution of research studies over the years

The count of papers is also analyzed from the perspective of their geographical distribution and year of publications. It has been observed from the figure that the maximum research articles are from China and Japan, where the former had 17 releases and later had 16 papers. The contribution of different countries in different years can be visualized by Fig. 11.

#### 4.5 Distribution of publications by authors

Figure 12 below shows the distribution of authors’ total research publication counts. The authors highlighted in this figure have made significant contributions to the field of research dedicated to identifying HMEs over the course of several years by publishing high-quality articles in the field. Researchers that place highly on this graph of study distributions likely have a keen interest in and considerable competence with the research topic of recognizing mathematical expressions. In addition to the metadata that was gathered, the analysis presented here is likely to be of great assistance to the domain experts who have dedicated themselves to this line of inquiry and developed the relevant recognition algorithms.

On analyzing the Fig. 12, it can be noticed that author Anh Duc Le has contributed most to the dictionary of selected studies, succeeded by authors H. Mouchere, M. Nakagawa,

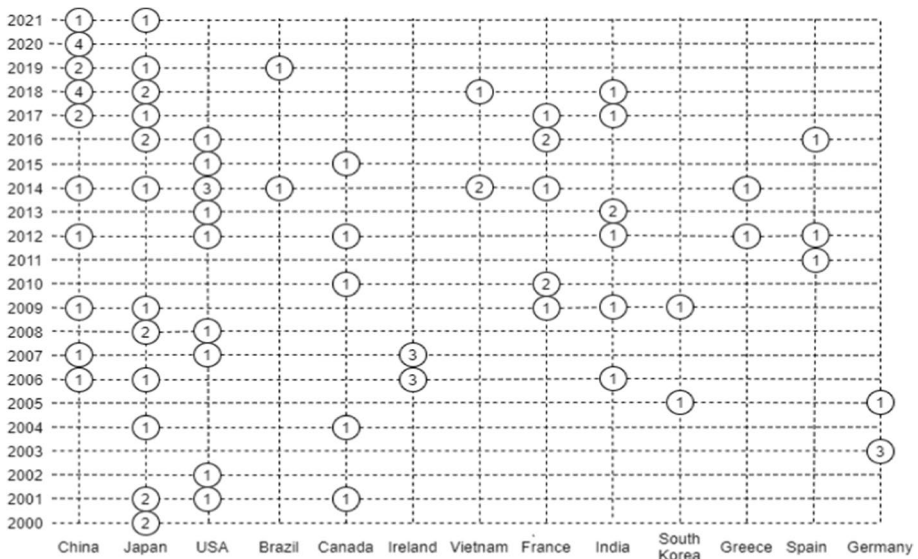
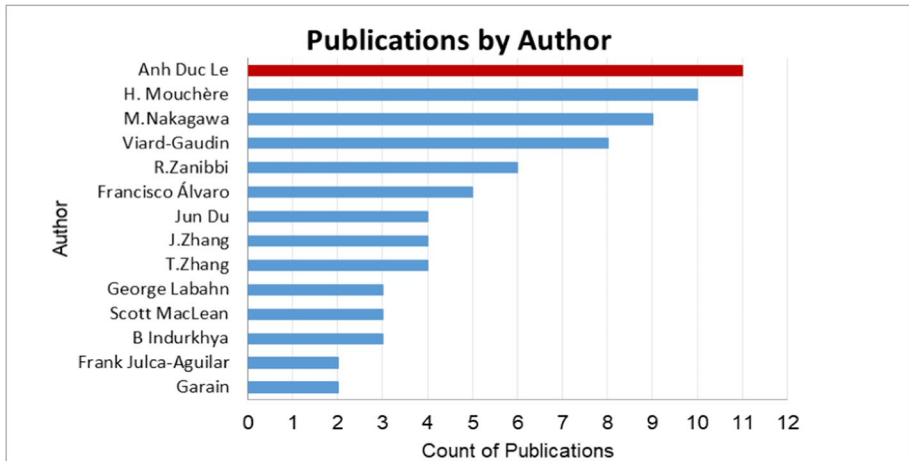


Fig. 11 Geographical distribution of research studies per year



**Fig. 12** Distribution of publication according to the authors of the publications

Viard-Gaudin, and Richard Zannibi. These are the authors who are leading the research in recognizing handwritten mathematical symbols and expressions.

#### 4.6 Frequency of keywords

The authors also tried to analyze the keywords count in the publications retrieved because these keywords are responsible for such high relevance ratio of the studies according to the search string. It has been observed that ‘*mathematical expression recognition*’ is the most common keyword found in the selected studies. This keyword has more relevance with the topic of the review objective.

Other specific keywords like ‘*mathematical expression*’, ‘*online handwritten mathematical expression*’, ‘*handwritten mathematical expression*’, and ‘*online recognition*’ are among the extreme top entries while tabulating the count of keywords found in the retrieved studies after the entire selection procedure. These frequently used keywords associated with the topic of this review study are equipped in contextualizing the keywords of our own review study as our study will revolve around these selected papers having above mentioned keywords in their abstract and keywords section. The graphical analysis of the keywords associated with the count of publications is presented in Fig. 13.

### 5 Results and discussion

The discussions around the extracted answers to the research questions are elaborated in this section.

*RQ1. Which ML/non-ML techniques are used in the studies for recognition?*

Before analyzing the type of ML and non-ML learning techniques used to recognize HMEs, the authors segregated the selected studies based on the approach used. It is found that from the total selected studies, 60 studies used ML algorithms for the task of



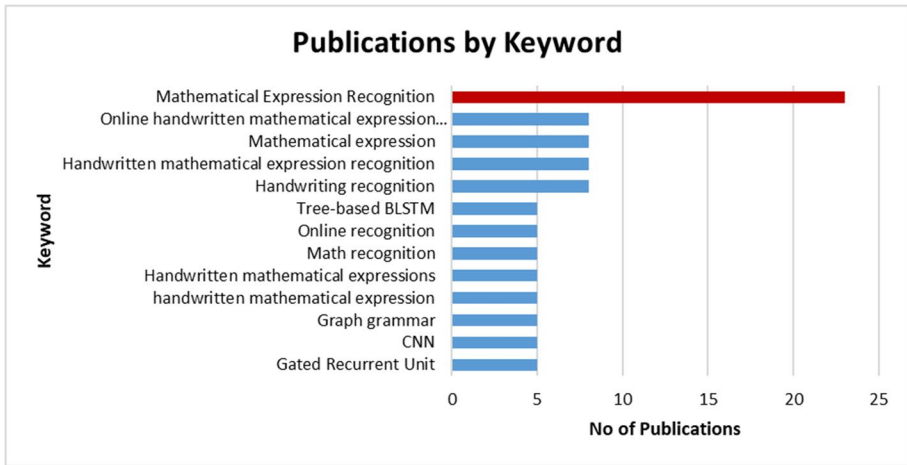


Fig. 13 Distribution of publications according to the frequency of keywords used

handwritten math recognition. In contrast, the count of studies implementing the non-ML approach is 38. As the popularity of ML and deep learning algorithms grew a bit late by the end of 2010, significant studies are using conventional and non-ML-based methods before that. An endeavor has been done to analyze the trend and popularity of the ML techniques used for recognition over the years. Figure 14 depicts the journey of the rising trends of machine learning techniques from the year 2000 to the present time. On the onset, it is found that not even a single study used ML methods of recognition in the year 2000, and a downward trend is observed until 2011. After that, the pace and implementation of ML algorithms ([45, 116, 143] increased and ending in almost every study using ML models by the year 2021. The stack frames colored in red depict the studies using ML techniques to recognize, and the stack colored in grey represents the studies implementing non-ML approaches. This graphical representation specifies the rising trend of ML techniques over the years. On analyzing, 60 studies are found using ML models for recognition tasks, and

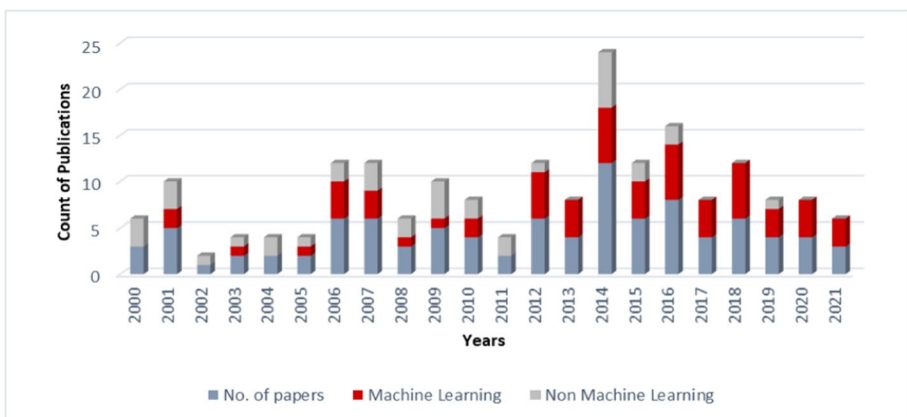


Fig. 14 Yearwise distribution based on ML/non-ML technique

the authors further analyzed several different ML techniques used in all the studies selected by the SLR. The list of the ML techniques used for the recognition of HMEs is as follows:

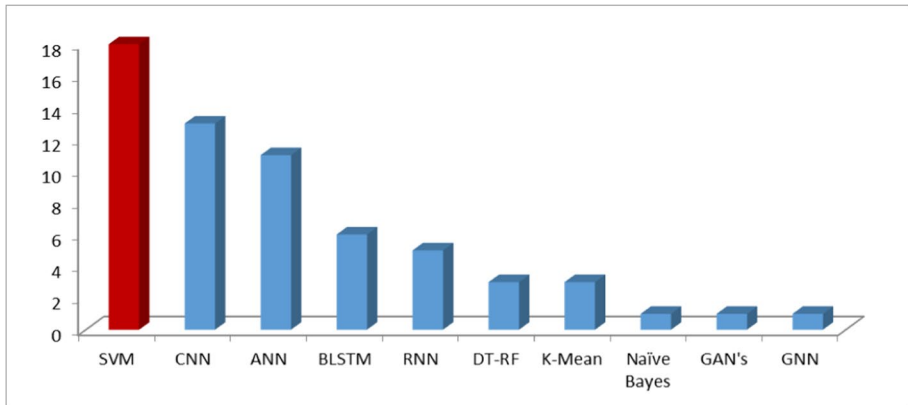
- Support Vector Machine (SVM)
- Artificial Neural Network (ANN)
- Convolutional Neural Network (CNN)
- Recurrent Neural Network (RNN)
- Bidirectional Long Short Term Memory (BLSTM)
- K-means neighbors (K-means)
- Decision Tree and Random Forest (DT + RF)
- Generative Adversarial Networks (GANs)
- Graph Neural Network (GNN)

Figure 14 shows for each technique the number of studies that applied it. It can be noted that the overall most frequently used ML technique is SVM. Still, suppose the trend is observed after the year 2013. In that case, it is found that ANN and CNN are more frequently applied to the recognition systems when the experimentation part of several papers is scrutinized intensely. As per our observations, almost eleven papers used ANN and thirteen studies that used CNN to classify and recognize math expressions. On overall analysis, it is observed that SVM and CNN are most frequently used as the count of the studies using these techniques ranged about 18 and 13 studies, respectively. Adding to this, SVM gave neck to neck competition to the application and implementation of ANN. Even after the year 2013, there is considered a decent number of papers using SVM for recognition. So, generalizing, it can be concluded that both SVM, ANN, and CNN are used primarily to recognize HMEs and symbols.

When the number of papers associated with each method is broken down, it is discovered that the ML method known as SVM is the one that is used the most frequently. This method has been utilized in approximately 18 research, which accounts for approximately 29% of all studies. CNN is the second most common estimating method, and it has been explored in approximately 13 (approximately 21%) distinct papers. Several distinct varieties of neural networks (NNs), such as backpropagation networks [167] and recurrent neural networks, have been utilized and incorporated in this research [11, 216, 219], and fuzzy NN [69, 90, 118], etc. ANN has been investigated in 11 (18%approx.). While BLSTM has been employed in 6 (11%approx.) and RNN has been implemented in four different studies. Decision trees and Random forests have been used in two and one studies, respectively. K-means neighbor has been used in three selected studies, and other techniques like Naïve Bayes and Generative Adversarial Networks have been scarcely used. These techniques witnessed a unit entry under each head. The detailed history of ML models is shown in Table 14, whereas Fig. 15 shows the algorithms used by the selected studies. Figure 16 shows the distribution of HMER based selected studies according to the ML techniques. It is also investigated and observed that the three dominating ML techniques are implemented after several years from when they came into the research innovation and implementation. For instance, the SVM algorithm was invented in 1963, published work reported by 1995, and the research experimentations investigated in the selected studies started implementing the technique in 2003 [179]. The second most frequently used ML technique is CNN. ‘Neocognitron,’ the origin of the CNN architecture, was introduced by Kunihiko Fukushima in 1980. Owing to the pioneering work done by Yann Le Cun, the field of Deep Learning was moved forward by the creation of one of the first convolutional neural networks in the year 1994. This

**Table 14** History and year of implementation of dominant ML models in the studies

Summary of implementation details of ML techniques					
ML technique	Type	Inventors	Year of invention	Year of implementation	Year of implementation in the selected studies
SVM	Supervised learning	Vladimir N. Vapnik and Alexey Ya. Chervonenkis	1963	1995	2003
ANN	Supervised learning	Frank Rosenblatt	1958	Late 1980's	2006
CNN	Deep learning	Kunihiko Fukushima	1980	1994	2015



**Fig. 15** Machine learning methods used by studies

network, which was given the name LeNet5 and was implemented in 1988 after many successful iterations that came before it, is credited with propelling the field forward. After conducting research, it was discovered that the very first study employing CNN took place in the year 2015. ANN, which was first developed in 1958 by psychologist Frank Rosenblatt, is the third most used machine learning technique. It wasn't until the late 1980s that many real-world institutes began using ANNs for a variety of applications, and the research that were included in SLR didn't start making use of this ML approach until the year 2006 [153].

Delving in the non-ML recognition techniques, it is found that there has been a blend of trends witnessed in the studies before the advent of the machine learning era. The major non-ML recognition approaches identified are grammar-based approaches like graph grammar ([56, 75, 89], stochastic context-free grammar ([7, 108, 134, 201], probabilistic context-free grammar ([35, 171], definite clause grammar [39], and another algorithmic approach ([141, 142, 210]. There have also been instances of the studies which are concentrated towards parsing ([41, 108, 206, 208], fuzzy methodologies ([58, 60, 70, 90, 102, 117, 118] and other methods based on relational grammars [120]. Though the total count of studies, 38% of the chosen studies, are purely based on non-ML approaches, the inclination towards these methods cannot be neglected as the initials of research on HMSER thoroughly engaged in their deployment on these approaches. But, undoubtedly, overall, ML techniques dominated the recognition trends in HMSER.

### Pros and Cons of using ML/ Non-ML approaches for HMSER Non-ML Techniques:

#### Pros:

- Non-ML methods, such as rule-based or grammar-based approaches or template-matching algorithms or other parsing techniques, can be more efficient and require less computational capacity than ML methods.
- These techniques can be useful for recognizing simple mathematical expressions or symbols with well-defined patterns and structures.

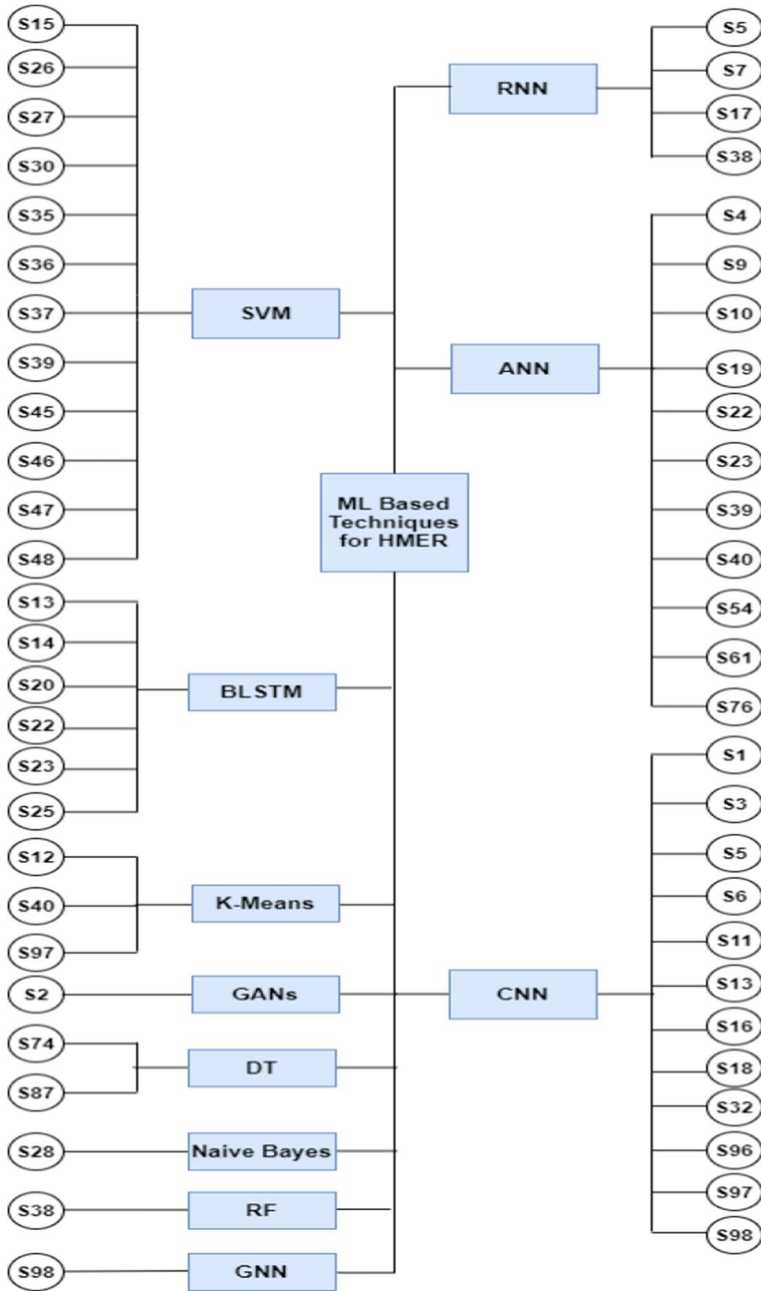


Fig. 16 Distribution of studies according to the ML technique implemented

Cons:

- Non-ML techniques are less precise when it comes to recognizing complex mathematical expressions or symbols, where handwriting, size, and orientation variations may exist.
- Non-ML techniques necessitate more manual intervention and specialized knowledge because they rely on predefined rules or templates that must be created and maintained.

ML Techniques:

Pros:

- ML techniques, such as deep learning algorithms, are at places extremely accurate at recognizing complex mathematical expressions because they can learn to recognize patterns and characteristics from large datasets of HMEs.
- The ML techniques can accommodate variations in handwriting, size, and orientation because they can learn to recognize the mathematical expression's underlying structure rather than its individual symbols.
- ML techniques are also capable of automatically adapting to new handwriting patterns and symbols.

Cons:

- To train and optimize ML models, large amounts of data and computational capacity are required, which can be time-consuming and costly.
- ML techniques can also be susceptible to overfitting, which occurs when models learn to recognize specific patterns in the training data but cannot generalize to new data.
- ML techniques may necessitate more specialized knowledge for model development and optimization.

To summarize, non-ML techniques may be effective for basic mathematical expressions or symbols, but they are less precise for complex expressions. ML techniques can accomplish greater accuracy and can handle complex expressions, but their development and optimization require more data, computational power, and expertise.

*RQ2. Which datasets are frequently used in the studies?*

Almost ten datasets have been used in the selected studies. The datasets that have been used at least once in the study have been considered. The listing of all datasets used is mentioned in Table 15, along with the count of studies that used these datasets.

**Table 15** Count of papers to different datasets

Name of the dataset	Count of papers	Name of the dataset	Count of papers
CROHME	39	Kaggle	1
Self-created, handwritten	33	MNIST	1
Aster	4	IM2LATEX-100 K dataset	1
INFTY	4	MfrDB	1
Mathbrush	3	Ekush	1
Hands-Math	3	Kaggle	1

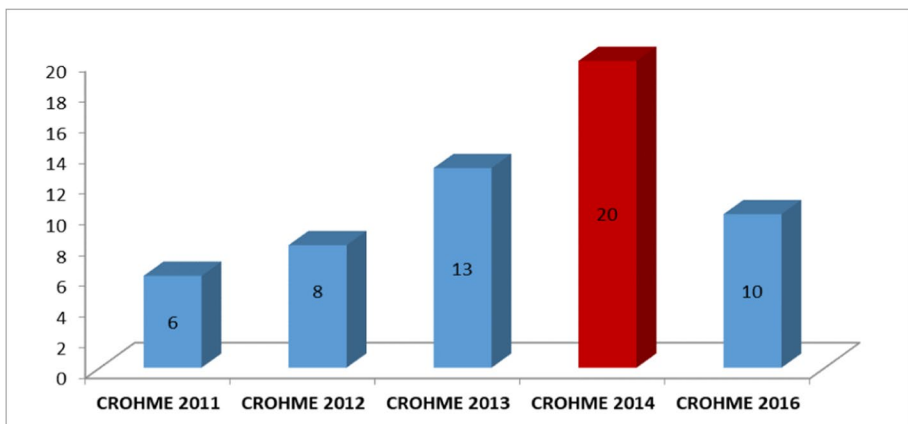
It has been observed that the most widely used dataset in the SLR, which has been employed in studies almost 38% of is CROHME. This dataset is provided by the CROHME series of competitions. CROHME is a competition that was first held in 2011 in Beijing as a part of the International Conference on Document Analysis and Recognition (ICDAR). In addition, CROHME is one of the organizations that has contributed to the mathematical formula symbol library. The CROHME encourages research in the area of HME recognition, which is one of its focus areas.

In addition to supplying the dataset, it also provides academics with a platform on which they can test their methods, analyze them, and then ultimately stimulate further development in this area. Prior to the development of CROHME, only a modest amount of research on math recognition had been carried out without the benefit of benchmark datasets, standard encodings, or evaluation tools. Researchers are able to effectively evaluate different systems and work on improving handwritten math recognition thanks to the CROHME competition. The self-created dataset is referred to as the second name in the table shown above. This type of dataset is extremely adaptable due to the fact that several authors have generated a unique and diversified dataset of their own. For example, some authors have included some written expressions from conventional math books, while others have included volunteer writers to compose HMEs to make their dataset. It has been noted that the majority of the studies conducted before to the year 2011 concentrated on taking into consideration a dataset that was produced by the researchers themselves because a common standard dataset was not readily available. But after the opening of the CROHME series, most of the studies majorly used the CROHME dataset only.

As the majority of the studies used the CROHME dataset, the authors decided to analyze the sub-datasets, which are part of the dataset of this series of competitions. CROHME is held in the years 2011, 2012, 2013, 2014, and 2016. In each year, it provided different datasets of handwritten mathematical symbols and expressions. The authors have investigated the count of studies using these other datasets launched in additional years. Figure 17 clarifies the trend of the different datasets of CROHME used by the selected studies.

### *RQ3. What type of handwritten symbols are used (online/offline)?*

Recognition of handwritten symbols written on paper or non-digital platforms is called offline recognition. When written on the digital platform where recordings of pen/fingertip



**Fig. 17** Distribution of studies according to CROHME sub-dataset

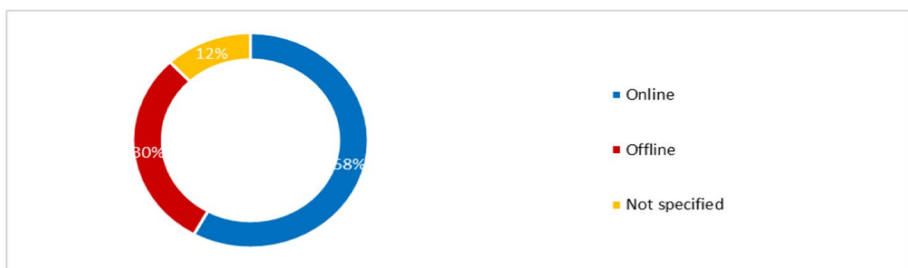
movements are there, called online recognition. With regard to the additional information that is offered online about how the writing is done, the accuracy of online recognition is typically higher than that of offline recognition. Despite this, offline handwriting recognition is employed in large-scale real-world systems, such as determining the monetary values on bank cheques or deciphering handwritten postal addresses [145]. In online expression recognition, the input to the system is often composed of a set of strokes that contain geometric and temporal information. In the event that online recognition is performed, the system is able to make use of the temporal information that is contained within the online input image. When working in offline mode, the image that is being input does not have such information in its geometric or temporal components. This mode of input is consequently made more difficult to access and is utilized significantly less frequently during the recognition process. Table 3 illustrates the primary distinction that exists between the two categories HMEs. Figure 18 illustrates the percentage breakdown.

It can be observed that most of the papers from the selected database of studies used and worked upon online handwritten math glyphs rather than offline. The reason could be the fact, as mentioned above, that the online data contains comparatively much more spatial and temporal information. Thus, most of the researchers prefer working on online HME.

*RQ4. What metric is used to measure accuracy, how much accuracy has been achieved, and which techniques?*

Accuracy measures are essential components of SLR. They are used to highlight how reliable a particular proposed model or system is when it comes to predicting and recognizing HMEs. Several accuracy measures have been applied in the excellent collection of the selected 95 studies. All the frequently used accuracy measures used in different studies are defined in Table 16.

On investigating and observing the trend of accuracy measures used, it is noticed that the studies after the year 2010 are more likely to give comparable results as the well-formed datasets came into implementation after this year only. Before the year 2010, almost 90% of the studies used self-created datasets. Among these self-developed datasets, it is hard to list out and compare the accuracy metric used. It is the reason for neglecting some portion of comprehensive studies that were published before the year 2010. Table 17 shows the list of several studies taken into consideration while comparing the accuracy metrics used by the studies for experimentation and implementation of the recognition techniques for HMEs. The review writers have compared the accuracy measures used by all the studies and observed the dataset on which the experimentation is carried out. Among these



**Fig. 18** Distribution of papers based on HME type



**Table 16** Description of Metrics/ Accuracy measures used by the reviewed studies

Metric	Representation	Observation	References
Accuracy	$\frac{TP+TN}{TP+FP+FN+TN}$	Measures the percentage of correct expressions classified	S1,S5,S37,S56,S96
Precision	$\frac{TP}{TP+FP}$	Percentage of correct predictions in recognition	S8,S14,S25,S20,S21,S29,S35,S46,S50
Recall	$\frac{TP}{TP+FN}$	Represents the proportion of positive samples that are correctly predicted	S39,S43,S55
Error rate	$\frac{\text{Total no. of wrong predictions}}{\text{Total no. of predictions}} * 100$	The ratio of no. of predictions gone wrong to total no. of predictions made	S28, S49
Stroke recognition rate	$\frac{\text{No. of correctly recognized strokes}}{\text{Total no. of strokes}} * 100$	Percentage of no. of corrected recognized strokes to total no. of strokes	S7,S15,S19,S20,S21,S23,S25,S28,S29,S33,S34,S35,S36,S40,S49,S50,S52,S53,S58
Segmentation rate	$\frac{\text{No. of correctly segmented symbols}}{\text{Total no. of symbols}} * 100$	Percentage of no. of corrected segmented symbols to total no. of symbols	S24,S28,S33,S34,S49,S50,S51,S52,S53,S58,S97
Symbol recognition rate	$\frac{\text{No. of symbols corrected recognized}}{\text{Total no. of symbols}} * 100$	Percentage of no. of corrected recognized symbols to total no. of symbols	S2,S4,S6,S9,S10,S13,S15,S16,S17,S18,S20,S23,S24,S25,S27,S28,S29,S33,S34,S35,S36,S40,S49,S50,S51,S52,S53,S58,S98
Expression recognition rate	$\frac{\text{No. of expressions corrected recognized}}{\text{Total no. of expressions}} * 100$	Percentage of no. of corrected recognized expressions to total no. of expressions	S15,S24,S33,S34,S47,S50,S51,S52
Structural recognition	$\frac{\text{No. of structures corrected recognized}}{\text{Total no. of structures in an expression}}$	Percentage of no. of corrected recognized structures to total no. of structures	S3,S12,S60
Purity	$\frac{1}{N} \sum_k \max_j  w_k \cap c_j $	The purity of a cluster represents the fraction corresponding to the largest class of documents assigned to that cluster.	

**Table 17** Accuracy analysis of the reviewed studies

Study name	Technique used	Dataset	Accuracy measure	Accuracy value
S1	CNN	CROHME (2013 2014 2016)	Accuracy	92.96%
S2	GAN	CROHME 2014	Expression recognition rate	48.88%
S3	CNN	CROHME 2014, CROHME 2016	Purity	0.99%
S4	ANN	CROHME 2014, CROHME 2016	Expression recognition rate	46.33%, 49.46%
S5	RNN	IM2LATEX-100 K dataset	Accuracy and BLEU	97.5%, 88.42%
S6	CNN	CROHME 2014, CROHME 2016	Expression recognition rate	48.78%, 45.60%
S7	ANN	MINIST, Kaggle	Symbol segmentation rate	96.90%
S8	Graph grammar	CROHME 2016	Recall and precision	segmentation(95.32% 96.84%) clas- sification (89.39% 90.81%)relations (78.68% 79.93%)
S9	ANN	CROHME 2014	Expression recognition rate	61.16%
S10	ANN	CROHME 2014,2016	Expression recognition rate	52.8%(2014),50.1%(2016)
S11	CNN	Self-created	Not specified	Not specified
S12	K MEANS	Self-created	Purity	0.92%
S13	CNN	CROHME 2014	Expression recognition rate	37.635
S14	BLSTM	CROHME 2016	Recall and Precision	segmentation(95.64% 91.44%) seg+class(89.84% 85.90%) rel tree (77.23% 74.08%)
S15	SVM	Handsmath, CROHME (2013,2014,2016)	Symbol segmentation rate, Segmentation+classification, Structural recognition, Expression recognition rate	92.46%, 88.24%, 89.79%, 68.07%
S16	CNN	CROHME 2014	Expression recognition rate	35.19%
S17	RNN	CROHME 2014	Expression recognition rate	52.43%
S18	CNN	CROHME 2014 CROHME 2016	Expression recognition rate	46.55%, 44.55%
S19	ANN	Self created	Recognition rate/ percentage	97% approx.

**Table 17** (continued)

Study name	Technique used	Dataset	Accuracy measure	Accuracy value
S20	BLSTM	CROHME 2014	Segmentation rate, Segmentation + Classification, Tree Relationship (in terms of Recall and precision), Expression recognition rate	(95.59%, 91.46%) (89.02%, 85.17%) (74.29%, 81.33%) 21.59%
S21	2D-CFG	CROHME 2013	Segmentation rate, Segmentation+Classification DAG relations(in terms of Recall and Precision)	(92.0%, 90.7%) (82.2%, 81.0%) (88.0%, 82.6%)
S22	BLSTM	CROHME (2013,2014)	n-best rate	87.35%(2013), 91.28%(2014)
S23	SVM,CYK+SCFG	CROHME (2013,2014), HAND-MATHS	Symbol segmentation rate, Segmentation +Classification, Relational Tree, Expression recognition rate, Updated Expression Recognition rate	2013[(88.49%),(79.63%)(71.17%) (32.34%)] 2014[(84.67%), (77.21%) (68.62%) (32.86%) ]handmath[(92.46%) (88.24%) (89.79%) (68.07%)]
S24	CYK + 2DCFG	CROHME (2012,2014)	Symbols Relationships Expressions	81.95%, 79.89%, 26.88%
S25	BLSTM	CROHME 2014	Segmentation, Segmentation + Classification, Tree Relationship. (in terms of Recall and Precision), Expression recognition rate	(92.77%, 85.99%), (85.17%, 78.95%) (67.79%, 67.33%) 13.02%
S26	SVM	Self created	Recognition rate	ALMOST 98%
S27	SVM	MATHBRUSH, CROHME (2012, 2013)	Expression recognition rate	50.22%
S28	NAIVE BAYES	MATHBRUSH, CROHME (2011, 2012)	Stroke rec, Symbol segmentation rate. Symbol recognition ratio. Expression recognition rate.	[2011] [2012] 93.16%, 91.54% 97.64%, 96.56% 95.85%, 94.86% 66.36%, 67.40%
S29	Graph grammar	CROHME 2013	Recognition rate, Segmentation rate, classification tree rel(in terms of Recall and Precision)	21.61, (75.70%, 83.41%) (62.63%, 69.00%), (44.67%, 49.73%)

Table 17 (continued)

Study name	Technique used	Dataset	Accuracy measure	Accuracy value
S30	SVM	Self created	Recognition rate	ALMOST 99%
S31	CYK algo	Self created	Recognition rate	62.74%
S32	CNN	CROHME	Not specified	NIL
S33	HMM, 2DCFG	CROHME 2011	Stroke rec, Symbol segmentation rate, Symbol recognition rate, Expression recog- nition rate	78.73%, 88.07%, 92.22%, 29.28%
S34	Gaussian structure, global classifier	CROHME 2011, ASTER	Segmentation rate, Recognition rate, ExpRate	89.79%, 95.21%, 63.54%
S35	SVM	CROHME 2013	Symbol segmentation rate, Seg+Class, Rel Tree, Exp Recogni- tion rate	81.95%, 67.06%, 49.12%, 20.72%
S36	SVM	HANDSMATH	Symbol segmentation rate, Segmentation+recognition, Rela- tions tree, Expression recognition rate	Not specified
S37	SVM	CROHME (2012, 2013), Math brush	Accuracy	93.52% CROHME 2013 96.36% CROHME 2012
S38	RNN	CROHME 2013	Recognition rate	PRHLT, FKI, and RWTH offline fea- tures improved up to 83.9%, 84.1%, and 83.4%
S39	SVM	CROHME 2012, Mathbrush	Error rate	2.87% (MathBrush)
S40	x-y cut algo	CROHME 2014	Symbol segmentation rate Symbol segmentation rate + Recognition Relation Tree Expression Recogni- tion	91.90%, 84.86%, 79.75%, 37.63%

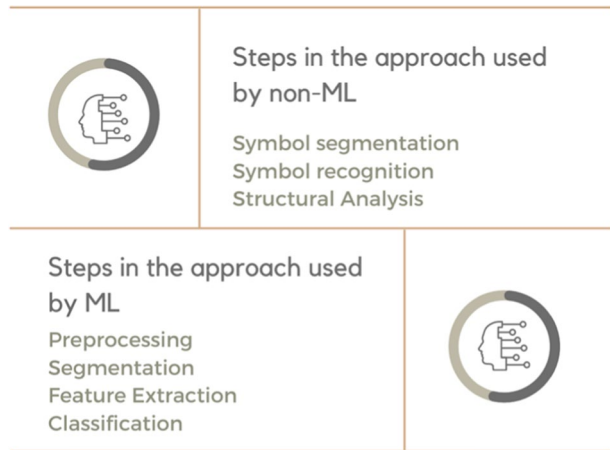
**Table 17** (continued)

Study name	Technique used	Dataset	Accuracy measure	Accuracy value
S41	Gaussian mixture model dynamic programming	Self-created	Mes Fully Recognized(%) Structure Analyzed (%) Symbols Segmented(%)	17.0%, 47.3%, 67.0%
S42	DRACULAE parser	Infty	Recognition rate	74%
S43	Statistical method	Self-created	Error rate	3.90%
S44	RF	CROHME 2014	Symbol classification rate	Not specified
S45	SVM	Not specified	Not specified	Not specified
S46	HMM, SVM	CROHME (2011, 2012)	Recall and Precision	Not specified
S47	SVM	Self-created	Symbol Recognition rate	94.833%
S48	SVM + Projection Histogram	Self-created	Recognition rate	98.26%
S49	FUZZY	CROHME 2011	Stroke rec., Symbol segmentation rate, Symbol recognition rate, Expression recognition rate.	71.73%, 84.09%, 85.99%, 32.04%
S50	SVM	Self-created	Recall, Precision, and F1	86.94%, 84.29%, 85.60%
S51	HMM + LL(1) PARSER	CROHME 2011	Stroke rec, Symbol segmentation rate, Symbol recognition rate, Expression recognition rate	41.69%, 47.35%, 87.00% 13.26%
S52	ELASTIC MATCHING DISTANCE	CROHME 2012	Stroke rec, Symbol segmentation rate, Symbol recognition rate, Expression recognition rate	75.90%, 86.67%, 81.76%, 9.94%
S53	SVM	MATHBRUSH	Symbol segmentation rate, Symbol recognition rate, Expression recognition rate	4.86%, 79.82%, 25.31%
S54	ANN	Self created	Recognition rate	90%
S55	2DSCFG	(InftyCDB-1).	Error rate	4.68%.
S56	GRAPH GRAMMAR 2D SCFG	Self-created	Correctly Recognized Expressions, Correct Structural Analysis	52%, 88%

**Table 17** (continued)

Study name	Technique used	Dataset	Accuracy measure	Accuracy value
S57	Parsing	Self-created	Not specified	Not specified
S58	GAUSSIAN MODEL	ASTER	Segmentation rate, Symbol recognition rate, Expression recognition rate	91.4%, 88.7%, 64.9%
S59	SVM	ASTER	Segmentation rate, Recognition rate, Expression recognition rate	73.6%, 72.9%, 50.6%
S96	CNN	Ekush	Accuracy	96.01%
S97	CNN and K-means Clustering	CROHME (2013, 2014, 2016), Self- created	Symbol recognition rate	89.39% (CROHME 2014)
S98	CNN, Graph Neural Network	CROHME (2014,2016)	Expression recognition rate	54.46% (CROHME 2014) 52.05% (CROHME 2016)

**Fig. 19** Generalized approach used by different ML and non-ML methods in the selected studies



measures, ExpRate, which means expression recognition rate, has been widely used in almost 43% of these studies.

*RQ5. What approach is followed in the study, or what is the proposed system?*

In the research question, the authors have tried to understand the kind of approach followed by the selected studies for recognizing the math expression in handwritten form. On investigating and generalizing the approaches, it is found that the recognition systems following non-ML approach, or we can say that the approach other than the ML, has used chiefly followed the conventional approach for solving the research problem in the steps of symbol segmentation, symbol recognition, structural analysis, and expression recognition. The last phase (extraction recognition) is common to the studies that used ML techniques. And for ML techniques, it is observed that the approaches proposed followed some steps (that are generalized by the authors) like preprocessing, segmentation, feature extraction, classification, and expression recognition. The generalized approach is shown in Fig. 19.

*RQ6. What kind of techniques is/are used in the sub-process?*

It has also been keenly observed that procedures or special techniques are followed to execute these phases mentioned above in both approaches efficiently. The focus of this research question is centralized to highlight the most frequently used specialized procedures and techniques for conducting and executing this problem stage in both approaches. The authors noticed that some studies used several different methods for preprocessing, segmentation, classification, etc. The summary of the experimental techniques used in the sub-process of the recognition is presented in Table 18. Note that those papers have only been considered here where some special techniques are used in the sub-process.

*RQ7. Which ML Technique outperforms other ML techniques?*

About 60% of the selected studies used ML techniques, which are compared with other ML methods. For reaching out to actual conclusions for this research question, the performances of different ML techniques have been compared based on the accuracy measure/

**Table 18** Sub-processes analysis concerning the reviewed studies

Preprocessing	Symbol segmentation	Symbol recognition	Structural Analysis	Feature Extraction	Classification	References
Filtering, Morphological Operations, Thinning (Skeletonization)	Not specified	Not specified	Not specified	. Zoning, Skeleton based direction, Projection histogram, Profiles, Structural features	Hidden Markov Models,SVM classifier	[71]
Not specified	Projection method, Cutting method	Recursive cortical network	Not specified	Not specified	Not specified	[52]
Noise Reduction, Binarization, Thinning	Recursive Projection projection cutting algorithm,	Not specified	Not specified	Not specified	CNN classifier	[54]
Not specified	An SVM classifier is used for segmentation, which takes 21 geometric features	Markov Random Field,(MRF) Modified Quadratic Discriminant Function(MQDF)	2D-CFG with CYK algorithm	Not specified	Not specified	[142]
Image binarization and normalization operations have been carried out in preprocessing.	edge detection, bounding box morphological operations	Not specified	Not specified	skew, standard deviation, mean, variance	Backpropagation neural network used as a classifier.	[167]
Not specified	GMM-model	RNN-BLSTM based classifier(symbol classification)	2D-PCFG based grammar	Not specified	Not specified	[12]
Not specified	SVM classifier	MRF	SCFG	Not specified	Not specified	[108]
Median filter, Morphological operations, image thinning	Not specified	Not specified	Not specified	Normalized Chain Code, Moment Invariant Features, Density Feature: UDRL density, Projection Histogram	SVM classifier	[27]
Not specified	Not specified	HMM, SVM(spatial relation classification)	2D-SCFG, CYK Parsing	Not specified	Not specified	[10]



**Table 18** (continued)

Preprocessing	Symbol segmentation	Symbol recognition	Structural Analysis	Feature Extraction	Classification	References
Not specified	Bounding box, the SVM classifier	Not specified	2D-SCFG, CYK Parsing	Not specified	Not specified	[110]
Not specified	SVM	Not specified	Body Box, 2D SCFG	Not specified	Not specified	[107]
Not specified	HMM	Not specified	Not specified	Not specified	SVM classifier	[127]
normalization,noise,compression	labeling and bounding boxes	Not specified	Not specified	Projection histogram	SVM classifier	[72]
Not specified	Not specified	Nearest neighbor, bounding box	2D-SCFG, CYK algo	Not specified	Not specified	[7]
Not specified	Not specified	2D dynamic programming (TDNN) symbol classifier	Bounding Box	Not specified	Not specified	[22]
Not specified	Graph model	GMM	Not specified	Not specified	Not specified	[83]

metric used. By comparing the accuracy results, the appropriate answer to this question can be retrieved well and analyzed. Still, it would be unfair to compare the distinct values belonging to different accuracy metrics. So, the authors decided to poll out the most frequently used measure, which could be taken as a standard metric for accuracy analysis. And on comparing the corresponding values of accuracy and performance, the authors could extract well which technique is outperforming the rest. Briefing the steps involved to reach conclusions for this RQ7 are as follows:

1. Construct a database/table detailing the used technique in the study, dataset, accuracy measure/metric, and an accuracy value.
2. Observe and identify the accuracy metric, which has been used most of the time.
3. Compare the accuracy values corresponding to that identified accuracy metric (observed in step2) w.r.t different ML techniques used.

Note: This procedure for comparison is designed as different studies used different datasets and various accuracy measures. So, it won't be purely justified if we observe the accuracy values irrespective of the accuracy metric used. The authors believed that these conclusions could be more acceptable and more exceptional if the performance could be evaluated and the same datasets and same accuracy measure. But the lack of standardization forced us to assess the results by the above-defined procedure

Table 17 allows for an easy analysis that reveals ExpRate to be the accuracy metric that is utilized the majority of the time. It is possible to quantify the accuracy of the expression recognition rate as the proportion of correctly recognized expressions to the total number of expressions. On analyzing the value against the ExpRate (accuracy metric commonly used), the highest accuracy value observed is, i.e., 68.07%. This ExpRate is found by applying SVM on the Handsmath dataset (using an augmented incremental approach) (refer to S15). Thus, it is observed that the highest accuracy rate is found as the outcome of the SVM and directing us to the conclusion that SVM outperformed all other ML models used by different studies.

NOTE: The research articles comprising summary of winning models of CROHME competition have been excluded for comparison purposes, and independent studies have been taken for performance analysis and comparisons

#### *RQ8. Which ML techniques outperform other non-ML methods?*

As for non-ML approaches, about 40% of the selected studies used non-ML techniques. The non-ML methods usually followed the recognition steps of symbol segmentation, symbol recognition, and structural analysis. There are fewer trends and frequencies observed while noticing the kind of non-ML techniques used. So, when comparing their performances with ML techniques, the studies that focused on evaluating the performance according to the identified accuracy metric are considered. The accuracy values against the expression recognition rate found the leading evaluated ExpRate as 64.9% obtained by implementing the Gaussian model. This accuracy rate is comparatively less than what accuracy rate produced by the SVM model. Thus, it can be concluded that SVM (an ML model) outperformed the other non-ML techniques.

It must be noted that a study abbreviated as S23 gave the ExpRate is 68.07%, the same as that offered by the study S15. The Cocke–Younger–Kasami (CYK) algorithm is used to parse two-dimensional (2D) structures of online handwritten MEs, and MEs are encoded in the form

of a stochastic context-free grammar (SCFG). However, this research utilized both SVM classifiers. This ExpRate is evaluated on the Handsmath dataset. Thus, it can be concluded that the resultant accuracy rate equivalent to that produced the S15(a study that holds the maximum expression recognition rate) is achieved by undertaking a hybrid approach, i.e., by applying ML and non-ML models. This combined approach is out of the scope of this review.

*RQ9. Which are the dominant journals/conference proceedings for papers analyzing HMEs recognition?*

We identified 98 studies in the field of recognition of HMEs. These papers were published during the time period 2000–2021. Among these selected studies, 45 (45%) papers are published in conference proceedings, 34 (35%approx.) articles appeared in journals, 8 (8%) articles are from workshops, and others are taken from symposiums, forums, meeting proceedings, technical reports, and thesis work. The analysis selected from journals is presented in Fig. 20.

The studies selected for the review are taken from 14 different journals presented in Fig. 20. It is observed that the dominating journal found in this research domain of HMEs are primarily from the renowned journal named International Journal on Document Analysis and Recognition (IJ-DAR), which contributed about 22% of the total studies selected from the journals—followed by Pattern Recognition Letters and Pattern Recognition, which ranked equally according to the figure and together contributed 42% of the journal chosen studies.

The authors have also analyzed the selected studies based on the proceedings of the conferences to which they belong. It is investigated and found that the specified database of papers is taken from 22 different conferences. The dominating conference in this category is the International Conference on Frontiers in Handwriting Recognition, preceded by the ICDAR.

Summary and details of the dominating journals, conferences, and workshops in this domain are presented in Table 19. International Journal on Document Analysis and Recognition (IJ-DAR) has been the most dominating journal. The International Association

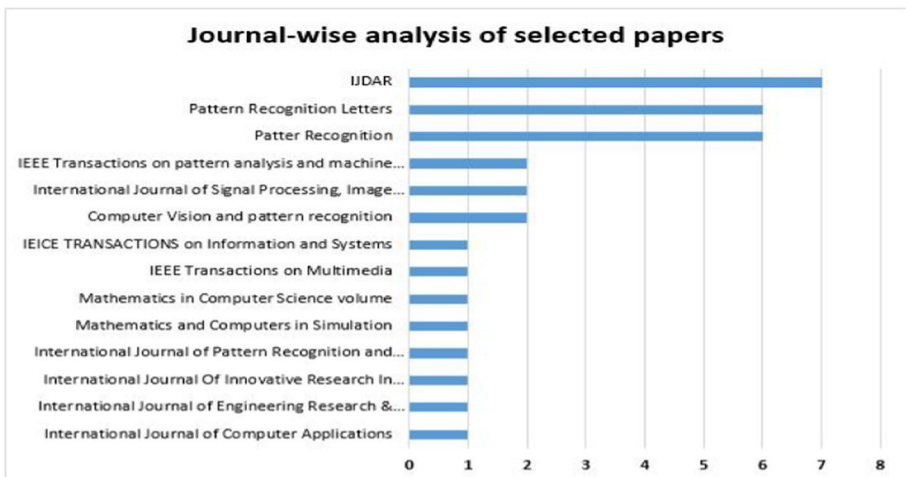


Fig. 20 Distribution of studies by different journals

**Table 19** List of dominating journals, conferences, and others used by reviewed studies

Titles	Category	Publisher	H-index	Country	Subject Area
International Journal on Document Analysis and Recognition(IJDAR)	Journal	Springer	44	Germany	Computer Science Applications Computer Vision and Pattern Recognition Software
Pattern Recognition Letters	Journal	Elsevier	139	Netherlands	Artificial Intelligence Computer Vision and Pattern Recognition Signal Processing Software
Pattern Recognition	Journal	Elsevier	180	United Kingdoms	Artificial Intelligence Computer Vision and Pattern Recognition Signal Processing Software
International Conference on Frontiers in Handwriting Recognition	Conference	–	18	United States	Computer Networks and Communications
International Conference on Document Analysis and Recognition (ICDAR)	Conference	–	60	United States, Japan	Computer Vision and Pattern Recognition
International Workshop on Frontiers in Handwriting Recognition(IWFHR)	Workshop	–	30	United States	Computer Science (miscellaneous)

for Pattern Recognition is the sponsoring agency for this reputed journal. The goal of this journal is to publish articles related to the areas of document analysis as well as document recognition. It invites articles for four different types, 'original research papers', 'system descriptions', 'correspondence', and 'overviews and summaries'. It also focuses on coming out with special issues that target active areas of research.

## 6 Summary of findings

The author has thoroughly reviewed and presented the results of the systematic review analysis performed on recognition techniques for HMEs. We have followed the guidelines designed by [96] and applied inclusion-exclusion criteria on the retrieved studies, fetched by using a formulated search string on digital libraries like Scopus, IEEE Explore, Science Direct, Wiley, Springer, and ACM digital library. The period of studies considered for this review is from the year Jan 2000 to June 2021. A detailed and related report on this research topic was published last in 2012 [204]. The aspects of analysis considered in the study are entirely different than the systematic approach used by this review. It should also be noted that the authors have chosen that potential studies out of candidate studies based on the quality by which each of these studies can fulfill the requirement of answering the formulated research questions. The authors in this review have given the least attention to citation count and strictly followed quality assessment scores for choosing a study in the review.

The authors also believe that such comprehensive and systematic work has not been performed in this research domain of recognition of HME. As in the initial stage of screening papers before performing the quality assessment, the number of candidate studies is as many as 202. These studies are collected, analyzed, and stored in the screening database for reference purposes. It is found that very few review studies have been performed to date on this research theme of recognition of HMEs. We tried to analyze these reviews to brief what aspects and review concepts have been covered. It could help us give us a better direction before we could plan to frame this study. During this analysis, it has also been observed that there has been no evident review study that used systematic literature review guidelines and frameworks to perform the research review on this challenging research domain. Thus, there is a need for a new analysis that covers the gap of 8 or 9 years and also adds a new systematic dimension to the analysis perceptible of this topic.

Indeed, the idea behind conducting this SLR is to summarize the work done in this field and present several perspectives of review to the present and future researchers that can help them generalize and revise the generic prerequisites required to know about this subject area. Moreover, it covers the techniques used and analyzes most of the details of metadata related to the selected studies for review. The quality assessment criteria are performed where the reviewers chose 98 studies out of 202 retrieved studies. These studies are determined based on the relevance and quality score according to quality assessment questions. These 98 articles are published in 14 leading journals and 22 premier conferences, eight established workshops, and other sources like symposiums, technical reports, and other thesis work.

The authors have sincerely analyzed all possible aspects required for reaching the conclusions and finding answers to the research questions. The data from the selected studies are analyzed qualitatively and quantitatively using narrative, vote counting, and grounded theory methods. The review aims to segregate the studies based on the technique used.

(ML/non-ML) to recognize handwritten maths. The entire focus is not concentrated and is limited to analyze the methods and compare their accuracies. Still, the authors also targeted to analyze the dataset used and the kind of handwritten expressions used in the dataset for experimentation. The primary findings of this study are summarized in Table 20.

Apart from the primary results and findings are drawn from the selected studies of the SLR, the writers have extended the limits of the findings by analyzing the other details of the research publications associating with the research zone. The summary of findings from metadata extracted from the studies included in the SLR is tabulated in Table 21.

The above summarizes the significant findings from the research studies referred to and reviewed while conducting this SLR. There is a lot much to add to the details about the results of the research. Still, we tried to summarize the research findings clearly and crisply by tabulating and listing the highlights and key conclusions extracted from the study. Finally, in a nutshell, we want to highlight that the primary purpose of this review study is to analyze the trend of techniques used to recognize HMEs. Though the research questions are formulated, especially targeting the ML-based studies published between the years 2000–2019, the SLR has also analyzed the studies using non-ML or deep learning models. Thus, extending the scope to a broader scale of review. This systematic SLR follows an entirely different approach than the other review studies performed since 2000. And to the best of our knowledge, this is the **first SLR** ever performed on this research domain of recognition of HMEs. This SLR aims and fulfills all the described objectives by determining

- What are the new priorities of researchers, considering the recognition techniques/datasets/type of handwritten expressions/accuracy metrics/approach/sub-techniques they employed in their studies?
- The experimentation results have been compared to analyze the technique that out-comes the rest of the methods employed for recognition.
- The studies are analyzed based on several temporal aspects related to the recognition technique/publication years/authors/country/journals/conferences/keywords/affiliations, etc.

## 7 Limitations of this study

Taking into account the above findings of this review, we have noticed certain limitations of this study. At certain places, it has been found that the scope would be made broader by formulating more research questions that could analyze in more depth the techniques used for recognition purposes on a larger scale of MEs taken into consideration. Researchers could explore an aspect in the extended version of this study by categorically analyzing the features extracted by several ML studies. This study has restricted this scope by examining the techniques based on the ML/non-ML methods used. Still, there could be another dimension of analyzing the studies based on the combined approach used, which uses some ML classifiers and uses some non-ML methods for recognition.

Further, this review study is constrained as it restricts the scope of generic HMEs, with no examining related to the identification and recognition of other types of MEs written in different languages like Arabic, Chinese, Gurmukhi Devanagari, and so on. Moreover, there is a need for a broader perspective of a more comprehensive review, mainly dealing with HMEs written using different scripts and languages. A detailed study of all features of written math expressions could be analyzed, and this review limits its horizons while

**Table 20** List of findings as per the formulated research goals

Sr.no.	Findings	Results
1.	The widely used technique for recognition of HMEs is	Support Vector Machine (SVM), (type: ML, supervised)
2.	The kind of expressions mainly used for experimentation are	Online MEs
3.	The dataset commonly used by the majority of the studies	CROHME dataset
4.	The type of approaches followed by the recognition techniques are	non-ML Approach: Segmentation, Symbol Recognition, Structural Analysis ML Approach: Preprocessing, Segmentation, Feature Extraction and Classification
5.	The technique that outperformed other techniques (ML/non-ML) in terms of accuracy rate is	SVM
6.	The most commonly used accuracy metric is	With an accuracy rate of 68.07%
7.	The dominant journal in the field of recognition of HMEs is	Expression Recognition rate (ExpRate) International Journal on Document Analysis and Recognition(IJDAR)

**Table 21** Meta-analysis listing and findings from the reviewed studies

Sr.no.	Findings	Result
1.	The trending technique used for recognition	ML(SVM)
2.	The number of studies using ML	57
3.	Most of the studies are published in the year	2014
4.	Trending ML technique used in the first half of the considered period	Support Vector Machine
5.	Trending ML technique used in the second half of the considered period	Artificial Neural Network
6.	Trending and popular ML techniques used in the recent studies	Convolutional Neural Networks (a type of deep learning model)
7.	The leading author of the selected studies	Anh Duc Le
8.	Other leading experts and active researchers contributing significantly quality research in this area	Masaki Nakagawa, Viard-Gaudin, H. Mouchere, Richard Zannibbi, Utpal Garain
9.	The country contributing the highest research articles in this domain	China
10.	The country contributing the second-highest research articles in this domain	Japan
11.	The year in which China had maximum publications	2019,(first quarter of 2020)
12.	The most commonly used keyword found in the studies	Mathematical expression recognition
13.	The count of survey and review articles published since 2000	6
14.	The count of citations of the experimentation based study cited the maximum number of times	226 (S87)
15.	The count of citations of the survey-based study cited the maximum number of times	362 (S1)
16.	The revolutionary event that paced and boosted the research in the domain of recognition of handwritten maths	CROHME
17.	The leading conference in this research area	International Conference on Frontiers in Handwriting Recognition
18.	The recommended workshops in this research domain	International Workshop on Document Analysis and Recognition (IWDAR)
19.	Some popular affiliations sources of the selected studies	The University of Nantes, Tokyo university of agriculture and technology, Rochester Institute of technology



discussing features less thoroughly. Hence we sum up these limitations and leave them to be resolved and fulfilled in the future studies.

## 8 Conclusions and future scope

- Many of the datasets used in the studies lack a large expression corpus. There is a need for a standard dataset, a complete corpus built using expressions of all forms, ranging from high school equations to complex scientific expressions.
- Although many of the offered methods and approaches have been successful in achieving high levels of accuracy in their outcomes, there is still a lack of unified procedures in this field to evaluate the effectiveness of those methods.
- Considering there was no freely accessible public dataset of HMEs before to 2011, the researchers were compelled to collect and construct the set of MEs on their own, which has a tendency to be restricted to a subset of expressions or particular domains. The result of this is that the comparative examination of the various methods is made more difficult. It is impossible and impractical to make a side-by-side comparison of the different recognition models and systems' respective performance levels.
- When it comes to the recognition tasks that are a part of this domain, there are not any standard accuracy measurements that have been developed. The absence of standard accuracy measurements is what causes the inconsistent metric trend that is utilized in the research to evaluate the subjects' performances and carry out the necessary tasks for the experiments. As a result of this, it is difficult to do direct comparisons between the accuracy values that were produced using various accuracy metrics because standard accuracy measurements are not readily available. Therefore there is a need to use a standard accuracy metric that should be well defined and experimented with.
- There are several ways of representation of MEs. Many systems make trees represent expressions resulting from structural analysis, while many others use parsing techniques and express the grammar using context-free grammar rules. Other representations are in binary trees and baseline structure trees, where the latter signifies the hierarchical structure baselines in a mathematical expression. Several procedures are proposed and implemented using representations like bounding box, body box, and hidden writing area (HWA) even for stroke recognition. The several representations used in the recognition process and different classifiers and feature extractors broaden the research zone. And there arises a need to perform a separate comparative study analyzing the recognition systems based on the kind of representation method used in the study.
- As the field of recognition of HMEs has been performing extensively well from the past decade. It becomes an issue of vital importance; thus, this review study concludes that there is a need of the hour to standardize different evaluation measures (accuracy metrics), make use of the standard and commonly used datasets so that there could be better analysis when it comes to extracting conclusions about the performance of different recognition systems. The public benchmarking system should be developed to ease and facilitate the comparative analysis of the achievements of varying recognition systems.
- It is observed that different classification techniques perform differently when combined with different models and other ML/non-ML methods and applied on different datasets. The outputs are varying and highly incomparable in these cases. So, the stand-

ardization of datasets, accuracy measures, and a benchmark system need to be established after thoroughly analyzing, revising, and examining the implementation of the concepts of pattern recognition techniques used in the research domain.

In the end, the authors call for the attention of future researchers on the implementation and accuracy of varying recognition techniques and a need for a more systematic review to be performed on deep learning methods, which have proven to produce a significantly enhanced efficiency. Also, as already mentioned, there is a need for a standard model for the comparison of different techniques is a requirement. Another central highlighting point is that different classification techniques are performed under other experimental conditions on different datasets with varying metrics of accuracy involved. Thus, standard accuracy measures should be adopted for better comparisons in the future.

## Appendix 1

**Table 22** Studies categorized under Quality Assessment Labels

Quality Assessment Score	References
No	[37],[159][220],[101],[2],[78],[192],[162],[188],[117],[33],[93],[100],[203],[158],[92],[193],[80],[156],[195],[99],[151],[212],[53],[210, 211],[200],[81],[133],[115],[137],[49],[34],[77],[132],[155],[106],[209],[131],[6],[130],[71],[176],[165],[76],[28],[190],[103],[146],[181],[113],[213],[182],[67],[66],[175],[183],[88],[65],[206, 207],[136],[145],[38],[168],[161]
Partly	S1,S12,S16,S18,S24,S37,S38,S39,S44,S45,S46,S53,S57,S64,S65,S66,S68,S69,S70,S73,S74,S78,S80,S83,S84,S88,S89,S91,S92
Rarely	[193, 194],[202],[62],[121],[186],[1],[125],[141],[166],[15],[74],[5],[138],[42],[9],[124],[47],[204],[123],[129],[134],[150],[147],[205],[160],[19],[105],[87],[163],[31],[197],[214],[63],[39],[41],[40],[97],[59],[26]
Yes	S2,S3,S4,S5,S6,S7,S8,S9,S10,S11,S13,S14,S15,S17,S19,S20,S21,S22,S23,S25,S26,S27,S28,S29,S30,S31,S32,S33,S34,S35,S36,S40,S41,S42,S43,S47,S48,S49,S50,S51,S52,S54,S55,S56,S58,S59,S60,S61,S62,S63,S67,S71,S72,S75,S76,S77,S79,S81,S82,S85,S86,S87,S90,S93,S94,S95, S96, S97,S98

## Appendix 2

**Table 23** Quality Assessment Questions

---

Questions

- Q1. Are the aims of the research clearly stated and are relevant to the objective of this review article?
- Q2. Does the research study describe the research context and recognition technique adequately?
- Q3. Does the research article explain the approach and methodology of the research with clarity?
- Q4. Is the experiment applied and tested on sufficient dataset(s)?
- Q5. Is the output of the proposed method significant accuracy results?
- Q6. How significant is the published work?
- Q7. Does the research study belong to a journal or a conference?
- Q8. Does the study have a sufficient number of average citation count?
- 

**Data Availability** My manuscript has no associated data.

## Declarations

**Conflicts of interest** The authors declare that they have no conflict of interest. The authors declare that no potential conflicts of interest (financial or non-financial).

## References

1. Afshan N, Afshar Alam M, Ali Mehdi S (2017) An Analysis of Mathematical Expression Recognition Techniques. *Int J Adv Res Comput Sci* 8(5):2021–2026 [www.ijarcs.info](http://www.ijarcs.info)
2. Aggarwal R, Harit G, Tiwari A (2020) Structural analysis of offline handwritten mathematical expressions. *Proceedings of 3rd International Conference on Computer Vision and Image Processing*. (Springer, Singapore) 1024:213–225
3. Ahmed M, Ward R, Kharna N (2004) Solving mathematical problems using knowledge-based systems. *Math Comput Simul* 67(1–2):149–161. <https://doi.org/10.1016/j.matcom.2004.05.015>
4. Ali A, Gravino C (2019) A systematic literature review of software effort prediction using machine learning methods. *Journal of Software: Evolution and Process* 31(10):1–25. <https://doi.org/10.1002/smr.2211>
5. Alvaro F (2015) Mathematical expression recognition based on probabilistic grammars [Departamento de Sistemas Informáticos y Computación, Universitat Politècnica de València]. 10.1.1.1031.2330
6. Álvaro F, Sánchez JA (2010) Comparing several techniques for offline recognition of printed mathematical symbols. *International Conference on Pattern Recognition* 1953–1956. <https://doi.org/10.1109/ICPR.2010.481>
7. Álvaro F, Sánchez JA, Benedí JM (2011) Recognition of printed mathematical expressions using two-dimensional stochastic context-free grammars. *International Conference on Document Analysis and Recognition*, 1225–1229. <https://doi.org/10.1109/ICDAR.2011.247>
8. Álvaro F, Sánchez JA, Benedí JM (2012) Unbiased evaluation of handwritten mathematical expression recognition. *International Conference on Frontiers in Handwriting Recognition* 181–186. <https://doi.org/10.1109/ICFHR.2012.287>
9. Álvaro F, Sánchez JA, Benedí JM (2013) An image-based measure for evaluation of mathematical expression recognition. *Iberian Conference on Pattern Recognition and Image Analysis IbPRIA 2013: Pattern Recognition and Image Analysis*. LNCS 7887:682–690. [https://doi.org/10.1007/978-3-642-38628-2\\_81](https://doi.org/10.1007/978-3-642-38628-2_81)

10. Álvaro F, Sánchez JA, Benedí JM (2014) Recognition of on-line handwritten mathematical expressions using 2D stochastic context-free grammars and hidden Markov models. *Pattern Recogn Lett* 35(1):58–67. <https://doi.org/10.1016/j.patrec.2012.09.023>
11. Alvaro F, Sanchez JA, Benedí JM, Sánchez J-A, Benedí JM (2014) Offline features for classifying handwritten math symbols with recurrent neural networks. 22nd International Conference on Pattern Recognition, 2944–2949. <https://doi.org/10.1109/ICPR.2014.507>
12. Álvaro F, Sánchez JA, Benedí JM (2016) An integrated grammar-based approach for mathematical expression recognition. *Pattern Recogn* 51:135–147. <https://doi.org/10.1016/j.patcog.2015.09.013>
13. Aly W, Uchida S, Suzuki M (2008) Identifying subscripts and superscripts in mathematical documents. *Math Comput Sci* 2(2):195–209. <https://doi.org/10.1007/s11786-008-0051-9>
14. Aly W, Uchida S, Fujiyoshi A, Suzuki M (2009) Statistical classification of spatial relationships among mathematical symbols. 2009 10th International Conference on Document Analysis and Recognition, i, 1350–1354. <https://doi.org/10.1109/ICDAR.2009.90>
15. Anderson RH (1967) Syntax-directed recognition of hand-printed two-dimensional mathematics. Symposium on interactive systems for experimental applied mathematics: proceedings of the Association for Computing Machinery Inc. Symposium 436–459. <https://doi.org/10.1145/2402536.2402585>
16. Arora S, Bhattacharjee D, Nasipuri M, Basu DK, Kundu M (2008) Combining multiple feature extraction techniques for Handwritten Devnagari Character recognition. Proceedings of IEEE Region 10 Colloquium and 3rd International Conference on Industrial and Information Systems, ICIS 2008 1–6. <https://doi.org/10.1109/ICIINF.2008.4798415>
17. Ashiquzzaman A, Tushar AK, Rahman A (2017) Handwritten Arabic numeral recognition using deep learning neural networks. Proceedings of IEEE International Conference on Imaging, Vision and Pattern Recognition, ICIVPR 2017(2):1–4. <https://doi.org/10.1109/ICIVPR.2017.7890866>
18. Awal AM, Mouchère H, Viard-Gaudin C (2009) Towards handwritten mathematical expression recognition. 2009, 10th International Conference on Document Analysis and Recognition 1046–1050. <https://doi.org/10.1109/ICDAR.2009.71>
19. Awal A-MM, Mouchère H, Viard-Gaudin C (2010) The problem of handwritten mathematical expression recognition evaluation. 12th International Conference on Frontiers in Handwriting Recognition 646–651. <https://doi.org/10.1109/ICFHR.2010.106>
20. Awal A-M, Mouchère H, Viard-Gaudin C (2010a) A hybrid classifier for handwritten mathematical expression recognition. *Document Recognition and Retrieval XVII* 7534:753410. <https://doi.org/10.1117/12.840023>
21. Awal A-M, Mouchère H, Viard-Gaudin C (2010b) Improving online handwritten mathematical expressions recognition with contextual modeling. Twelfth International Conference on Frontiers in Handwriting Recognition 427–432. <https://doi.org/10.1109/ICFHR.2010.73>
22. Awal AM, Mouchère H, Viard-Gaudin C (2014) A global learning approach for an online handwritten mathematical expression recognition system. *Pattern Recogn Lett* 35(1):68–77. <https://doi.org/10.1016/j.patrec.2012.10.024>
23. Bage DD, Adhiya KP, Gharde SS (2010) a New Approach for Recognizing Offline Handwritten Mathematical Symbols using character geometry. *International Journal of Innovative Research in Science, Engineering and Technology* 2(7):2823–2830
24. Baker JB, Sexton AP, Sorge V (2010) Faithful mathematical formula recognition from PDF documents. 9th IAPR International Workshop on Document Analysis Systems 485–492. <https://doi.org/10.1145/1815330.1815393>
25. Basu S, Das N, Sarkar R, Kundu M, Nasipuri M, Basu DK (2005) An MLP based approach for recognition of handwritten “Bangla” numerals. Proceedings of the 2nd Indian International Conference on Artificial Intelligence, IICAI 2005 407–417
26. Belaid A, Haton JP (1984) a syntactic approach for handwritten mathematical formula recognition. *IEEE Trans Pattern Anal Mach Intell* PA(1):105–111. <https://doi.org/10.1109/TPAMI.1984.4767483>
27. Bharambe M (2015) recognition of offline handwritten mathematical expressions. National Conference on Digital Image and Signal Proceeding 35–39
28. Bott JN, Gabriele D, LaViola JJ (2011) Now or later: An initial exploration into user perception of mathematical expression recognition feedback. Proceedings - SBIM 2011: ACM SIGGRAPH / Eurographics Symposium on Sketch-Based Interfaces and Modeling, 1:125–132. <https://doi.org/10.1145/2021164.2021187>
29. Brereton P, Kitchenham BA, Budgen D, Turner M, Khalil M (2007) Lessons from applying the systematic literature review process within the software engineering domain. *J Syst Softw* 80(4):571–583. <https://doi.org/10.1016/j.jss.2006.07.009>

30. Bryant A, Charmaz K, Urquhart C (2012) The evolving nature of grounded theory method: The case of the information systems discipline. In: The SAGE Handbook of Grounded Theory <https://doi.org/10.4135/9781848607941.n16>
31. Büyükbayrak H, Yanikoglu B (2007) Online handwritten mathematical expression recognition. Document recognition and retrieval XIV (International Society for Optics and Photonics), 6500, 65000
32. Campanelli AS, Parreiras FS (2015) Agile methods tailoring - A systematic literature review. *J Syst Softw* 110:85–100. <https://doi.org/10.1016/j.jss.2015.08.035>
33. Cao Y, Xie Z, Li L (2019) Research on identification of handwritten mathematical formulas. International Conference on Applications and Techniques in Cyber Security and Intelligence, Springer, Cham., 1, 1494–1500.
34. Celar S, Stojkic Z, Seremet Z, Marusic Z, Zelenika D (2015) Classification of test documents based on handwritten student ID's characteristics. *Procedia Engineering* 100:782–790. <https://doi.org/10.1016/j.proeng.2015.01.432>
35. Celik M, Yanikoglu B (2011) Probabilistic mathematical formula recognition using a 2D context-free graph grammar. International Conference on Document Analysis and Recognition 161–166. <https://doi.org/10.1109/ICDAR.2011.41>
36. Chajri Y, Bouikhalene B (2016) Handwritten mathematical expressions recognition. *International Journal of Signal Processing, Image Processing and Pattern Recognition* 9(5):69–76. <https://doi.org/10.14257/ijsp.2016.9.5.07>
37. Chan C (2020) Stroke extraction for offline handwritten mathematical expression recognition. *IEEE Access* 8:61565–61575. <https://doi.org/10.1109/ACCESS.2020.2984627>
38. Chan K-F, Yeung DYD (2000a) An Efficient syntactic approach to structural analysis of on-line handwritten mathematical expressions. *Pattern Recogn* 33(3):375–384. [https://doi.org/10.1016/S0031-3203\(99\)00067-9](https://doi.org/10.1016/S0031-3203(99)00067-9)
39. Chan KF, Yeung DY (2000b) Mathematical expression recognition: A survey. *International Journal on Document Analysis and Recognition(IJDAR)* 3(1):3–15. <https://doi.org/10.1007/PL00013549>
40. Chan K-F, Yeung D (2001a) Pencil: A novel application of on-line mathematical expression recognition technology. *Proceedings of Sixth International Conference on Document Analysis and Recognition*. IEEE, 274–278
41. Chan K-F, Yeung DY (2001b) Error detection, error correction and performance evaluation in on-line mathematical expression recognition. *Pattern Recogn* 34(8):1671–1684. [https://doi.org/10.1016/S0031-3203\(00\)00102-3](https://doi.org/10.1016/S0031-3203(00)00102-3)
42. Chaudhuri BB, Garain U (2013) An approach for recognition and interpretation of mathematical expressions in printed document. *Pattern Analysis & Applications* 3(2):120–131
43. Chen Y, Okada M (2001) Structural analysis and semantic understanding for offline mathematical expressions. *Int J Pattern Recognit Artif Intell* 15(EC06):967–987. <https://doi.org/10.1142/S021800140100126X>
44. Chen X, Sun H, Tobe Y, Zhou Z, Sun X (2015) Coverless information hiding method based on the chinese mathematical expression. *Proceedings of International Conference on Cloud Computing and Security*, Springer Cham 9483:133–143. <https://doi.org/10.1007/978-3-319-27051-7>
45. Choudhary A, Ahlawat S, Gupta H, Bhandari A, Dhall A, Kumar M (2021) Offline handwritten mathematical expression evaluator using convolutional neural network. *International Conference on Innovative Computing and Communications*, pp. 527–537
46. Clark R, Kung Q, Van Wyk A (2013) System for the recognition of online handwritten mathematical expressions. *Eurocon 2013*:2029–2035. <https://doi.org/10.1016/j.ympev.2006.04.014>
47. Code C, Asst EO, Naik B (2013) A shape-based layout descriptor for classifying spatial relationships in handwritten math. *Proceedings of the 2013 ACM Symposium on Document Engineering* pp 123–126
48. Corbin J, Strauss A, Department of Social and Behavioral Sciences. University of California (2016) *Grounded Theory Research: Procedures, Canons and Evaluative Criteria*. *Qual Sociol* 19(6):418–427. <https://doi.org/10.1515/zfsoz-1990-0602>
49. Da Silva W, Silva CE (2014) Mathematical models to describe thin-layer drying and to determine drying rate of whole bananas. *J Saudi Soc Agric Sci* 13(1):67–74 <https://www.sciencedirect.com/science/article/pii/S1658077X13000040>
50. Dagenais B, Ossher H, Bellamy RKE, Robillard MP, De Vries JP (2009) A qualitative study on project landscapes. *Proceedings of the 2009 ICSE Workshop on Cooperative and Human Aspects on Software Engineering, CHASE*. 2009:32–35. <https://doi.org/10.1109/CHASE.2009.5071407>

51. Dai Nguyen H, Duc Le A, Nakagawa M (2016) Recognition of online handwritten math symbols using deep neural networks. *IEICE Trans Inf Syst*, 3110–3118. <https://doi.org/10.1587/transinf.2016E DP7102>
52. Dai J, Sun Y, Su G, Ye S, Sun Y (2019) Recognizing offline handwritten mathematical expressions efficiently. 10th International Conference on E-Education, E-Business, E-Management and E-Learning, pp 198–204. <https://doi.org/10.1145/3306500.3306543>
53. Davila K, Zanibbi R (2017) Layout and semantics: Combining representations for mathematical formula search. *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval(SIGIR)*, pp 1165–1168. <https://doi.org/10.1145/3077136.3080748>
54. Drsouza L, Mascarenhas M (2018) Offline handwritten mathematical expression recognition using convolutional neural network. *International Conference on Information, Communication, Engineering and Technology*, pp 1–3. <https://doi.org/10.1109/ICICET.2018.8533789>
55. Eto Y, Suzuki M (2001) Mathematical formula recognition using virtual link network. *Proceedings of Sixth International Conference on Document Analysis and Recognition*, pp 762–767. <https://doi.org/10.1109/icdar.2001.953891>
56. Fahmy H, Blostein D (1993) A graph grammar programming style for recognition of music notation. *Mach Vis Appl* 6(2–3):83–99. <https://doi.org/10.1007/BF01211933>
57. Fang D, Zhang C (2020) Multi-feature Learning by Joint Training for Handwritten Formula Symbol Recognition. *IEEE Access* 8(2):48101–48109. <https://doi.org/10.1109/ACCESS.2020.2979346>
58. Farulla GA, Armano T, Capietto A, Murru N, Rossini R (2016) Artificial neural networks and fuzzy logic for recognizing alphabet characters and mathematical symbols. *International Conference on Computers Helping People with Special Needs*, 7–14. [https://doi.org/10.1007/978-3-319-41264-1\\_1](https://doi.org/10.1007/978-3-319-41264-1_1)
59. Feng X, Shiiba K, Okazaki Y, Okamoto M, Kondo H (2001) Java based on-line handwriting interface for mathematical expression and its character recognition performance character recognition. 85th Technology Research Meeting of JSISE (Japanese Society for Information and Systems in Education), 1–8
60. Fitzgerald J, Geiselbrechtinger F, Kechadi M (2006) Structural analysis of handwritten mathematical expressions through fuzzy parsing. *ACST* 6:151–156
61. Fitzgerald JA, Geiselbrechtinger F, Kechadi T (2007) Mathpad: A fuzzy logic-based recognition system for handwritten mathematics. *Ninth International Conference on Document Analysis and Recognition* 2:694–698. <https://doi.org/10.1109/ICDAR.2007.4377004>
62. Fontenele Marques Junior FDC, Pontes De Araujo T, Moura Sousa JV, Carvalho Da Costa NJ, Teixeira Melo R, Martins Pinto A, Andrade Saraiva A, F MJFC, Ara D, P T (2019) Recognition of simple handwritten polynomials using segmentation with fractional calculus and convolutional neural networks. 8th Brazilian Conference on Intelligent Systems. pp 245–250. <https://doi.org/10.1109/BRACIS.2019.00051>
63. Garain U (2005) Automatic recognition of printed and handwritten mathematical expressions. *Indian Institute of Statistics, Kolkata*
64. Garain U (2009) Identification of mathematical expressions in document images. In: 10th International Conference on Document Analysis and Recognition, pp 1340–1344. <https://doi.org/10.1109/ICDAR.2009.203>
65. Garain U, Chaudhuri B (2002) On development and statistical analysis of a corpus for printed and handwritten mathematical expressions *Computer Science Preprint Archive Available at SSRN: Htpps://Ssrn.Com/Abstract=3125396* 2002(7):689–699
66. Garain U, Chaudhuri BB (2005) A corpus for OCR research on mathematical expressions. *Int J Doc Anal Recognit* 7(4):241–259. <https://doi.org/10.1007/s10032-004-0140-5>
67. Garain U, Chaudhuri B (2007) *OCR of printed mathematical expressions*. Springer, London, *Digital Document Processing*. <https://doi.org/10.1109/isscc.1986.1156961>
68. Garst, P. (2004). Modeless gesture-driven editor for handwritten mathematical expressions (Patent No. US 2004/0054701 A1). In *Mathsoft Engineering & Education Inc*, 2004, U.S. Patent Application No. 10/378,386. (US 2004/0054701 A1). <https://doi.org/10.4324/9781315853178>
69. Genoe R, Fitzgerald J, Kechadi T. (2006a) A purely online approach to mathematical expression recognition. *International Workshop on Frontiers in Handwriting Recognition* pp 1–6. *iffinria-00104890f*
70. Genoe Ray, Fitzgerald JA, Kechadi T (2006b) An online fuzzy approach to the structural analysis of handwritten mathematical expressions. *IEEE International Conference on Fuzzy Systems* pp 244–250. <https://doi.org/10.1109/FUZZY.2006.1681721>
71. Gharde SS (2012) Evaluation of classification and feature extraction techniques for simple mathematical equations. *International Journal of Applied Information Systems* 1(5):34–38

72. Gharde SS, Baviskar PV, Adhiya KP (2013) Identification of handwritten simple mathematical equation based on SVM and projection histogram. *ACM Symposium on Document Engineering* 3(2):425–429
73. Glaser BG, Strauss AL, Strutzel E (1968) The discovery of grounded theory: Strategies for qualitative research. *Nursing Research Journal* 17(4):364
74. Gong Y, Li S, Wang X, Wang X (2015) Real-time recognition method of understanding on-line handwritten mathematical expression. *Computer Engineering and Applications Journal* 7:43
75. Han F, Zhu S (2005) Bottom-up / Top-Down Image Parsing by Attribute Graph Grammar. Tenth IEEE International Conference on Computer Vision 1778–1785. <https://doi.org/10.1109/ICCV.2005.50>
76. Hirata N, Honda W (2011) Automatic labeling of handwritten mathematical symbols via expression matching. *Graph-Based Representations in Pattern Recognition*. GbRPR 2011. *Lecture Notes in Computer Science* pp 295–304. <https://doi.org/10.1177/107808747000500401>
77. Hirata NST, Julca-Aguilar NST, Julca-Aguilar F (2015) Matching based ground-truth annotation for online handwritten mathematical expressions. *Pattern Recogn* 48(3):837–848. <https://doi.org/10.9790/7388-05621214>
78. Hong Z, You N, Tan J, Bi N (2019) Residual BiRNN based Seq2Seq model with transition probability matrix for online handwritten mathematical expression recognition. *International Conference on Document Analysis and Recognition* pp 635–640. <https://doi.org/10.1109/ICDAR.2019.00107>
79. Hordri NF, Samar A, Yuhaziz SS, Shamsuddin SM (2017) A systematic literature review on features of deep learning in big data analytics. *Proceedings of International Journal of Advances in Soft Computing and Its Applications* 9(1):32–49
80. Hossain MB, Naznin F, Joarder YA, Zahidul Islam M, Uddin MJ (2018) Recognition and solution for handwritten equation using convolutional neural network. 2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition pp 250–255. <https://doi.org/10.1109/ICIEV.2018.8640991>
81. Hu L, Zanibbi R (2016) MST-based visual parsing of online handwritten mathematical expressions. 15th International Conference on Frontiers in Handwriting Recognition 337–342. <https://doi.org/10.1109/ICFHR.2016.0070>
82. Hu L, Hart K, Pospelov R, Zanibbi R (2012) Baseline extraction-driven Parsing of handwritten mathematical expressions. 21st International Conference on Pattern Recognition pp 326–330
83. Hu Y, Peng L, Tang Y (2014) On-line handwritten mathematical expression recognition method based on statistical and semantic analysis. 11th IAPR International Workshop on Document Analysis Systems pp 171–175. <https://doi.org/10.1109/DAS.2014.47>
84. Huang BQ, Kechadi MM (2007) A structural analysis approach for online handwritten mathematical expressions. *International Journal of Computer Science and Network Security* 7(7):47–56. [https://doi.org/10.1142/9789812837042\\_0014](https://doi.org/10.1142/9789812837042_0014)
85. Huang BQ, Zhang YB, Kechadi MT (2007) Preprocessing Techniques for Online Handwriting Recognition. Seventh International Conference on Intelligent Systems Design and Applications pp 793–800. <https://doi.org/10.1109/isda.2007.31>
86. Jain C, Zanibbi R (2017). Recognition of Online Handwritten Math Symbols using Density Features
87. Jianyu X, Qianping W, Liurong H (2008) Recognition of mathematical expressions based on convex hull and fuzzy recognition. *Computer Applications and Software* 5:82
88. Jin J, Han X, Wang Q (2003) Mathematical formulas extraction. In *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR* (pp 1138–1141). <https://doi.org/10.1109/ICDAR.2003.1227834>
89. Julca-Aguilar F, Mouchère H, Viard-Gaudin C (2015) Top-down online handwritten mathematical expression parsing with graph grammar. *Proceedings of IberoAmerican Congress on Pattern Recognition*. Springer, Cham. pp 144–151.
90. Kacem A, Belaid A, Ben Ahmed M (2001) Automatic extraction of printed mathematical formulas using fuzzy logic and propagation of context. *Int J Doc Anal Recognit* 4(2):97–108. <https://doi.org/10.1007/s100320100064>
91. Kanahori T, Tabata K, Cong W, Tamari F, Suzuki M (2000) On-line recognition of mathematical expressions using automatic rewriting method. *International Conference on Multimodal Interfaces* pp 394–401. [https://doi.org/10.1007/3-540-40063-x\\_52](https://doi.org/10.1007/3-540-40063-x_52)
92. Khuong VTM, Phan M (2019) Interactive user interface for recognizing online handwritten mathematical expressions and correcting misrecognition. In: *Proceedings of International Conference on Document Analysis and Recognition Workshops (ICDARW)*, IEEE, 2, pp 26–30

93. Khuong VTM, Huy U, Masaki N (2019) Generating synthetic handwritten mathematical expressions from a LaTeX Sequence or a MathML script. *Proceedings of International Conference on Document Analysis and Recognition (ICDAR)*, IEEE, pp 922–927
94. Khuong V-T-m, Member S, Phan K, Ung H (2021) Clustering of handwritten mathematical expressions for computer-assisted marking. *IEICE Trans Inf Syst* 2:275–284
95. Kim K, Rhee TH, Lee JS, Kim JH (2009) Utilizing consistency context for handwritten mathematical expression recognition. *International Conference on Document Analysis and Recognition* 1051–1055. <https://doi.org/10.1109/ICDAR.2009.140>
96. Kitchenham B, Pearl Brereton O, Budgen D, Turner M, Bailey J, Linkman S (2009) Systematic literature reviews in software engineering - A systematic literature review. *Inf Softw Technol* 51(1):7–15. <https://doi.org/10.1016/j.infsof.2008.09.009>
97. Kosmala A, Rigoll G (2000) Online handwritten formula recognition with integrated correction recognition and execution. In: *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000 IEEE*, pp 590–593
98. Sakshi, Kukreja V (2021) A retrospective study on handwritten mathematical symbols and expressions : Classification and recognition. *Eng Appl Artif Intell* 103:104292. <https://doi.org/10.1016/j.engappai.2021.104292>
99. Kumar PP, Agarwal A, Bhagvati C (2018) Isolated structural error analysis of printed mathematical expressions. *Pattern Anal Applic* 21(4):1097–1107. <https://doi.org/10.1007/s10044-017-0667-y>
100. Kumar P, Shreekanth T, Shashank NS, Sneha S (2019) A simplified research for mathematical expression recognition and its conversion to speech. *International Journal of Recent Technology and Engineering* 8(2S8):1033–1038. <https://doi.org/10.35940/ijrte.B1008.0882S819>
101. Kundu S, Paul S, Kumar Bera S, Abraham A, Sarkar R (2020) Text-line extraction from handwritten document images using GAN. *Expert Syst Appl* 140:112916. <https://doi.org/10.1016/j.eswa.2019.112916>
102. Labahn SMG (2013) A new approach for recognizing handwritten mathematics using relational grammars and fuzzy sets. *Int J Doc Anal Recognit* 16(2):139–163. <https://doi.org/10.1007/s10032-012-0184-x>
103. Lapointe A, Blostein D (2009) Issues in performance evaluation: A case study of math recognition. *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR* pp 1355–1359. <https://doi.org/10.1109/ICDAR.2009.247>
104. LaViola JJ, Zeleznik RC (2007) A practical approach for writer-dependent symbol recognition using a writer-independent symbol recognizer. *IEEE Trans Pattern Anal Mach Intell* 29(11):1917–1926. <https://doi.org/10.1109/TPAMI.2007.1109>
105. LaViola JJ, Leal A, Miller TS, Zeleznik RC (2008) Evaluation of techniques for visualizing mathematical expression recognition results. *Proceedings of Graphics Interface* pp 131–138
106. Le A, Nakagawa M (2013) A tool for ground-truthing online handwritten mathematical expressions. *16th International Graphonomics Society Conference*. <https://doi.org/10.9790/487X-171214553>
107. Le AD, Nakagawa M (2015) Improving structure analysis for online handwritten mathematical expressions. *18th Meeting on Image Recognition and Understanding*, 1–2. %60
108. Le AD, Nakagawa M (2016) A system for recognizing online handwritten mathematical expressions by using improved structural analysis. *Int J Doc Anal Recognit* 19(4):305–319. <https://doi.org/10.1007/s10032-016-0272-4>
109. Le A, Nakagawa M (2017) Training an end-to-end system for handwritten mathematical expression recognition by generated patterns. *2017 14th IAPR International Conference on Document Analysis and Recognition* 1:1056–1061. <https://doi.org/10.1109/ICDAR.2017.175>
110. Le AD, Phan T Van, Nakagawa M (2014) A system for recognizing online handwritten mathematical expressions and improvement of structure analysis. *11th IAPR International Workshop on Document Analysis Systems* pp 51–55. <https://doi.org/10.1109/DAS.2014.52>
111. Le AD, Indurkha B, Nakagawa M (2019a) Pattern generation strategies for improving recognition of Handwritten Mathematical Expressions. *Pattern Recogn Lett* 128:255–262. <https://doi.org/10.1016/j.patrec.2019.09.002>
112. Le AD, Nguyen HD, Indurkha B, Nakagawa M (2019b) Stroke order normalization for improving recognition of online handwritten mathematical expressions. *Int J Doc Anal Recognit* 22(1):29–39. <https://doi.org/10.1007/s10032-019-00315-2>
113. Li C, Zeleznik R, Miller T, LaViola JJ (2008) Online recognition of handwritten mathematical expressions with support for matrices. *Proceedings - International Conference on Pattern Recognition* pp 1–4. <https://doi.org/10.1109/icpr.2008.4761825>



114. Lin X, Gao L, Tang Z, Hu X, Lin X (2012) Identification of embedded mathematical formulas in PDF documents using SVM. *Document Recognition and Retrieval XIX* 8297:82970D. <https://doi.org/10.1117/12.912445>
115. Liu C, Zuo L, Li X, Tian X (2016) An improved algorithm for identifying mathematical formulas in the images of PDF documents. *Proceedings of 2015 IEEE International Conference on Progress in Informatics and Computing, PIC 2015* pp 252–256. <https://doi.org/10.1109/PIC.2015.7489848>
116. Liu H, Yu C, Wu H (2021) Smart non-intrusive device recognition based on deep learning methods. In *Smart Device Recognition* (pp 229–258). Springer
117. Lods A, Anquetil E, Mace S (2019) Fuzzy visibility graph for structural analysis of online handwritten mathematical expressions. *International Conference on Document Analysis and Recognition* pp 641–646. <https://doi.org/10.1109/ICDAR.2019.00108>
118. MacLean S, Labahn G (2010) Recognizing handwritten mathematics via fuzzy parsing (Issue Tech. Rep.CS-2010-13)
119. MacLean S, Labahn G (2015) A Bayesian model for recognizing handwritten mathematical expressions. *Pattern Recogn* 48(8):2433–2445. <https://doi.org/10.1016/j.patcog.2015.02.017>
120. MacLean S, Labahn G, Labahn SMG, MacLean S, Labahn G (2013) A new approach for recognizing handwritten mathematics using relational grammars and fuzzy sets. *Int J Doc Anal Recognit* 16(2):139–163. <https://doi.org/10.1007/s10032-012-0184-x>
121. Mahdavi M, Zanibbi R, Mouchere H, Viard-Gaudin C, Garain U (2019) ICDAR 2019 CROHME + TFD: Competition on recognition of handwritten mathematical expressions and typeset formula detection. *International Conference on Document Analysis and Recognition* pp 1533–1538. <https://doi.org/10.1109/ICDAR.2019.00247>
122. Malon C, Uchida S, Suzuki M (2008) Mathematical symbol recognition with support vector machines. *Pattern Recogn Lett* 29(9):1326–1332. <https://doi.org/10.1016/j.patrec.2008.02.005>
123. Medjkoune S, Mouchère H (2012) Using speech for handwritten mathematical expression recognition disambiguation. *International Conference on Frontiers in Handwriting Recognition, IEEE* pp 187–192. <http://plc-solutions.blogspot.com/p/block-diagram-of-plc.html>
124. Medjkoune S, Mouchère H, Petitrenaud S (2013) Multimodal mathematical expressions recognition: Case of speech and handwriting. In: *International Conference on Human-Computer Interaction*. Springer, Berlin, Heidelberg, pp 77–86
125. Medjkoune S, Mouchere H, Petitrenaud S, Viard-Gaudin C (2017) Combining speech and handwriting modalities for mathematical expression recognition. *IEEE Transactions on Human-Machine Systems* 47(2):259–272. <https://doi.org/10.1109/THMS.2017.2647850>
126. Memon J, Sami M, Khan RA (2019) Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR). <http://arxiv.org/abs/2001.00139>
127. Mohan K, Lu C (2013) Recognition of online handwritten mathematical expressions. In: *Stanford University*
128. Mohan K, Lu C (2015) Recognition of online handwritten mathematical expressions using convolutional neural networks. In: *Stanford University*
129. Mouchère H, Viard-Gaudin C, Kim DH, Kim JH, Garain U (2011) CROHME2011: Competition on recognition of online handwritten mathematical expressions. *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR* pp 1497–1500. <https://doi.org/10.1109/ICDAR.2011.297>
130. Mouchère H, Viard-Gaudin C, Kim DH, Kim JH, Garain U (2012) ICFHR 2012 - Competition on Recognition of On-line Mathematical Expressions (CROHME 2012). *Proceedings - International Workshop on Frontiers in Handwriting Recognition, IWFHR, Crohme* pp 811–816. <https://doi.org/10.1109/ICFHR.2012.215>
131. Mouchere H, Viard-Gaudin C, Zanibbi R, Garain U, Kim DH, Kim JH (2013) ICDAR 2013 CROHME: Third international competition on recognition of online handwritten mathematical expressions. *2013 12th International Conference on Document Analysis and Recognition* pp 1428–1432. <https://doi.org/10.1109/ICDAR.2013.288>
132. Mouchere H, Viard-Gaudin C, Zanibbi R, Garain U (2014) ICFHR 2014 Competition on Recognition of On-Line Handwritten Mathematical Expressions (CROHME 2014). *2014 14th International Conference on Frontiers in Handwriting Recognition* pp 791–796. <https://doi.org/10.1109/ICFHR.2014.138>
133. Mouchère H, Viard-Gaudin C, Zanibbi R, Garain U (2016) ICFHR2016 CROHME: Competition on recognition of online handwritten mathematical expressions. *Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR* pp 607–612. <https://doi.org/10.1109/ICFHR.2016.0116>

134. Muñoz FÁ (2010) Off-line Recognition of Printed Mathematical Expressions Using Stochastic Context-Free Grammars. Universidad Politecnica de Valencia
135. Nguyen CT, Khuong VTM, Nguyen HT, Nakagawa M (2020) CNN based spatial classification features for clustering offline handwritten mathematical expressions. *Pattern Recogn Lett* 131:113–120. <https://doi.org/10.1016/j.patrec.2019.12.015>
136. Okamoto M, Imai H, Takagi K (2001) Performance evaluation of a robust method for mathematical expression recognition. Sixth International Conference on Document Analysis and Recognition pp 121–128. <https://doi.org/10.1109/ICDAR.2001.953767>
137. Pacheco-Venegas ND, López G, Andrade-Aréchiga M (2015) Conceptualization, development and implementation of a web-based system for automatic evaluation of mathematical expressions. *Computers and Education* 88:15–28. <https://doi.org/10.1016/j.compedu.2015.03.021>
138. Pattaniyil N, Zanibbi R (2014) Combining TF-IDF Text Retrieval with an Inverted Index over Symbol Pairs in Math Expressions: The Tangent Math Search Engine at {NTCIR 2014}. NTCIR Workshop 11 Meeting, 135–142
139. Pfleeger SL (2005) The Role of Evidential Force in Empirical Software Engineering. *IEEE Softw* 22(1):66–73 <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1377126>
140. Phan K, Nguyen C, Le A (2015) An incremental recognition method for online handwritten mathematical expressions. 3rd IAPR Asian Conference on Pattern Recognition pp 171–175
141. Phan KM, Le AD, Nakagawa M (2016) Semi-incremental recognition of online handwritten mathematical expressions. 15th International Conference on Frontiers in Handwriting Recognition 258–264. <https://doi.org/10.1109/ICFHR.2016.0057>
142. Phan KM, Le AD, Indurkha B, Nakagawa M (2018) Augmented incremental recognition of online handwritten mathematical expressions. *International Journal on Document Analysis and Recognition (IJ DAR)* 21(4):253–268. <https://doi.org/10.1007/s10032-018-0306-1>
143. Phong BH, Hoang TM, Le T-L (2021) Mathematical variable detection in scientific document images. *International Journal of Computational Vision and Robotics* 11(1):66–89
144. Pillay A (2014) Intelligent combination of structural analysis algorithms: Application to mathematical expression recognition. Rochester Institute of Technology
145. Plamondon RR, Srihari SN (2000) On-line and off-line handwriting recognition: A comprehensive survey. *IEEE Trans Pattern Anal Mach Intell* 22(1):63–84. <https://doi.org/10.1109/34.824821>
146. Plötz T, Fink GA (2009) Markov models for offline handwriting recognition: A survey. *Int J Doc Anal Recognit* 12(4):269–298. <https://doi.org/10.1007/s10032-009-0098-4>
147. Predovic G, Abdulkader A, Dresevic B, Viola P, (U.S. Patent No. 8, 009, 915. Washington, D. U. S. P. and T. O. (2011) Recognition of mathematical expressions (Patent No. U.S. Patent No. 8,009,915. Washington, DC: U.S. Patent and Trademark Office.). <https://patentimages.storage.googleapis.com/47/9b/41/6846dbfd8ea793/US8009915.pdf>
148. Průša D, Hlaváč V (2007) Mathematical formulae recognition using 2D grammars. 2017 Ninth International Conference on Document Analysis and Recognition 2:849–853. <https://doi.org/10.1109/ICDAR.2007.4377035>
149. Qi X, Pan W, Yusup WY (2009) The study of structure analysis strategy in handwritten recognition of general mathematical expression. *International Forum on Information Technology and Applications* 2:101–107. <https://doi.org/10.1109/IFITA.2009.169>
150. Quiniou S, Mouchère H, Saldarriaga SP, Viard-gaudin C, Morin E, Petitrenaud S, Medjkoune S (2011) HAMEX - A handwritten and audio dataset of mathematical expressions. Proceedings of the International Conference on Document Analysis and Recognition, ICDAR pp 452–456. <https://doi.org/10.1109/ICDAR.2011.97>
151. Rahmi S, Nadia R, Hasibah B, Hidayat W (2017) The relation between self-efficacy toward math with the math communication. *Infinity Journal* 6(2):177–182. <https://doi.org/10.22460/infinity.v6i2.p177-182>
152. Raman TV (1994) Aster: Audio system for technical readings. *Inf Technol Disabil* 1(4):1–11
153. Ramteke RJ, Mehrotra SC (2006) Feature extraction based on moment invariants for handwriting recognition. *IEEE Conference on Cybernetics and Intelligent Systems 2006*:1–6. <https://doi.org/10.1109/ICCIS.2006.252262>
154. Ramteke S, Patil D, Patil N (2012) Neural network approach to Mathematical Expression Recognition System. *International Journal of Engineering Research & Technology (IJERT)* 7(10):2278–0181
155. Reichenbach MS, Agarwal A, Zanibbi R (2014) Rendering expressions to improve accuracy of relevance assessment for math search. SIGIR 2014 - Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval pp 851–854. <https://doi.org/10.1145/2600428.2609457>

156. Ren H, Wang W, Liu C (2019) Recognizing online handwritten Chinese characters using RNNs with new computing architectures. *Pattern Recogn* 93:179–192. <https://doi.org/10.1016/j.patcog.2019.04.015>
157. Rhee TH, Kim JH (2009) Efficient search strategy in structural analysis for handwritten mathematical expression recognition. *Pattern Recogn* 42(12):3192–3201. <https://doi.org/10.1016/j.patcog.2008.10.036>
158. Rohatgi S, Zhong W, Zanibbi R, Wu J, Giles CL (2019) Query auto completion for math formula search 6–11. <http://arxiv.org/abs/1912.04115>
159. Ruiz V, Linares I, Sanchez A, Velez J (2020) Off-line handwritten signature verification using compositional synthetic generation of signatures and Siamese Neural Networks. *Neurocomputing* 374:30–41. [https://doi.org/10.1007/1-4020-0613-6\\_11008](https://doi.org/10.1007/1-4020-0613-6_11008)
160. Sain K, Dasgupta A, Garain U (2010) EMERS: A tree matching-based performance evaluation of mathematical expression recognition systems. *Int J Doc Anal Recognit* 14(1):75–85. <https://doi.org/10.1007/s10032-010-0121-9>
161. Savchenkov P, Savinov E, Mikhail T, Kiyas S, Esin A (2018) Neural Network Based Recognition of Mathematical Expressions (Patent No. 15 / 187 , 723). In United States Patent (15 / 187 , 723) Google Patents
162. Shan G, Wang H, Liang W, Chen K (2021) Robust encoder-decoder learning framework towards offline handwritten mathematical expression recognition based on multi-scale deep neural network. *SCIENCE CHINA Inf Sci* 64(3):1–12
163. Shi Y, Soong F (2008) A symbol graph-based handwritten math expression recognition. In: 19th International Conference on Pattern Recognition, pp 1–4
164. Shi Y, Li HY, Soong FK (2007) A unified framework for symbol segmentation and recognition of handwritten mathematical expressions. In: 9th International Conference on Document Analysis and Recognition, 2, pp 854–858. <https://doi.org/10.1109/ICDAR.2007.4377036>
165. Shi Y, Soong F, Zhou J (2011) Symbol graph generation in handwritten mathematical expression recognition. In: US Patent No 7,885,456
166. Shinde S, Waghulade R (2016) Handwritten mathematical expressions recognition using back propagation artificial neural network. *Communications on Applied Electronics* 4(7):1–6. <https://doi.org/10.5120/cae2016652125>
167. Shinde S, Waghulade RB, Bormane DS (2018) A new neural network based algorithm for identifying handwritten mathematical equations. *International Conference on Trends in Electronics and Informatics* pp 204–209. <https://doi.org/10.1109/ICOEI.2017.8300916>
168. Shuvo SN, Hasan F, Ahmed MU, Hossain SA, Abujar S (2021) MathNET: Using CNN Bangla Handwritten Digit, Mathematical Symbols, and Trigonometric Function Recognition. In: *Soft Computing Techniques and Applications*. Springer, pp 515–523
169. Simistira F, Papavassiliou V, Katsouros V, Carayannis G (2012). A system for recognition of online handwritten mathematical expressions. *International Conference on Frontiers in Handwriting Recognition*, pp 193–198. <https://doi.org/10.1109/ICFHR.2012.172>
170. Simistira F, Papavassiliou V, Katsouros V, Carayannis G (2014) Recognition of spatial relations in mathematical formulas. In: 14th International Conference on Frontiers in Handwriting Recognition pp 164–168. <https://doi.org/10.1109/ICFHR.2014.35>
171. Simistira F, Katsouros V, Carayannis G (2015) Recognition of online handwritten mathematical formulas using probabilistic SVMs and stochastic context free grammars. *Pattern Recogn Lett* 53:85–92. <https://doi.org/10.1016/j.patrec.2014.11.015>
172. So CMY (2005) An Analysis of mathematical expressions used in practice, technical report [University of Western Ontario]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.367.529&rep=rep1&type=pdf>
173. Stria J, Pruša D, Hlavác V (2014) Combining structural and statistical approach to online recognition of handwritten mathematical formulas. *Nineteenth Computer Vision Winter Workshop* pp 103–109
174. Suzuki T (2000) A new system for the real-time recognition of handwritten mathematical formulas. 15th International Conference on Pattern Recognition, 4 pp 515–518. <https://doi.org/10.1109/icpr.2000.902970>
175. Suzuki M, Uchida S, Nomura A (2005) A ground-truthed mathematical character and symbol image database. *Proceedings of Eighth International Conference on Document Analysis and Recognition (ICDAR)* pp 1–5
176. Takiguchi Y, Okada M, Miyake Y (2005) A fundamental study of output translation from layout recognition and semantic understanding system for mathematical formulae. *Eighth International*

- Conference on Document Analysis and Recognition pp 745–749. <https://doi.org/10.1109/ICDAR.2005.10>
177. Tapia E (2005) Understanding mathematics: A system for the recognition of on-line handwritten mathematical expressions
  178. Tapia E. (2007) Handwritten Mathematical notation a survey on recognition of on-line handwritten mathematical. January, 17
  179. Tapia E, Rojas R (2003) Recognition of on-line handwritten mathematical formulas in the e-chalk system. Seventh International Conference on Document Analysis and Recognition 3:980–984. <https://doi.org/10.1109/ICDAR.2003.1227805>
  180. Tapia E, Rojas R (2004) Recognition of On-line Handwritten Mathematical Expressions Using a Minimum Spanning Tree Construction and Symbol Dominance. International Workshop on Graphics Recognition 3088:329–340. [https://doi.org/10.1007/978-3-540-25977-0\\_30](https://doi.org/10.1007/978-3-540-25977-0_30)
  181. Tapia E, Rojas R (2007a) A survey on recognition of on line handwritten mathematical notation. In: Technical Report B-07-01 Freie Universität Berlin, Institut für Informatik Takustr. 9, 14195 Berlin, Germany
  182. Tapia E, Rojas R (2007b) MathFoR: The Mathematical Formula Recognition System. <http://cites.eerx.ist.psu.edu/viewdoc/download?doi=10.1.1.94.2663&rep=rep1&type=pdf>
  183. Thimbleby W (2004) A better calculator: Processing handwritten mathematical expressions to solve problems
  184. Tian X-D, Li H-Y, Li X-F, Zhang L-P (2006) Research on symbol recognition for mathematical expressions. First International Conference on Innovative Computing, Information and Control pp 357–360. <https://doi.org/10.1109/iciic.2006.506>
  185. Toyozumi K, Yamada N, Kitasaka T (2004) A study of symbol segmentation method for handwritten mathematical formula recognition using mathematical structure information. Proceedings of the 17th International Conference on Pattern Recognition, 2, pp 630–633. <https://doi.org/10.1109/ICPR.2004.1334327>
  186. Tran GS, Huynh CK, Le TS, Phan TP, Bui KN (2018) Handwritten mathematical expression recognition using convolutional neural network. 3rd International Conference on Control, Robotics and Cybernetics pp 15–19. <https://doi.org/10.1109/CRC.2018.00012>
  187. Ung HQ, Vu KTM, Le AD, Nguyen CT, Nakagawa M (2018) Bag-of-features for clustering online handwritten mathematical expressions. International Conference on Pattern Recognition and Artificial Intelligence pp 127–132
  188. Veres O, Rishnyak I, Rishniak H (2019) Application of methods of machine learning for the recognition of mathematical expressions. CEUR Workshop Proceedings 2362:1–12
  189. Vuong BQ, Hui SC, He Y (2008) Progressive structural analysis for dynamic recognition of on-line handwritten mathematical expressions. Pattern Recogn Lett 29(5):647–655. <https://doi.org/10.1016/j.patrec.2007.11.017>
  190. Vuong B-QQ, He Y, Hui SC (2010) Towards a web-based progressive handwriting recognition environment for mathematical problem solving. Expert Syst Appl 37(1):886–893. <https://doi.org/10.1016/j.eswa.2009.05.091>
  191. Wang H, Shan G (2020) Recognizing handwritten mathematical expressions as LaTeX sequences using a multiscale robust neural network. In Computer Vision and Pattern Recognition (Issue 37)
  192. Wang J, Du J, Zhang J, Wang ZR (2019) Multi-modal attention network for handwritten mathematical expression recognition. Proceedings of the International Conference on Document Analysis and Recognition, ICDAR pp 1181–1186. <https://doi.org/10.1109/ICDAR.2019.00191>
  193. Wang J, Du J, Zhang J (2020a) Stroke constrained attention network for online handwritten mathematical expression recognition. Pattern Recogn 119:1–29 <http://arxiv.org/abs/2002.08670>
  194. Wang Z, Du J, Wang J (2020b) Writer-aware CNN for parsimonious HMM-based offline handwritten Chinese text recognition. Pattern Recogn 100(10702). <https://doi.org/10.1016/j.patcog.2019.107102>
  195. Wangperawong A (2018) Attending to Mathematical Language with Transformers
  196. Wen J, Li S, Lin Z, Hu Y, Huang C (2012) Systematic literature review of machine learning based software development effort estimation models. Inf Softw Technol 54(1):41–59. <https://doi.org/10.1016/j.infsof.2011.09.002>
  197. Wu W, Li F, Kong J, Hou L, Zhu B (2006) A bottom-up OCR system for mathematical formulas recognition. International Conference on Intelligent Computing pp 274–279. [https://doi.org/10.1007/11816157\\_27](https://doi.org/10.1007/11816157_27)
  198. Wu JW, Yin F, Zhang YM, Zhang XY, Liu CL (2020) Handwritten mathematical expression recognition via paired adversarial learning. Int J Comput Vis <https://doi.org/10.1007/s11263-020-01291-5>

199. Wu J, Yin F, Zhang Y, Zhang X, Liu C (2021) Graph-to-Graph : Towards Accurate and Interpretable Online Handwritten Mathematical Expression Recognition. AAAI Conference on Artificial Intelligence 35:2925–2933
200. Xiangwei Q, Abaydulla Y (2010) The study of mathematical expression recognition and the embedded system design. *Journal of Software* 5(1):44–53. <https://doi.org/10.4304/jsw.5.1.44-53>
201. Yamamoto R, Sako S, Nishimoto T, Sagayama S (2006) On-line recognition of handwritten mathematical expressions based on stroke-based stochastic context-free grammar. Tenth International Workshop on Frontiers in Handwriting Recognition
202. Yan L, Ratra P, Khanna H, Yan L (2019) Recognizing handwritten mathematical expressions. *International Journal of Engineering Applied Science and Technology* 4(3):1–7
203. Yogatama BW, Lee J, Harimurti S, Adiono T (2018) FPGA-based optical character recognition for handwritten mathematical expressions. *International SoC Design Conference* pp 125–126. <https://doi.org/10.1109/ISOC.2018.8649966>
204. Zanibbi R, Blostein D (2012) Recognition and retrieval of mathematical expressions. *Int J Doc Anal Recognit* 15(4):331–357. <https://doi.org/10.1007/s10032-011-0174-4>
205. Zanibbi R, Yuan B (2011) Keyword and image-based retrieval of mathematical expressions. *Document Recognition and Retrieval XVIII*, 78740I. <https://doi.org/10.1117/12.873312>
206. Zanibbi R, Blostein D, Cordy JR (2001a) Baseline structure analysis of handwritten mathematics notation. *Sixth International Conference on Document Analysis and Recognition* 768–773. <https://doi.org/10.1109/ICDAR.2001.953892>
207. Zanibbi R, Novins K, Arvo J, Zanibbi K (2001b) Aiding manipulation of handwritten mathematical expressions through style-preserving morphs. *Proceedings - Graphics Interface* 127–134
208. Zanibbi R, Blostein D, Cordy JR (2002) Recognizing mathematical expressions using tree transformation. *IEEE Trans Pattern Anal Mach Intell* 24(11):1455–1467. <https://doi.org/10.1109/TPAMI.2002.1046157>
209. Zanibbi R, Mouchère H, Viard-Gaudin C (2013) Evaluating structural pattern recognition for handwritten math via primitive label graphs. *Document Recognition and Retrieval XX* 8658:865817. <https://doi.org/10.1117/12.2008409>
210. Zanibbi R, Davila K, Kane A, Tompa FW (2016a) Multi-stage math formula search: Using appearance-based similarity metrics at scale. *SIGIR 2016 - Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval* pp 145–154. <https://doi.org/10.1145/2911451.2911512>
211. Zanibbi R, Hu L, Zanibbi R (2016b) Line-of-sight stroke graphs and Parzen shape context features for handwritten math formula representation and symbol segmentation. *15th International Conference on Frontiers in Handwriting Recognition* pp 180–186. <https://doi.org/10.1109/ICFHR.2016.0044>
212. Zhang T (2017) *New Architectures for Handwritten Mathematical Expressions Recognition*. Image Processing [Eess.IV]. Université de Nantes, 2017., English. f.
213. Zhang J, Hong L (2008) A survey on recognition of on-line handwritten mathematical expression. *Journal of Huaibei Coal Industry Teachers College (Natural Science Edition)* 29(3)
214. Zhang L, Blostein D, Zanibbi R (2005) Using fuzzy logic to analyze superscript and subscript relations in handwritten mathematical expressions. *Eighth International Conference on Document Analysis and Recognition* pp 972–976. <https://doi.org/10.1109/ICDAR.2005.250>
215. Zhang T, Mouchere H, Viard-Gaudin C (2016) Online handwritten mathematical expressions recognition by merging multiple 1D interpretations. *15th International Conference on Frontiers in Handwriting Recognition* pp 187–192. <https://doi.org/10.1109/ICFHR.2016.0045>
216. Zhang J, Du J, Dai L (2017) A GRU-Based encoder-decoder approach with attention for online handwritten mathematical expression recognition. *Fourteenth IAPR International Conference on Document Analysis and Recognition* pp 902–907. <https://doi.org/10.1109/ICDAR.2017.152>
217. Zhang J, Du J, Dai L (2018a) Track, Attend, and Parse (TAP): An End-to-End Framework for Online Handwritten Mathematical Expression Recognition. *IEEE Transactions on Multimedia* 21(1):221–233. <https://doi.org/10.1109/TMM.2018.2844689>
218. Zhang T, Mouchère H, Viard-Gaudin C (2018b) A tree-BLSTM-based recognition system for online handwritten mathematical expressions. *Neural Comput & Applic* 2(1). <https://doi.org/10.1007/s00521-018-3817-2>
219. Zhang W, Bai Z, Zhu Y (2019) An improved approach based on CNN-RNNs for mathematical expression recognition. *4th International Conference on Multimedia Systems and Signal Processing* 57–61. <https://doi.org/10.1145/3330393.3330410>
220. Zhelezniakov D, Cherneha A, Zaytsev V, Ignatova T, Radyvonenko O, Yakovchuk O (2020a) Evaluating new requirements to pen-centric intelligent user interface based on end-to-end mathematical

- expressions recognition. International Conference on Intelligent User Interfaces pp 212–220. <https://doi.org/10.1145/3377325.3377482>
221. Zhelezniakov D, Zaytsev V, Radyvonenko O (2020b) Online handwritten mathematical expression recognition and applications: A survey. *IEEE Access* 24:1–24
  222. Zhelezniakov D, Zaytsev V, Radyvonenko O (2021) Online handwritten mathematical expression recognition and applications : A survey 9

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.