Original Research Paper

A CNN-KNN Based Recognition of Online Handwritten Symbols within Physics Expressions Using Contour-Based Bounding Box (CBBS) Segmentation Technique

[1]Ujwala Kolte, [1]Sachin Naik and [2]Vidya Kumbhar

[1]*Symbiosis Institute of Computer Studies and Research, Symbiosis International (Deemed University), Pune, India*
[2]*Symbiosis Institute of Geoinformatics, Symbiosis International (Deemed University), Pune, India*

Corresponding Author:
Sachin Naik
Symbiosis Institute of
Computer Studies and
Research, Symbiosis
International (Deemed
University), Pune, India
Email: sachin.naik@sicsr.ac.in

Abstract: The task of recognizing symbols poses a significant challenge owing to the wide variability in human handwriting. Complexity in terms of the structural representation of symbols used in physics expressions is a major challenge in the recognition process The emergence of online handwriting, fueled by the widespread adoption of handheld digital devices, particularly in educational contexts, highlights the critical importance of precise symbol recognition, especially in the teaching and learning process. In contemporary literature, there is a notable emphasis on LaTex sequencing, symbol recognition and parsing. However, deep learning continues to yield promising results in this domain. The convenience of user input provides benefits to e-learning applications. In this study, we propose three approaches for the recognition of physics symbols within physics expressions (1) A proposed Java user interface for taking input from the user, as convenience of user input provides benefits to e-learning applications. (2) Contour-based bounding box segmentation algorithm, which deals with broken symbols within physics expressions. (3) For recognition, we propose a Convolution Neural Network-K-Nearest Neighbor (CNN-KNN) recognition model, as CNN plays an important role in extracting features, which are further provided as input to the K-NN classifier using the dropout method. Combining these three approaches into a symbol recognition model provides state-of-arts results. Handwritten physics symbols were collected from 20 different writers and each writer has written 5 types of physics expressions under different categories like electric flux, Maxwell's equations, inductance and pointing vector and moment of Interia. There were 25 classes identified from the 780 samples collected from the users. The recognition rate is identified using (1) Using CNN model, which shows an accuracy of 91.48 and (2) Using the proposed hybrid CNN-KNN model the accuracy reported is 98.06.

Keywords: CNN, Contour Based Bounding Box Segmentation (CBBS), K-NN, Physics Expression, Symbol Recognition

Introduction

During the last two decades, online and offline data classification has been a major research area. This study particularly focuses on a variety of symbols used in engineering and scientific documents. Especially post-COVID pandemic online e-learning has flown in a new dimension of the teaching and learning process. Online handwritten devices such as tablets and mobile applications are widely used. Physics formulas are widely used in engineering and scientific documents.

The importance of mathematics is in all areas including physics expressions (Nguyen *et al*., 2021) which is still an untouched area by the researchers. Several authors propose recognition systems for online and offline handwritten mathematical symbols, musical symbols and graphics symbols. Physics symbols consist of characters, symbols and numbers, which have little distinction from the basic mathematical symbols. The researchers study symbol recognition, still it is still a challenging task (Ptucha *et al*., 2019). Many researchers use mathematical expressions written and coded using

LateX or MathML as input to the recognition process (Shuvo *et al.* 2021), which is difficult to understand for the users. A recent study focuses on deep learning approaches for the mathematical symbol recognition process which has promising results (Li *et al.*, 2024). Researchers propose symbol segmentation recognition techniques for online, offline and audio mathematical expression recognition (Kherdekar *et al.*, 2023) Similar to mathematical expressions, physics expressions have different categories of expressions. In view of this considering the complexity of symbol structures, recognition of symbols is a challenging task. The complexity varies in terms of the occurrence of different symbols as superscript and subscript positions in the expressions. This study aims to segment four categories of physics expressions (1) Electric flux, (2) Maxwell's equations, (3) Inductance (4) Pointing vector. The proposed contour-based bounding box segmentation algorithm segments the input symbols from the physics expressions and further these are given to the proposed CNN-KNN classification model with the dropout method. The subsequent sections will cover a literature review, methodology, the proposed Contour Bounding Box Segmentation (CBBS) algorithm, CNN-KNN-based classification, results and discussion. Integrating the dropout method into the CNN-KNN classification model utilized in this study is crucial. Dropout acts as a regularization technique frequently employed in deep learning architectures to counter overfitting and enhance generalization capabilities. During the training phase, dropout selectively deactivates a portion of neurons, fostering redundancy within the network and promoting the acquisition of more robust and adaptable features. Specifically within the CNN-KNN framework, dropout is implemented within the convolutional layers to discourage the co-adaptation of feature detectors, thus improving the model's ability to generalize across diverse datasets. By harnessing dropout, the CNN-KNN classification model maintains resilience and flexibility, essential qualities that bolster its effectiveness in precisely identifying and categorizing physics symbols within expressions.

While existing research primarily concentrates on recognizing mathematical symbols, this study fills a notable void by directing its attention toward the recognition of physics expressions, an area that has received comparatively less attention from scholars. Furthermore, the integration of advanced deep learning techniques, exemplified by the CNN-KNN classification model proposed herein, capitalizes on recent progress in the field and exhibits encouraging outcomes. Employing the contour-based bounding box segmentation algorithm to delineate symbols, coupled with the robust capabilities of deep learning, the research endeavors to surmount the complexities inherent in physics symbols' intricate structures, encompassing variations in superscript and subscript placements within expressions. This focused approach, centered on four distinct categories of physics expressions, emphasizes the study's thoroughness and precision, laying a strong groundwork for tackling the nuances of symbol recognition within this domain. Through a comprehensive methodology encompassing literature review, algorithmic development and empirical validation, this research aims to offer valuable insights into symbol recognition within physics expressions.

Literature Review

A three-stage framework to recognize geological symbols from geological maps using deep learning has been proposed by Coquenet *et al.* (2023). The framework introduced comprises three essential elements: The automated assembly of datasets, the training of a Convolutional Recurrent Neural Network (CRNN) model and the establishment of a geo-symbol index. Among the symbols present in the dataset, images are uppercase and lowercase letters, Greek letters, numerical subscripts and special symbols. To assess the effectiveness of the novel deep geological symbol technique, an experimentation phase was conducted utilizing automatically generated datasets created through a randomized combination strategy. Various analyses were performed on the proposed approach's module parameters and subsequently compared. Impressively, when employing the optimal hyperparameters, an average precision score of 94% was attained. This outcome not only signifies commendable feature extraction performance but also showcases superiority over the baseline CNN-based method (Qiu *et al.*, 2023). A variation of the YOLO, an object detection framework that is specialized for the detection of symbols that are not uniform symbols and regions demonstrated in historic maps (Smith and Pillatt, 2023).

An innovative approach has been introduced for classifying Javanese characters, involving a combination of Convolutional Neural Networks (CNN) and Support Vector Machine (SVM) with dropout regularization. The assessment encompassed three distinct CNN architectures and hybrid CNN-SVM models, evaluating both classification accuracy and training time. The most notable outcome was achieved by the hybrid model, which combined the third CNN architecture with the SVM classifier, resulting in an impressive accuracy rate of 98.35% in classifying the testing data. This amalgamated CNN-SVM approach effectively elevates the accuracy of Javanese character recognition, showcasing its potential for enhancement in this domain (Putri *et al.*, 2023). A robust approach for the classification of handwritten musical symbols using Convolution Neural Networks (CNN), k-nearest Neighbor (kNN), Support Vector Machine (SVM) and Random Forest (RaF) is proposed and three network

topologies namely a baseline network, generic object recognition and HMSnet (Baró *et al*., 2019). A fast floating search method with a mixed classifier which is a combination of HMM and Multilayer Perceptron neural network (MLP) proposed, this study uses a UNIPEN dataset of size 10×1000 samples of digits, 840×26 samples of uppercase characters and 1440×26 samples of lower case letters used as dataset. This study has 97.45% accuracy for digits, 91.81% for uppercase characters and 91.03% for lowercase characters. This proposed model used to optimize large-scale feature sets in the future (Huang *et al*., 2006). A stroke-based recognition of musical symbols has been proposed. In this study, a data set has been collected using a pen-based computer interface. Two feature extraction methods are time-series features and the second is stroke-based methods are used. The recognition accuracy is 97.60 and 98.80% for 7666 test strokes and 250 test symbols (Miyao and Maruyama, 2007). Symbol recognition algorithm for online hand-drawn graphic symbols based on the hidden Markov model has been studied using global distance measures and local angle features which are applied to rearranged drawing points. This study achieved an 85% recognition rate for the graphic library of 110 symbols (Xin *et al*., 2003). A multi-class classifier SVM is used to get probabilistic output. Two multi-class SVM classifiers run in parallel using a weighted sum. A symbol set consists of 137 and results for online recognizer perform better with assigning little more weight. In the future authors want to integrate context information with the proposed recognition system (Keshari and Watt, 2007). To recognize chemical expressions, the authors proposed an online hybrid Support Vector Machine Elastic Matching (SVMEM) approach. The proposed method contains five processes. Stroke partitioning; stroke preprocessing, SVM recognition, elastic matching variation and user feedback. This method is evaluated on the chemical elastic symbol library. The average accuracy of the hybrid SVM-EM method is 89.7% (Tang *et al*., 2013). To support user personalization, a new method for online shape recognition has been investigated. This method is based on a two-layered segmental HMM architecture. The recognition system operates on a stroke-level representation of characters and the representation is based on a dictionary of base shapes, which is used to describe various graphical signs. Unipen and the Kaist databases are used for performance evaluation. The recognition system shows an accuracy of 94% for 200 writers (Artieres *et al*., 2007). A new approach for recognition of hand-drawn graphical sketches with structure is proposed, this method is comprised of two phases. The first phase detects some mutually conflicting symbols. Then using the max-sum problem, the best interpretation of input is selected. This method causes fewer recognition errors. The experiments are evaluated on the FC database. The achieved accuracy is 82.7%. The authors reported that this method has good potential and can compete with state-of-the-art methods (Bresler *et al*., 2013). The study has been conducted to recognize open vocabulary, isolated, online handwritten Tamil words and paragraphs of writing (Urala *et al*., 2014). A comprehensive study has been conducted on the strokes while writing an expression, in context with normalizing the order of strokes, which helps in improving the recognition rate of mathematical formulae. This proposed figure outs vertical symbols and upper and lower MEs and divides input strokes into vertical and horizontal strokes. For stroke order, the normalization X-Y cut method is used. The method is tested on the CHROME 2014 database and has given 92% accuracy. This stroke order normalization system shows a remarkable improvement in the results from 35.80-37.63% (Le *et al*., 2019). A study on Tamil symbol recognition, comprising allocating multiple experts to review certain decisions of the primary support vector machine classifier in order to reduce the error rate. These techniques are especially effective in resolving ambiguities that commonly occur in base consonants, pure consonants and vowel modifiers. With the implementation of these techniques, a significant reduction in ambiguities related to these elements is achieved. Dynamic Time Warping (DTW) technique automatically fetches their discriminative regions to deal with confused pairs. The reevaluation methods are tested on the IWFHR test set and the symbols are segmented from a set of 10000 Tamil words. The recognition rate is improved by 1.9% for isolated test symbols of the IWFHR database (Sundaram and Ramakrishnan, 2014). The recognition of handwritten symbols in context with mathematical expressions is an important task in optical character recognition, which includes structural parsing and symbol recognition (Li *et al*., 2024).

Based on the literature review, most of the studies are focused on geological symbols, while there is a lack of research with respect to other domains like Physics expressions and chemical expressions. Handling non-uniform symbols is the challenge posed by irregularly shaped symbols in various contexts. There will be more exploration opportunities for hybrid models combining different machine learning techniques that could be beneficial for improving recognition accuracy across different symbol types like CNN-SVM and CNN-KNN.

Proposed Work

The methodology includes major steps including pre-processing, segmentation, two-layer classification using CNN and K-NN and Symbol prediction as described in Fig. 1.

Data Collection and Pre-Processing

A Java-based GUI has been developed to input online handwritten physics symbols from different categories like electric flux, Maxwell's equations, inductance and pointing vectors. Online handwriting is considered a set of strokes, which are identified using pen up and pen down in Fig. 2.
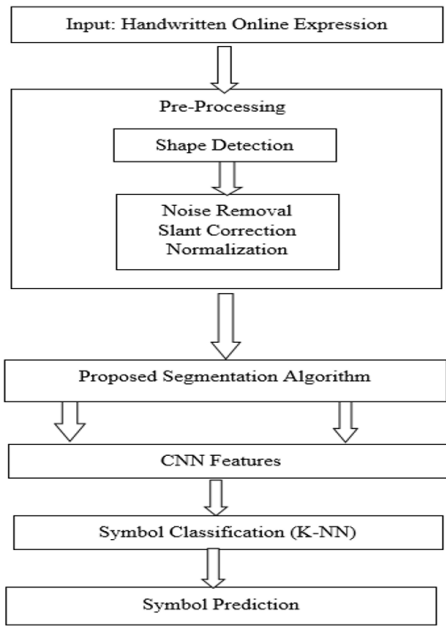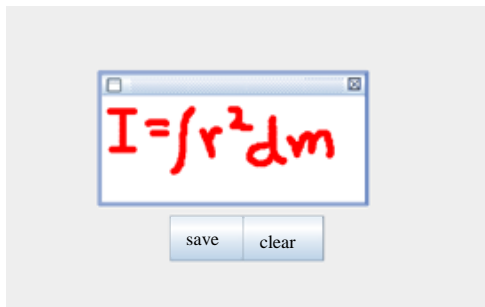

Fig. 1: Steps in symbol recognition

Fig. 2: Java-based UI to input physics expressions

Table 1: Expression types

Expression name	Online handwritten expression
Electric flux	$\phi_E = \int \vec{E} \cdot d\vec{A}$
Maxwell's equations	$\lambda_m = \frac{2L}{m}$
Inductance	$L = \frac{\mu_0 N^2 A}{\ell}$
The pointing vector	$\vec{S} = \frac{1}{\mu_0}(\vec{E} \times \vec{B})$
Moment of Interior	$I = \int r^2 dm$

Table 2: Data collection from five categories of OHPE

Class number	Symbol	Frequency	Total symbols
1	Phi	1	20
2	E	2	40
3	Equal	4	80
4	Integration	1	20
5	Arrow	5	100
6	.	1	20
7	d	1	20
8	A	1	20
9	lamda	1	20
10	m	2	40
11	2	2	40
12	L	2	40
13	Division	2	40
14	I	1	20
15	mu	2	40
16	0	2	40
17	small L	1	20
18	N	1	20
19	r	1	20
20	X	1	20
21	(	1	20
22	)	1	20
23	B	1	20
24	s	1	20
25	1	1	20
Total			780

The process of collecting data involves distributing a Java-based application to 20 writers; each one wrote a single expression related to each of four types: electric flux, Maxwell's equations, inductance and the pointing vector. In total, there were 80 expressions comprising 25 different symbol types. The dataset contains 780 occurrences of various symbols. Following Table 1 shows the types of expressions considered for the experiment.

The data collection includes five categories of physics expressions comprising 25 classes with its frequency in the overall five categories of Online Handwritten Physics Expressions (OHPE) has been shown in Table 2.

Pre-processing of expression consists of smoothening, sharpening, skew correction and resizing. To smoothen the image median filter is used to decrease the intensity disparity between pixels which helps to remove the noise from the image. The median filter is the most common method to remove the image noise.

To sharpen the image, a high pass filter maximizes the brightness of the center pixel relative to neighbor pixels. It decreases low-frequency information within an image and preserves high-frequency information. Skew detection and correction have been performed using the skew detection and correction method. The skew angle including the slant is corrected by rotation and then by shear transformation in the horizontal direction.

Proposed Extended Contour-Based Bounding Box Segmentation (CBBS) Algorithm

In the existing literature, online handwriting recognition is obtained by identifying the structural and special representation of symbols used within handwritten formulas, most of them recognized by sequence-to-sequence methods (Coquenet *et al.*, 2023; Johannes Michael *et al.*, 2019; Lai *et al.*, 2017; Zhu and Wang, 2012). In context with more simplicity towards the segmentation process for the connected symbols, CBBS shows better performance. This proposed algorithm treats physics symbols using a classical approach, which is similar to symbols having character-like properties. This classical approach is more suitable as compared to the template matching and recognition of whole physics expression. The proposed algorithm addresses the segmentation issues related to broken symbols, for example, the symbol '=' is segmented as two times '-', which is ambiguous to the minus symbol. The contour-based approach. The existing literature proposes the following methods for symbol segmentations:

1. Connected Component Analysis (CCA)
2. Deep Learning Based Approaches
3. Graph-based approaches

The conventional technique of Connected Component Analysis (CCA) offers efficiency, while deep learning approaches provide cutting-edge performance and adaptability to diverse datasets. Graph-based methods, on the other hand, integrate spatial and contextual information, enhancing segmentation accuracy, especially for intricate symbol structures. It is important to evaluate the hybrid approaches, as well as the integration of domain-specific knowledge and advanced optimized segmentation techniques to improve the accuracy of the mathematical symbol segmentation process (Long *et al.*, 2015). The proposed CBBS segmentation algorithm is a hybrid model, which segments the symbols prior to CNN, which then, provides a more refined classification accuracy.

The proposed algorithm takes input for handwritten physics expression using Java UI and converts it into greyscale to simplify the image and reduce dimensionality. Otsu's thresholding method is used to convert an image into a binary image, which is used to optimize the threshold value. The kernel function is defined to perform morphological operations including dilation and erosion. This algorithm also embeds the features of identifying the contours of the image, which determines the boundaries of the connected components of the symbols. The algorithm adjusts the contours of the connected symbols by counting the connected components and thereafter extension of this uses a bounding box algorithm to segment the symbols within physics expressions.

Proposed Segmentation Algorithm: CBBS

Input: Online Mathematical Expression (AME) image taken by JAVA UI.
Output: Segments of Physics Expression
  Begin: Read the image
    Convert the image to a grayscale
    Convert gray image into binary using Oststu's threshold method
    Define Kernel function
    Apply morphological operations (Dilation and erosion)
    cnts ← findContours(dilated-img, retrieval mode, approximation method)
    IF length (cnts) == 2
    Cnts= cnts [0]
    ELSE
        Cnts= cnts[0]
Initialize min-val to max-value
Initialize num to 0
For c in cnts
    Area ← contourArea(c)
    IF area > min-val:
        x, y, width, height ← boundingRect(c)
        Draw Rectangle (img, (x, y), (x+width, y+height), color, thickness)
        Extract ROI (Region of Interest)
        Increment num by 1
      Display an image with contours around it.
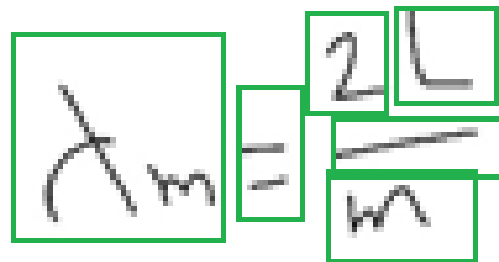 End

Fig. 3: Output of proposed CBBS algorithm

In this algorithm, morphological operations like dilation and erosion are applied in order to connect the dis-connected symbols. It overcomes the problem of segmentation of broken symbols. The erroneous situation that occurred during the segmentation of broken symbols is shown in Fig. 3.

CNN-K-NN Based Classification

The segmented images of the physics expressions are provided to CNN-KNN architecture for recognition purposes. CNN has proven efficient performance for spatial correlation of images. CNN is designed to extract the relevant features from the images.

Table 3: Types of symbols

Symbol	Class	Symbol	Class	Symbol	Class
φ	1	ʌ	9	ʟ	17
E	2	m	10	N	18
⁼	3 19	2	11	ɣ	19
∫	4	ʟ	12	×	20
→	5	—	13	(	21
·	6	⊥	14	)	22
d	7	⅏	15	B	23
A	8	O	16	S	24
ǀ	25				

The convolution layers in CNN learn a hierarchical representation of the input character or expression and each layer extracts more complex features than the previous layer. Handwritten characters have variations in strokes, width, size, orientation and shape, which it makes challenging to build a model to recognize the handwritten characters. CNN comprises local receptive, shared weight and pooling.

The proposed model consists of six stages: Preprocessed segmented symbols as an input, convolution layer 1, 2 and 3, generation of the 3D feature map, converting 3D feature map into 1D feature map, KNN classifier with dropout method and the prediction of symbols. The dataset consists of segmented symbols, which are segmented using the proposed extended contour-based bounding box method. There are 24 unique symbols identified and labeled with 24 classes. The dataset comprises 780 symbols, which are split into training and testing datasets.

Table 3. shows the class assignment for each of the 24 uniquely extracted symbols from the four categories of physics expression.

Results and Discussion

A convolution layer consists of fully connected network layers, in the proposed system a kernel of 32 filters of size 3×3, using batch normalization is used to extract the most distinguishable features from the raw input image. The output of each layer becomes the input of the layer. In convolution layer 2 of 64 filters of size 3×3 using batch normalization and convolution layer 3 of 128 filters of size 3×3 using batch normalization produces a 3D feature map. The 3D feature map was further converted into a 1D feature map and provided with a K-NN classifier with 180 features extracted using the dropout method. This

CNN-KNN coupling is to improve the overall performance of the prediction process, using the dropout technique. With the addition of dropout to the CNN-based KNN model, only fully connected hidden layers provide remarkable results. To implement the CNN-KNN model dropout layer is added after each CNN layer with the parameter called dropout rate, which has been adjusted between 0.1 and 0.5. To reduce overfitting dropout techniques are widely used (Lai *et al.*, 2017). The overall network is trained using layers and filter size with dropout shown in Table 4.

The overall network parameters used to build the network are shown in Table 5.

To avoid the gradient vanishing problem, the *ReLU* activation function is used, which causes preventing the training of the network during gradient-based learning or backpropagation (Jeon and Yang, 2021; Sanida *et al.*, (2022):

$$ReLu(x) = 0 \; for \; x < 0, x \; for \; x \geq 0 \tag{1}$$

The batch normalization works with mean zero and standard deviation, it is used to normalize data and bring it into a common scale. The mean of the hidden activation from the input layer *h* using Eq. (2):

$$\mu = \frac{1}{n} \sum hi \tag{2}$$

CNN generates the feature map using the filter, in which the max pooling method extracts the maximum element from the feature map regions.

Gamma is a parameter that affects the KNN algorithm's performance by controlling the distance metric used to measure the proximity between two points in the feature space. This feature has been adjusted to 2 to evaluate the performance under dropout shown in Table 6.

Table 4: Network layers and filter size with dropout

Layer	Size	Filter size	Dropout rate
Input layer	40×40×3	-	-
Conv1 3×3	40×40×3	32	0.25
Conv2 3×3	40×40	64	0.25
Conv3 3×3	40×40	128	0.50

Table 5: Network size with dropout

Layers	Activation shape	Learnable parameters
Conv1	40×40×32	896
batchnorm1	40×40×32	32
relu1	40×40×32	0
maxpool1	40×40×32	0
Conv2	40×40×64	73,856
batchnorm2	40×40×64	128
relu2	40×40×64	0
maxpool2	40×40×64	0
Conv3	40×40×128	295,168
batchnorm3	40×40×128	256
relu3	40×40×128	0
maxpool3	40×40×128	0
Total		370336

Table 6: Dropout parameters

CNN based-KNN model with dropout

Learning rate	0.001	Kernel parameter γ	0.1
Batch size	30.000	Penalty parameter C	20.0

Table 7: Classification accuracy with fold

Fold	CNN accuracy %	CNN-KNN accuracy %
1-Fold	93.30	97.80
2-Fold	94.40	98.10
3-Fold	89.30	98.30
4-Fold	87.20	97.90
5-Fold	93.20	98.20
Average	91.48	98.06

Table 8: Performance evaluation for CNN

		CNN-model		
Class no.	Symbol	Precision	Recall	F1-score
1	Phi	1.00	0.90	0.95
2	E	1.00	0.95	0.97
3	Equal	1.00	0.81	0.90
4	Integration	1.00	0.90	0.95
5	arrow	1.00	0.80	0.89
6	Dot	1.00	1.00	1.00
7	d	1.00	0.90	0.95
8	A	1.00	1.00	1.00
9	Lambda	1.00	0.95	0.97
10	m	1.00	0.95	0.97
11	2	1.00	0.90	0.95
12	L	1.00	0.93	0.96
13	Division	1.00	0.63	0.77
14	I	1.00	0.90	0.95
15	mu	1.00	0.83	0.90
16	0	1.00	0.88	0.93
17	Small L	1.00	0.95	0.97
18	N	1.00	0.85	0.92
19	r	1.00	0.85	0.92
20	X	1.00	0.90	0.95
21	(	1.00	0.80	0.89
22	)	1.00	0.80	0.89
23	B	1.00	1.00	1.00
24	s	1.00	0.90	0.95
25	1	1.00	0.70	0.82
Average		1.00	0.88	0.93

The CNN helps to obtain features in terms of feature maps from each layer of CNN. Using CNN two levels of features were obtained, these are low-level features and high-level features. The low-level features are extracted from the first convolution layer having weight 32. The obtained feature maps are the local features, which concentrate on special features of the characters. The convolution takes into the deeper level i.e., at level 3,

which helps to obtain the high-level features. These features are given as input to the K-NN classifier. To train the CNN model an Adam classifier is used with a learning rate of 0.001. A fivefold cross-validation technique is used to represent the accuracy of the prediction. The dataset is divided into five equal parts or folds; this data is given for training and testing purposes. The five-fold cross-validation technique is used to prevent overfitting (Sejuti and Islam, 2023). Table 7 represents the accuracy reported using cross validation k-fold technique. The dataset is divided into k-folds and the model is tested and trained k-times. Which helps for performance evaluation of the model and generalization ability. The average accuracy for CNN is 91.48 and while for CNN-KNN it is 98.06, which suggest that CNN-KNN model accuracy is better than CNN.

This proposed model is designed to predict online handwritten physics symbols. A self-learning CNN connected with KNN is used to predict the accuracy of the physics symbols. The k-fold cross-validation technique is applied to generalize the symbols database. The high-level CNN features are further given as an input to the KNN algorithm with a considerable dropout rate. In this case, KNN is used to classify the 3rd convolution layer high-level features of each symbol. The average accuracy is obtained at 98.06%. This proposed method is an efficient method for recognizing online handwritten physics symbols. The CNN-KNN model reports better accuracy than only using the CNN model. Tables 8-9 show the comparison of both models with respect to fold 5.

Table 9: Performance evaluation for CNN

		CNN-KNN model		
Class no.	Symbol	Precision	Recall	F1-score
1	Phi	1.00	0.95	0.97
2	E	1.00	0.95	0.97
3	Equal	1.00	0.98	0.99
4	Integration	1.00	0.90	0.95
5	Arrow	1.00	0.99	0.99
6	dot	1.00	1.00	1.00
7	d	1.00	0.90	0.95
8	A	1.00	0.95	0.97
9	Lamda	1.00	0.95	0.97
10	m	1.00	0.98	0.99
11	2	1.00	0.98	0.99
12	L	1.00	0.98	0.99
13	Division	1.00	0.98	0.99
14	I	1.00	0.95	0.97
15	mu	1.00	0.98	0.99
16	0	1.00	0.98	0.99
17	Small L	1.00	0.95	0.97
18	N	1.00	0.95	0.97
19	r	1.00	0.95	0.97
20	X	1.00	0.95	0.97
21	(	1.00	0.90	0.95
22	)	1.00	1.00	1.00
23	B	1.00	1.00	1.00
24	s	1.00	1.00	1.00
25	1	1.00	1.00	1.00
Average		1.00	0.96	0.98

The comparison table assesses the performance of the CNN and CNN-KNN models in symbol recognition. Both models demonstrate a precision of 1.00 (100%) for most classes, indicating rare misclassification of symbol instances. However, the symbol "division" shows lower precision, with 0.77 in the CNN model and 0.99 in the CNN-KNN model, suggesting occasional misclassification. Recall values vary across classes. Symbols like "dot", "d", "A", "B" and "s" achieve perfect recall (1.00) in both models, indicating successful identification of all instances. Notably, the CNN-KNN model significantly outperforms the CNN-Model in the recall of the "Division" symbol, with values of 0.98 and 0.63, respectively. Symbols "Phi" and "arrow" exhibit slightly lower recall values in both models, indicating these symbols are more challenging to identify accurately. high F1 scores across most classes signify a balanced trade-off between precision and recall. The CNN-KNN model generally achieves higher F1 scores, suggesting a better balance between precision and recall compared to the CNN model. On average, both models demonstrate high precision (1.00) for symbol recognition. However, the CNN-KNN Model shows slightly superior recall (0.96) and F1-score (0.98) compared to the CNN Model, indicating better overall performance across all classes. In conclusion, both models exhibit robust performance in symbol recognition, with minor variations in precision, recall and F1-score across different classes. The CNN-KNN model demonstrates slightly superior average performance, particularly in recall and F1-score, compared to the CNN model.

Conclusion

In this study, we propose three approaches for physics symbol recognition: A Java user interface for user input, a contour-based bounding box segmentation algorithm to handle broken symbols and a CNN-KNN recognition model that combines CNN feature extraction with K-NN classification. By integrating these approaches, we achieve state-of-the-art results in symbol recognition. Experimental results, obtained from handwritten physics symbols collected from various writers, demonstrate the effectiveness of the proposed hybrid CNN-KNN model, achieving an impressive accuracy of 98.06%. This underscores the potential of our approach to significantly enhance symbol recognition in physics expressions, offering valuable insights for e-learning applications and beyond. In further research, more categories of physics expressions can be experimented with hybrid segmentation and classification techniques.

Author's Contributions

Ujwala Kolte: Contributed to the literature reviewed, methodology, experimentation and result analysis.

Sachin Naik: Contributed to the conceptualization, supervision and conclusion.

Vidya Kumbhar: Contributed to result analysis and evaluation.

Ethics

I undersigned that this article has not been published elsewhere. The authors declare no conflict of interest.

References

Artieres, T., Marukatat, S., & Gallinari, P. (2007). Online Handwritten Shape Recognition Using Segmental Hidden Markov Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(2), 205-217. https://doi.org/10.1109/tpami.2007.38

Baró, A., Riba, P., Calvo-Zaragoza, J., & Fornés, A. (2019). From Optical Music Recognition to Handwritten Music Recognition: A baseline. *Pattern Recognition Letters*, *123*, 1-8. https://doi.org/10.1016/j.patrec.2019.02.029

Bresler, M., Prua, D., & Hlavác, V. (2013). Modeling Flowchart Structure Recognition as a Max-Sum Problem. *2013 12th International Conference on Document Analysis and Recognition*, 1215-1219. https://doi.org/10.1109/icdar.2013.246

Coquenet, D., Chatelain, C., & Paquet, T. (2023). End-to-End Handwritten Paragraph Text Recognition Using a Vertical Attention Network. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(1), 508-524. https://doi.org/10.1109/tpami.2022.3144899

Huang, B., Du, C., Zhang, Y., & Kechadi, M. T. (2006). A Hybrid HMM-SVM Method for Online Handwriting Symbol Recognition. *Sixth International Conference on Intelligent Systems Design and Applications*, 887-891. https://doi.org/10.1109/isda.2006.61

Jeon, H. K., & Yang, C. S. (2021). Enhancement of Ship Type Classification from a Combination of CNN and KNN. *Electronics*, *10*(10), 1169. https://doi.org/10.3390/electronics10101169

Michael, J., Labahn, R., Grüning, T., & Zöllner, J. (2019). Evaluating Sequence-to-Sequence Models for Handwritten Text Recognition. *2019 International Conference on Document Analysis and Recognition (ICDAR)*, 1286-1293. https://doi.org/10.1109/icdar.2019.00208

Keshari, B., & Watt, S. (2007). Hybrid Mathematical Symbol Recognition Using Support Vector Machines. *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, 859-863. https://doi.org/10.1109/icdar.2007.4377037

Kherdekar, V. A., Naik, S. A., & Bafna, P. (2023). A Segmentation-based Token Identification for Recognition of Audio Mathematical Expression. *International Journal of Advanced Computer Science and Applications*, *14*(7), 298-303. https://doi.org/10.14569/ijacsa.2023.0140733

Lai, S.-H., Lepetit, V., Nishino, K., & Sato, Y. (Eds.). (2017). Computer Vision-ACCV 2016 (Vol. *10112*). *Springer Cham*. https://doi.org/10.1007/978-3-319-54184-6

Li, Z., Wang, X., Liu, Y., Jin, L., Huang, Y., & Ding, K. (2024). Improving Handwritten Mathematical Expression Recognition via Similar Symbol Distinguishing. *IEEE Transactions on Multimedia*, *26*, 90-102. https://doi.org/10.1109/tmm.2023.3260648

Le, A. D., Nguyen, H. D., Indurkhya, B., & Nakagawa, M. (2019). Stroke Order Normalization for Improving Recognition of Online Handwritten Mathematical Expressions. *International Journal on Document Analysis and Recognition (IJDAR)*, *22*(1), 29-39. https://doi.org/10.1007/s10032-019-00315-2

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3431-3440. https://doi.org/10.1109/cvpr.2015.7298965

Miyao, H., & Maruyama, M. (2007). An Online Handwritten Music Symbol Recognition System. *International Journal of Document Analysis and Recognition (IJDAR)*, *9*(1), 49-58. https://doi.org/10.1007/s10032-006-0026-9

Nguyen, C. T., Truong, T. N., Nguyen, H. T., & Nakagawa, M. (2021). Global Context for Improving Recognition of Online Handwritten Mathematical Expressions. In J. Lladós, D. Lopresti, & S. Uchida (Eds.), *Document Analysis and Recognition-ICDAR 2021* (Vol. *12822*, pp. 617-631). Springer, Cham. https://doi.org/10.1007/978-3-030-86331-9_40

Ptucha, R., Petroski Such, F., Pillai, S., Brockler, F., Singh, V., & Hutkowski, P. (2019). Intelligent Character Recognition Using Fully Convolutional Neural Networks. *Pattern Recognition*, *88*, 604-613. https://doi.org/10.1016/j.patcog.2018.12.017

Putri, D. U. K., Pratomo, D. N., & Azhari, A. (2023). Hybrid Convolutional Neural Networks-Support Vector Machine Classifier with Dropout for Javanese Character Recognition. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, *21*(2), 346-353. https://doi.org/10.12928/telkomnika.v21i2.24266

Qiu, Q., Tan, Y., Ma, K., Tian, M., Xie, Z., & Tao, L. (2023). Geological Symbol Recognition on Geological Map Using Convolutional Recurrent Neural Network with Augmented Data. *Ore Geology Reviews*, *153*, 105262. https://doi.org/10.1016/j.oregeorev.2022.105262

Sanida, T., Sideris, A., Tsiktsiris, D., & Dasygenis, M. (2022). Lightweight Neural Network for COVID-19 Detection from Chest X-ray Images Implemented on an Embedded System. *Technologies*, *10*(2), 37. https://doi.org/10.3390/technologies10020037

Sejuti, Z. A., & Islam, M. S. (2023). A Hybrid CNN–KNN Approach for Identification of COVID-19 with 5-Fold Cross Validation. *Sensors International*, *4*, 100229. https://doi.org/10.1016/j.sintl.2023.100229

Shuvo, S. N., Hasan, F., Ahmed, M. U., Hossain, S. A., & Abujar, S. (2021). MathNET: Using CNN Bangla Handwritten Digit, Mathematical Symbols and Trigonometric Function Recognition. In S. Borah, R. Pradhan, N. Dey, & P. Gupta (Eds.), *Soft Computing Techniques and Applications* (Vol. *1248*, pp. 515-523). Springer, Singapore. https://doi.org/10.1007/978-981-15-7394-1_47

Smith, W. A. P., & Pillatt, T. (2023). You Only Look for a Symbol Once: An Object Detector for Symbols and Regions in Documents. In G. A. Fink, R. Jain, K. Kise, & R. Zanibbi (Eds.), *Document Analysis and Recognition - ICDAR 2023* (Vol. *14191*, pp. 227-243). Springer, Cham. https://doi.org/10.1007/978-3-031-41734-4_14

Sundaram, S., & Ramakrishnan, A. G. (2014). Performance enhancement of online handwritten Tamil symbol recognition with reevaluation techniques. *Pattern Analysis and Applications*, *17*(3), 587-609. https://doi.org/10.1007/s10044-013-0353-7

Tang, P., Hui, S. C., & Fu, C. W. (2013). Online chemical symbol recognition for handwritten chemical expression recognition. *2013 IEEE ACIS 12th International Conference on Computer and Information Science (ICIS)*, 535-540. https://doi.org/10.1109/icis.2013.6607894

Urala, K. B., Ramakrishnan, A. G., & Mohamed, S. (2014). Recognition of open vocabulary, online handwritten pages in Tamil script. *2014 International Conference on Signal Processing and Communications* (*SPCOM*), 1-6. https://doi.org/10.1109/spcom.2014.6984002

Xin, G., Cuiyun, L., Jihong, P., & Weixin, X. (2003). HMM based online hand-drawn graphic symbol recognition. *6th International Conference on Signal Processing, 2002*, 1067-1070. https://doi.org/10.1109/icosp.2002.1179973

Zhu, R., & Wang, Y. (2012). Application of Improved Median Filter on Image Processing. *Journal of Computers*, *7*(4), 838-841. https://doi.org/10.4304/jcp.7.4.838-841