International Journal of Engineering in Computer Science

E-ISSN: 2663-3590
P-ISSN: 2663-3582
www.computersciencejournals.com/ijecs
IJECS 2024; 6(1): 204-215
Received: 17-08-2024
Accepted: 22-09-2024

Supreetha Patel Tiptur Parashivamurthy
Research Scholar, Department of Electronics and Communication Engineering, Kalpataru Institute of Technology, Tiptur, Visvesvaraya Technological University, Belagavi, Karnataka, India

Dr. Sannangi Viswaradhya Rajashekararadhya
Professor and Former Principal, Department of Electronics and Communication Engineering, Kalpataru Institute of Technology, Tiptur, Visvesvaraya Technological University, Belagavi, Karnataka, India

Corresponding Author:
Supreetha Patel Tiptur Parashivamurthy
Research Scholar, Department of Electronics and Communication Engineering, Kalpataru Institute of Technology, Tiptur, Visvesvaraya Technological University, Belagavi, Karnataka, India

A comprehensive literature review on handwritten character recognition of Kannada scripts

Supreetha Patel Tiptur Parashivamurthy and Sannangi Viswaradhya Rajashekararadhya

DOI: https://doi.org/10.33545/26633582.2024.v6.i2c.142

Abstract
One among the southern states of India is Karnataka. The official language of Karnataka is Kannada. In the image processing (IP) and also the pattern recognition's field, the Kannada handwritten character recognition (KHCR) is an active and difficult research field. When contrasted to the other languages, KHCR is complex. For understanding the insight into the existing research's development in the Character Recognition (CR) field, a comprehensive review of the KHCR is proffered by this paper that will be useful for the new researchers. The different IP methods utilized in Kannada handwritten CR and challenges from literature analysis are covered in this survey. It also addresses and reviews interpretations and probable suggestions.

Keywords: Kannada handwritten character recognition, feature extraction techniques, preprocessing and segmentation of CR, classification techniques, deep learning, machine learning

1. Introduction

The growth of technologies that make the tools more human is the most significant effect possessed by the cognitive science field in the computer science field. Handwritten recognition technology is an extremely pertinent current field of natural interface research [1]. An infinite amount of styles is possessed by handwritten characters (HC) from one person to another. It is hard to be identified by a machine because of this broad range of variability [2]. The machine-printed digits along with characters are initially recognized by the Optical Characters Recognition (OCR). After that, it is developed for the identification of characters together with digits handwritten [3].

Asamiya, Bodo, Bangla, Dogri, Hindi, Gujarati, Kannada, Konkani, Kashmiri, Maithili, Malayalam, Marathi, Manipuri, Nepali, Punjabi, Odia, Sanskrit, Sindhi, Santali, Tamil, Telugu, along with Urdu are the 22 official languages in India [4]. Only some works could be traced for handwritten CR of Indian scripts although adequate studies are conducted in foreign scripts namely Chinese, Japanese along with Arabic characters [5]. Owing to several probable variations namely, number, direction, and also the shape of the constituent strokes, the handwritten Kannada characters are more complicated for identification when contrasted to English characters. In Chinese, English, along with Urdu, numerous works are conducted [6].

Occasionally, characters might not be identified properly by computer systems on account of these difficulties [7]. Therefore, in the CR field, IP methods are utilized. The common phases for IP are preprocessing, segmentation, feature extraction (FE) along with classification. FE is a vital component of CR in IP methods. Identifying the different features of digitized character images that increased the recognition accuracy (RA) is the FE phase's major aim [8]. Template matching, Unitary Image transforms, Projection Histograms, Graph description, Deformable templates, Contour profiles, Zoning geometry moments invariants, Zernike moments, Fourier descriptors, Gradient feature, Spline curve approximation, along with Gabor feature are the FE techniques [9, 10]. In the KHCR field, numerous remarkable works are performed in the preceding years on Kannada base characters. By offering a complete literature review of handwritten CR for Kannada scripts, this paper tries to elaborate on these sorts of studies [71].

Addressed multi-orientation scene text detection using an effective OBD method. Using discretization methodology, OBD, can solve the inconsistent labelling issue by discretizing the point-wise prediction into order less key edges. To decode accurate vertex positions, they have proposed a simple but effective MTL method to reconstruct the quadrilateral bounding box. Benefiting from OBD, they improved the reliability of the confidence of the bounding box and adopted more effective post-processing methods to improve performance. Results verified that their method can significantly outperform recent state-of-the-art methods.

[72] an efficient technique is proposed, which recognizes the characters in real-time by first applying a smoothing technique to remove the trembling noise, followed by a feature extracting process to extract geometric invariant features from the characters and then apply a multiclass Naïve Bayesian Classifier to recognize handwritten characters. It has been observed through experiments that the proposed approach achieves recognition accuracy of 97.5% in real-time [73].

Presents a sword-like model based on a convolutional neural network with a sword structure to recognize font styles for Chinese characters. It includes 15 convolutional layers. For each layer, they gradually increased the number of convolutional kernels to better extract the classification features of the input image. Four down sampling layers are used in the model. For each down sampling operation, the length and width of the image become half of their original values while the number of channels gradually increases, leading to a sword-like shape. Hence, they named their model as SwordNet. They compared the accuracy of their model with six other state-of-the-art network models. The experiments showed that SwordNet could achieve an average recognition accuracy of 99.03% in multiple experiments, while the other six models can only achieve accuracy up to 94.91%.

[74] proposes a collaborative trainable mechanism that integrates a global image feature extraction-based super-resolution neural network with a character recognition neural network. This collaborative trainable mechanism helps the character recognizer to be robust to inputs with varying quality in the real world. The alternative collaborative learning and character recognition performance test was conducted using the license plate image dataset among various character images, and the effectiveness of the proposed algorithm was verified using a performance test.

[75] focused on the noise in the character images that could decrease the accuracy of handwritten character recognition. Many types of noise penalties influence the recognition performance, for example, low resolution, Gaussian noise, low contrast, and blur. First, they proposed a method that learns from the noisy handwritten character images and synthesizes clean character images using the robust deblur generative adversarial network (DeblurGAN). Second, they combined the DeblurGAN architecture with a convolutional neural network (CNN), called DeblurGAN-CNN. Subsequently, two state-of-the-art CNN architectures are combined with DeblurGAN, namely DeblurGAN-DenseNet121 and DeblurGAN-MobileNetV2, to address many noise problems and enhance the recognition performance of the handwritten character images. Finally, the DeblurGAN-CNN could transform the noisy characters to the new clean characters and recognize clean characters simultaneously. For the n-THI-C68 dataset, the DeblurGAN-CNN achieved above 98% and outperformed the other existing methods. For the n-MNIST, the proposed DeblurGAN-CNN achieved an accuracy of 97.59% when the AWGN+Contrast noise method was applied to the handwritten digits. They have evaluated the DeblurGAN-CNN on the THCC-67 dataset. The result showed that the proposed DeblurGAN-CNN achieved an accuracy of 80.68%, which is significantly higher than the existing method, approximately 10%.

[76] Presented a classification framework for automatic recognition of handwritten Urdu character and digits with higher recognition accuracy by utilizing theory of transfer learning and pre-trained Convolution Neural Networks (CNN). The performance of transfer learning is evaluated in different ways: by using pre-trained AlexNet CNN model with Support Vector Machine (SVM) classifier, and fine-tuned AlexNet for extracting features and classification. They have fine-tuned AlexNet hyper-parameters to achieve higher accuracy and data augmentation is performed to avoid over-fitting. Experimental results and the quantitative comparisons demonstrate the effectiveness of the proposed research for recognition of handwritten characters and digits using fine-tuned AlexNet. This proposed research is based on fine-tuned AlexNet outperforms the related state-of-the-art research thereby achieving a classification accuracy of 97.08%, 98.21%, 94.92% for urdu characters, digits and hybrid datasets respectively.

[77] proposed ShuiNet-A, a lightweight artificial neural network model based on the attention mechanism, which combines channel and spatial dimensions to extract key features and finally recognize Shui manuscript characters. The effectiveness and stability of ShuiNet-A were verified through multiple sets of experiments. Their results showed that, on the Shui manuscript dataset with 113 categories, the accuracy of ShuiNet-A was 99.8%, which is 1.5% higher than those of similar studies. Experiments verified that the obtained classification accuracy was higher than those of similar studies.

[78] presented an end-to-end DL-based Tamil handwritten document recognition (ETEDL-THDR) model. Methods: Segmentation is used, first at the word level and then at the line level. ETEDL-THDR text recognition can be accomplished using two modules: line segmentation and line recognition. Initially, the ETEDL-THDR model targets improving input image quality using the median filtering (MF) technique. To create meaningful regions, more line and character segmentation activities are performed. A deep convolutional neural network (DCNN) based MobileNet approach is also applied to derive feature vectors. Finally, the water strider optimization (WSO) algorithm with a bidirectional gated recurrent unit (BiGRU) model is used to identify the Tamil characters.

Figure 1 depicts the general architecture for handwritten recognition.

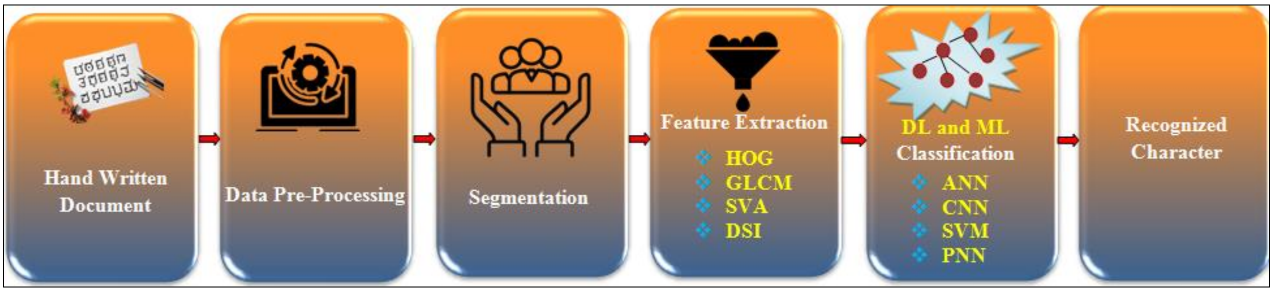


Fig 1: General architecture for handwritten recognition

2. Literature Review

Regarding KHCR, numerous research papers have been written recently. Different Handwritten CR methods are reviewed in this paper. Section 2.1 deemed the Preprocessing along with Segmentation methods. Section 2.2 reviewed the FE methods for CR. Section 2.3 addressed Classification methods of Handwritten Recognition.

2.1 Preprocessing and Segmentation

The first step for the image recognition system is preprocessing as it enhances the image's reliability. For obtaining an appropriate image for the succeeding steps, the preprocessing method is utilized in a CR system. It alleviates or intensifies particular image details for quicker assessment. The objects present inside an image are identified by using the *segmentation process*. The different preprocessing and segmentation techniques together with disadvantages are examined here.

Mamatha H R and Srikantamurthy K [11] presented morphological operations along with projection profiles for segmenting the Kannada scripts (handwritten) into lines, words, along with characters. On account of the complexity implied within the script, the technique was checked on completely unrestrained handwritten Kannada characters, which was more challenging. The text line's extraction was made effective by the morphology's utilization via a 94.5% average extraction rate. They could obtain an average segmentation rate of 82.35% and also 73.08% for words along with characters, correspondingly as a consequence of the changing inter and intra-word gaps. But, a single technique for segmentation was utilized by the method. For segmenting every sort of image, a single method couldn't be implemented efficiently.

Mayur M Patil and Akkamahadevi R Hanni [12] calculated the skew angle together with the image was rotated by the skew angle in the reverse direction. For enhancing the preprocessing and segmentation phases, noise removal methods were applied. For obtaining excellent performance and also the outcome's accuracy, effective FE and also classification techniques were utilized. Satisfactory results are attained for the algorithms utilized in the system. However, for real-time applications, this method was not appropriate.

Mamatha Hosalli Ramappa and Srikantamurthy Krishnamurthy [13] created a bounding box method for skew detection and also correction, together with the Kannada document's segmentation (handwritten). Hough transform and also contour detection was applied, correspondingly for segmenting the document into lines along with words. As a result of the complexity incorporated in the script, the technique was examined on fully unrestrained handwritten Kannada characters, which was more difficult. A 91% line segmentation rate was attained. For words, they could attain a 70% average segmentation rate on account of the changing inter and intra-word gaps. Nevertheless, the system had less convergence.

2.2 Feature Extraction for Kannada Handwritten Documents

By forming new features from the existent ones, FE decreases the total features in a dataset. The different FE methods of Kannada handwritten identification and limitations were deemed in this section.

Niranjan S.K *et al.* [14] proffered a Fisher Linear Discriminate Analysis (FLD) for unrestrained handwritten Kannada CR. Features were extracted by the system as of well-known FLD, Two-dimensional FLD (2D-FLD), along with Diagonal FLD. Disparate distance measure methods were explored and contrasted with their results for classifying the characters. Concerning several character classes, the system was checked on unrestrained handwritten Kannada characters. The method's effectiveness and also feasibility was exhibited by the system. A better recognition rate (RR) was attained by the angle merged with 2D-FLD for the outcomes merged with modified characters. But, for huge datasets, the method offered was unstable.

Dr. Balaji Rajendra Bombade and Dr. D.D Doye [15] presented the Four Side Border Features (FSBF) method, which extracted the pre-processed image's features. For finding the image's unattended information, these features were utilized. Using the Segmented Binary Block (SBB) technique, this information was processed. A few characters in the Kannada script were more complex in their structure. Thus, for FE, the Segmented FSBF (SFSBF) method was needed. Utilizing the Support Vector Machine (SVM) with a 3-fold cross-validation method, every feature was identified. This experiment's produced results were contrasted with similar sorts of algorithms, which rendered that the CR's accuracy was 92%. However, when the dataset was complicated, the RR was unstable.

Karthik S and Srikanta Murthy K [16] presented a Histogram of Oriented Gradient (HoG) for CR in amalgamation with the SVM. Utilizing a horizontal projection profile along with windowing, the segmentation was performed. The result attained was transferred to the recognition module. A segmentation accuracy of 94% and also a RA of 95.02% were attained by this technique. The outcomes were weighted against the prevailing techniques which were promising. But, the touching character's segmentation in the document was not addressed by the method.

Zhuo Chen *et al.* [17] proffered a framework called Multilingual Text Recognition Networks (MuLTReNets) for handwriting detection. Feature extractor, handwriting recognizer, script identifier, along with auto-weightier were the system's 4 major modules. Spatial along with temporal knowledge was integrated by the feature extractor for encoding text images into features shared via the script identifier along with the recognizer. The script category as of a variable-length sequence was predicted by the script identifier comprising an auto-weighter aimed at balancing disparate scripts, whilst the long-short term memory (LSTM) along with Connectionist Temporal classification was adopted by the handwriting recognizer for accomplishing sequence decoding. The script detection's accuracy had attained 99.9%. However, more features were needed by this method for precise classification.

A. Sushma and Veena.G.S [18] examined a Hidden Markov model-centered FE for KHCR. The word's image will be segmented into letters, or every word in a line was segmented into frames after preprocessing method. For converting the image into frames that were in sequences, the procedure of this segmentation method was utilized. Binarization was incorporated by the preprocessing methods. This method had focused only upon simple words.

M.S. Patel and Sanjay Linga Reddy [19] discussed the grid-centered approach in Kannada offline handwritten word detection. Initially, the input word was split by the method into '4' grids. For every grid, the well-known subspace learning (SL) approach namely Principal Component Analysis (PCA) was implemented for better representation. The distance measure method i.e., Euclidean distance was calculated for categorization purposes. The standard PCA approach was outperformed by the grid-centered approach with the SL approach as revealed by the experiment result. However, more features were extracted by this approach, which degraded the recognition performance. The different FE methods for handwritten CR were deemed in table 1.

Table 1: FE Techniques

Author	Feature extraction Technique	Dataset	Accuracy achieved (%)	demerits
Leena R Ragha and M Sasikumar [20]	Gabor wavelets and also Statistical features	7650 images along with 225 sets of samples were created with 34 consonants features in every set.	75%	Less number of features was extracted by the methods that degraded the performance.
V. N. Manjunath Aradhya and C. Naveena [21]	Mathematical morphology	250 handwritten documents	85%	For challenges namely text overlapping and also the existence of characters, this method wasn't so robust.
Rajni Kumari Sah and Dr. K Indira [22]	Normalized coordinates, normalized trajectory, along with normalized deviation	OHKC dataset	98.38%	The accuracy level was decreased by the technique for a lower number of training data.
Parashuram Bannigidad and Chandrashekar Gudada [23]	Gray Level Co-occurrence Matrix (GLCM)	1200 image dataset from disparate dynasties, like, Vijayanagar, Hoysala, and Mysore Wodeyar was comprised by the dataset.	96.7%	A specific gray level pair was time-consuming.
M. Mahadeva Prasad *et al.* [24]	Principle Compound Analysis (PCA)	Own writing style, without exercising any limitations utilizing Tablet PC along with Genius digital note pads.	81%	The character-specific shape features were not deemed, which decreased the accuracy.
M. C. Padma and Saleem Pasha [25]	Quad tree-centered FE method	Data set of 1,200 samples	85.43%	The Kannada compound characters were not focused by this technique.
Saleem Pasha and M.C.Padma [26]	Discrete wavelet transform (DWT)	4800 images of handwritten Kannada characters	97.60%	It had poor directionality along with sensitivity.
L R Ragha and M Sasikumar [27]	Gabor Transformation	200 samples as of the corpus	80%	Higher time was taken for executing features since its dimension of the feature vector was very long.
Ganapatsingh and G. Rajput and Rajeswari Horakeri [28]	Crack codes along with Fourier descriptors	Dataset of handwritten Kannada character	91.02%	The proffered technique had experimented on vowels only, which was the drawback.
Sk Md Obaidullah *et al.* [29]	Structural and also visual appearance (SVA) along with directional stroke identification (DSI)	PHDIndic_11	98.5%	A tedious along with time-consuming task was data collection.

Shivanand Rumma *et al.* [30] presented the Radial Basis Function (RBF) for handwritten numeral identification of Kannada script. It was a feed-forward neural network that estimated the activation at the hidden neuron in a manner that was diverse as of the vector (input) along with the weight vector. For character categorization, an RBF classifier was utilized. Average accuracy of 91.87% was attained by this system. However, only utilizing local features, this method was executed.

2.3 Classification Techniques

For classifying the disparate HC, the classification technique is utilized. The different deep learning (DL) and also machine learning (ML) classifiers for identifying the characters together with disadvantages are deemed in this section.

Vishweshwrayya C *et al.* [31] formed a Deep Convolutional Neural Network (DCNN) for the handwritten Kannada Numerals detection process. In document fashion, the Kannada characters (handwritten) were enthralled and they were accustomed to pre-processing along with attribute extraction processes. Using exploited strategies namely Drift Length Count, Direction-associated progression code, DWT, along with Curvelet Transfiguration Wrapping, the features were extracted. A DCNN classifier was desired for an extraordinary classification process. 96% accuracy was attained by the Kannada numeral's isolation accuracy. However, low accuracy was proffered by the system for complex input.

Mallikarjun Hangarge *et al.* [32] presented Gabor wavelets centered on word retrieval as of Kannada documents. For capturing directional energy features of the word image, Gabor wavelets were utilized. For improving the document image's quality, a pre-processing step was executed. Simultaneously, utilizing 294 document images, word segmentation was attained. Next, for representing the candidate word and a query word, Gabor wavelets were utilized. Concerning average precision 81.25%, average recall 82.09%, along with F measure 84.53% at a threshold of 97%, the results attained were hopeful. Nevertheless, higher redundancy problems were possessed by this technique.

Keerthi Prasad G *et al.* [33] executed PCA for an online handwritten Kannada CR system. Utilizing '2' disparate approaches, like PCA and also Dynamic Time Wrapping (DTW), the system was applied on a mobile device. Numerous experiments were executed and complete analysis was executed on the acquired results for finding the appropriateness of these '2' approaches for handheld devices. When analogized to DTW, the results attained for the PCA approach were quite hopeful. For the PCA approach, RA up to 88% was obtained, and up to 64% was acquired for the DTW approach. In addition, 0.8 sec was the time taken for the unknown character's detection for the PCA approach, and approximately 55 sec for the DTW approach. Hence, for real-time applications, the PCA was appropriate. However, the new features were not interpreted easily.

V.N. Manjunath Aradhya *et al.* [34] examined a Fourier transform for IP method for KHCR. The suitable frequency band images were trained by the system succeeded by a PCA FE scheme. A Probabilistic Neural Network (PNN) was utilized for the following classification techniques. On a huge database, the system was checked comprising Kannada characters. It was also checked upon a standard COIL-20 object database. When contrasted to standard methods, the outcomes were better. However, more memory space was needed by the system for storing the model.

S.V. Rajashekararadhya and Vanaja Ranjan P [35] introduced a zone-centered hybridized FE system aimed at handwritten numeral detection. The character centroid was calculated and also the image (character/numeral) was separated into n equal zones. Likewise, the zone centroid was also assessed (two features). For every zone, grid, or box in the numeral image, this procedure had successively recurred. For the following classification along with recognition purposes, the nearest neighbor along with SVM classifiers was employed. Employing an SVM, 97.85%, 96.8%, 95.1%, along with 95% recognition rates were attained for Kannada, Telugu, Tamil, along with Malayalam numerals correspondingly. However, inadequate results were produced by this system.

Rakesh Rampalli and Angarai Ganesan Ramakrishnan [36] introduced a PCA aimed at FE system for handwritten CR. For the samples, the features for offline recognition along with a disambiguation procedure were utilized in the offline system where the classifier's confidence level was low. The individual systems were outperformed via the classifier, which was caused by the outputs that were merged probabilistically. For Kannada, experiments were conducted for a database of 295 classes. When the amalgamation with the offline system was employed, the online recognizer's accuracy was enhanced by 11%. A low RR was possessed by this technique. Although less RR was possessed by this technique.

Pawan Kumar Singh *et al.* [37] introduced a distance-Hough transform (DHT) for HC categorization. Utilizing a tree-centered approach, the categorization task was executed at the word level wherein the Matra-centered scripts were initially divided as of non-Matra scripts employing the DHT algorithm. Utilizing modified log-Gabor filter-centered features implemented at multi-scale along with multi-orientation, the Matra along with non-Matra-centered scripts were individually detected. Aimed at the handwritten Indic script's categorization, the present tree-centered approach's efficacy was established by the encouraging outcomes. The top overall RA of 94.23% was attained by this approach on 12 official Indic scripts. The features learned across the data sample's disparate positions were not shared by this technique.

Abhishek S. Rao *et al.* [38] presented a Convolution Neural Network (CNN). The features were efficiently extracted by the DL method as of the image (input) which assisted the model in the superior classification of KHCR. For assessing the error rate within the model, a categorical Cross-entropy loss function was utilized. During the model's testing, real-time handwritten data was employed. 86% accuracy was obtained. But, the classification along with prediction accuracy was low.

Asha K and Krishnappa H K [39] presented a CNN centered on CR system for documents written in Kannada language. For the application, the CNN was utilized. For training the model, the Chars74K dataset was employed. For documents (handwritten), CNN with suitable configuration and handful dataset had led to an adequate recognition rate. 98% accuracy was attained by the system aimed at the document comprising non-overlapping lines of characters. But, the system was computationally expensive.

Soumyadeep Kundu *et al.* [40] examined a Non-Fully-Connected Network (NFC-Net), centered on CNN architecture for Indic script detection as of handwritten word images. In consequence of its easy, efficient, and powerful architecture, NFC-Net was most helpful in a resource-limited environment. The convolution layers performed the FE in the system; therefore it offered exact, pertinent, and precise features. The RA of 96.30% was held by this system. However, simple datasets were utilized by the system for training.

S. Karthik and K. Srikanta Murthy [41] presented a deep belief network (DBN) for identifying the remote HC of Kannada. For the HC, a greater degree of RA was yet to be obtained. Distributed average of gradients (DAGs) was employed as the features and also RA of 97.04% was acquired. The best in class was the RA and it was very hopeful. However, backpropagation issues were possessed by the system.

Swapnil A. Vaidya and Balaji R. Bombade [42] introduced a Generalized Regression Neural Networks (GRNN) classifier-centered system towards the detection of offline

Devnagari along with Kannada HC. For obtaining the categorization process's precise outcomes in handwritten CR, the pre-processing methods utilized were normalization along with binarization. This method was centered on the positional properties of every pixel within the character image. GRNN was utilized for resulting feature vectors and also classification performance was attained in the CR task. 82.89% along with 85.62% accuracies on Devnagari and also Kannada character databases correspondingly were presented by the recognition scheme. But, for a complex task, this system was time-consuming. Table 2 deemed the classifiers utilized for KHCR.

Table 2: Classification Techniques

Author	Classification Technique	Dataset	Results	Advantage	Disadvantage
N. Venkateswara Rao *et al.* [43]	Multilayer Perceptron (MLP)	MNIST database	Accuracy- 95.6%	For enhancing the recognition outcomes, this FE and also classification technique was utilized.	Less accuracy for inaccurate data was possessed by this technique.
Gururaj Muharambi *et al.* [44]	SVM	2800 Kannada consonants	Accuracy- 83.02%	From Kannada consonants, high RA was attained.	When the data had more noise, it didn't execute very well.
Victor Carbune *et al.* [45]	LSTMs	IBM-UB-1 dataset	Accuracy- 86.12%	Relying upon the language, the system's RA was enhanced by 20–40% while utilizing smaller along with quicker models.	A huge quantity of memory bandwidth was needed by linear layers to be computed.
Roshan Fernandes and Anisha P Rodrigues [46]	CNN	Publicly accessible dataset	Accuracy- 87%	For any Artificial Intelligence associated research work on the Kannada language, this could well be a primary step.	The character's wrong recognition for noisy data was offered.
Deepika Gupta and Soumen Bag [47]	CNN	ImageNet datasets	Accuracy was 96.23%.	When CNN had offered such adequate results, higher precision attained in recognition compels us to think, whether multilingual still exist.	For training a network, several data was needed by CNN. However, a small dataset was utilized by this technique.
Parikshith H *et al.* [48]	CNN	ICDAR-2013	Accuracy was 93.06%.	They had to augment and fine-tune the accuracy attained on training the CNN for creating a good handwritten CR system.	For processing long sequences, it was tedious.
Saleem Pasha and M.C. Padma [49]	K–nearest neighbor	3600 handwritten samples	Accuracy was 87.33%.	Since the features were extracted as of the partitioned image as local features and the whole image as global features, the approach was easy and was efficient.	Only the handwritten Kannada characters were deemed for experimentation purposes.
G. Rajput and Rajeswari Horakeri [50]	Multi-level SVM	For every character, Kannada characters were comprised of 24500 images with 500 samples.	The RR was 91.02%.	The system was to eliminate the confusion amongst similarly shaped classes and thus augmented the RR.	When more noise was possessed by the dataset, it did not execute very well.
N. Shobha Rani *et al.* [51]	Transfer learning	Chars74K dataset	Accuracy was 73.51%.	Adequate accuracy with validation was exhibited by the usage of a huge quantity of training knowledge for the task of very large classes of 188.	Lower recognition rate.
Ramesh. G *et al.* [52]	Capsule Networks	Manually Kannada handwritten images	Accuracy - 98.7%	Both these methods were outperformed by this technique and it offered the advantages with the capacity of handling raw images.	Less accuracy for classification consonants.

N. Shobha Rani *et al.* [53] created DCNN aimed at handwritten Kannada CR. For experimentation, every base character merged with consonant or vowel modifiers in text (handwritten) with overlapping or touching diacritics was presumed as a separate class in the Kannada script. But, the individual classes were deemed as the compound characters that are non-overlapping or touching where the semantic analysis was conducted throughout the OCR's post-processing phase. It was perceived that the Alex net's performance in the printed character sample's categorization was stated as 91.3%. 92% accuracy was recorded concerning the handwritten text.

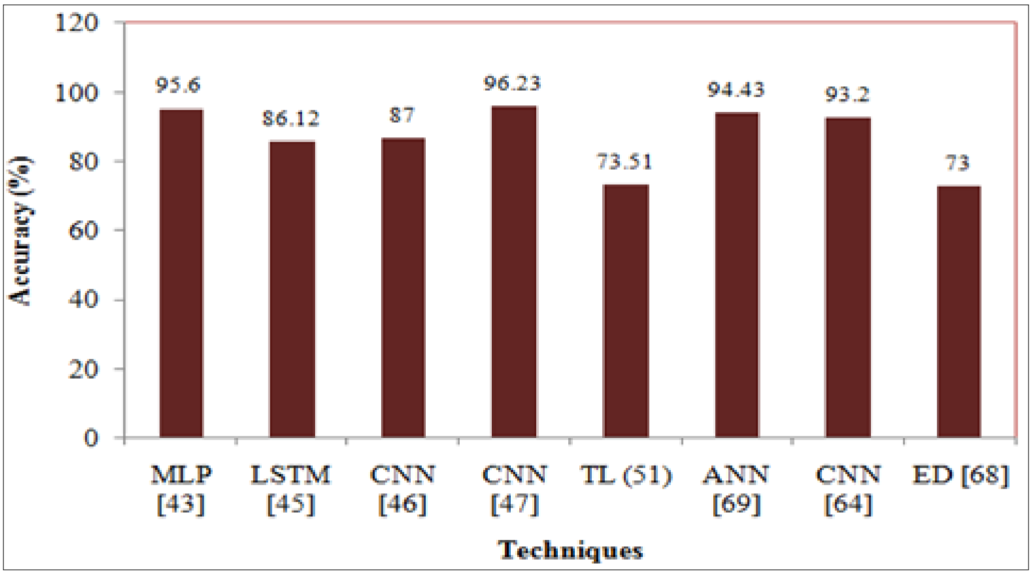
A K Sampath and N Gomathi [54] examined the fuzzy-centered multi-kernel spherical SVM offline handwritten CR. For obtaining appropriate images, the image (input) was initially offered to the pre-processing step. Next, for FE, a HOG descriptor was employed. Histogram estimation along with normalization calculation was constituted by the HOG descriptor. For CR, the features were categorized utilizing the classifier. For enhancing the performance considerably, a multi-kernel function was involved in the spherical SVM. 99% higher accuracy was attained by the system's outcome. However, for texture-centered images, this technique was not appropriate. Table 3 depicts the review of other FE and classification methods.

Table 3: Different FE and Classification Techniques

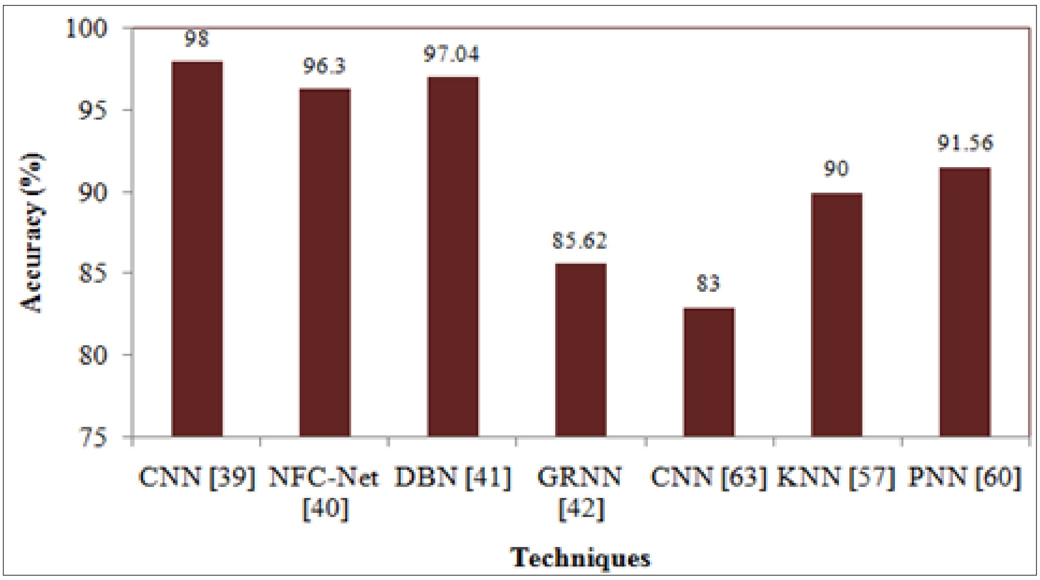
Author	Feature Extraction Technique	Classifiers	Disadvantage
S.A. Angadi *et al.* [55]	Structural feature vectors	SVM	This sort of recognition was hard to perform and also it was a very slow technique.
Aravinda C.V and Dr.H.N. Prakash [56]	Correlation method	Structural classifier	For unrestrained handwriting styles, template-centered matching had failed to produce significant results.
Vishwaas M *et al*. [57]	Direction centered Stroke Density principle(DSD)	Kohonen Neural Network(KNN)	The accuracy was decreased by mapping.
Devendra Pratap Yadav and Mayank Kumar [58]	HOG	Artificial neural network (ANN)	The network's duration was not deemed.
Rituraj Kunwar [59]	HOG	Statistical DTW (SDTW)	From the classifier, the accuracies stated are raw symbol recognition outcomes and also the usage of any language models or post-processing was not involved.
B.V. Dhandra *et al.* [60]	Directional density estimation	PNN	The complex input was hard to categorize.

3. Results and Discussion

In the below figures, the result's comparison obtained by the disparate DL, ML classification algorithms, and FE methods employed in the survey is done. Disparate sizes of datasets are employed by the studied algorithm for the prediction. They executed training along with testing on the input dataset. For showing their efficiency in detection, the algorithm's performance is contrasted and examined. Centered on classification accuracy, various ML along with DL techniques are examined.



(a)



(b)

Fig 2: CA of the DL techniques

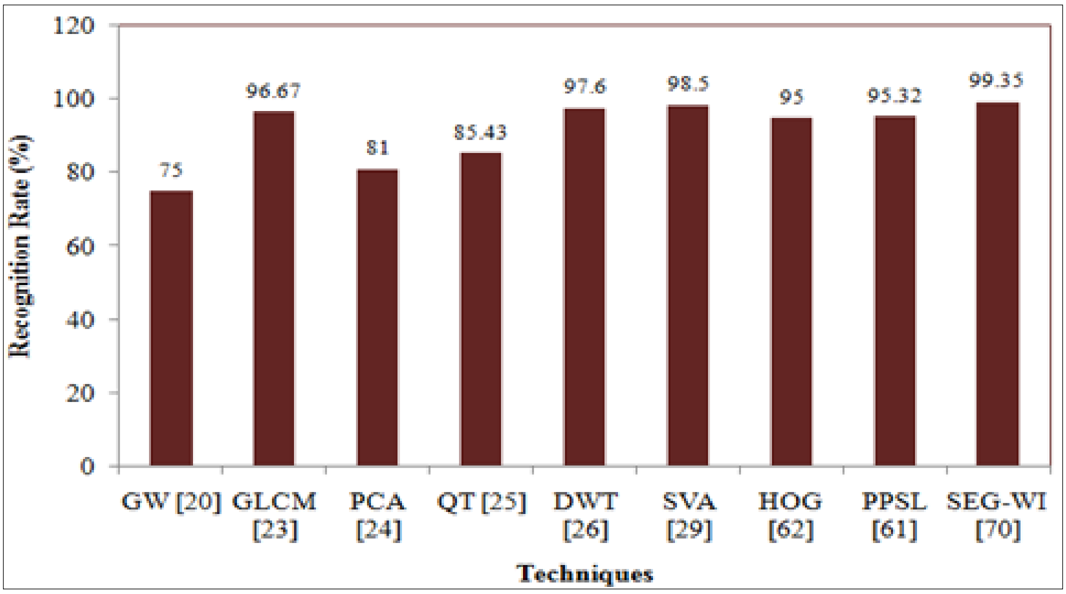


Fig 3: FE techniques

Utilizing DL, the Kannada HC classification accuracy (CA) is rendered in figure 2 (a). 95.6% accuracy is attained by MLP [43]. CA of 86.12% is possessed by LSTM [45], which is lesser when analogized to MLP. Afterward, 87% accuracy is acquired by CNN [46]. 96.23% of accuracy is attained by another CNN [47], which is greater when contrasted to every other method. 73.51% accuracy is accomplished by TL [51]. Then, ANN [69] attains 94.43% of accuracy. The Kannada handwritten document classification methods are depicted in Figure 3 exhibits the RR of FE methods. A 75% of RR is attained by GW [20]. Next, 96.70% of recognition is possessed by GLCM [23]. After that, DWT [26] obtains a 97.6% of recognition rate, which has higher recognition when analogized to other techniques. The ML's CA is figure 2 (b). 98% accuracy is acquired by CNN [39]. NFC-Net [40] obtains 96.3% and 85.62% of accuracy is attained by GRNN [42]. Next, 97.04% accuracy is possessed by DBN [41], which is greater when analogized to every other method. Next, 83% of accuracy is attained by CNN [63], which is lesser when contrasted with other methods. CA of 90% is possessed by KNN [57]. Lastly, PNN [60] obtains 91.56% of accuracy.

exhibited in figure 4. 83.02% of accuracy is attained by SVM [44]. Then, 87.33% of accuracy is acquired by KNN [49]. MSVM [50] acquires 91.02% of accuracy. LS-SVM [66] obtains 91% accuracy. Lastly, 80% accuracy is held by SVM [67], which is lesser when contrasted to other methods.

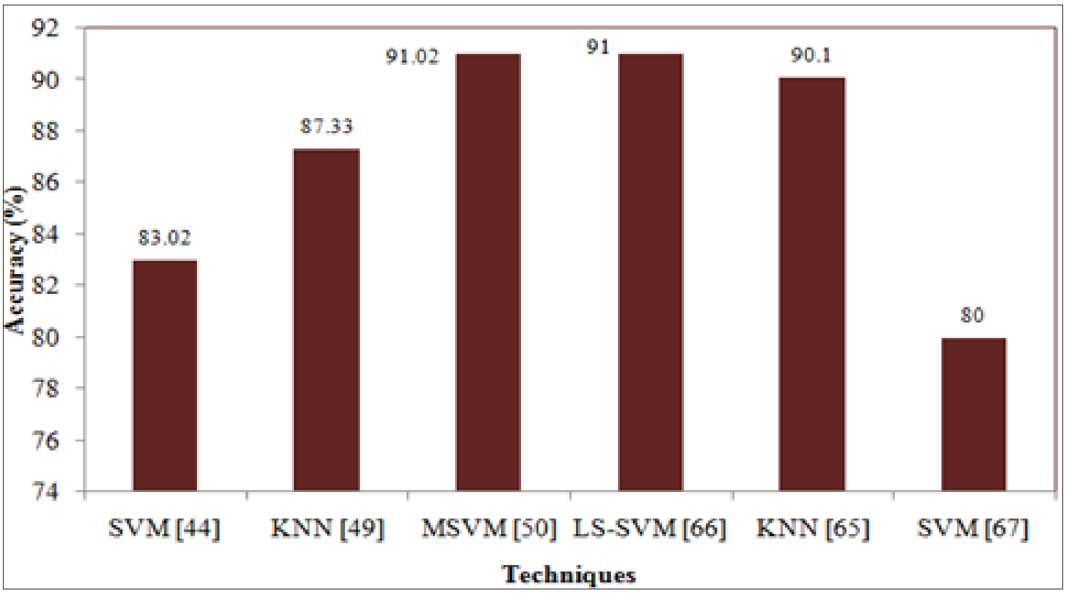


Fig 4: CA of the ML techniques

Different ML and also DL classifiers are reviewed in this paper. The maximum CA is offered by DL than ML methods as illustrated by comparison results as shown in Figure 4. However, the RR, dataset collections along with computational cost are still difficult tasks.

3. Conclusion

For the handwritten Kannada scripts identification, a complete overview of different FE along with classification methods is proffered in this paper. A detailed explanation of the associated benchmark along with standardized datasets is also offered. By contrasting numerous existent FE and also classification methods centered on their recognition rates, a systematic analysis is executed for deciding the recognition scheme's efficiency. Finally, disparate challenges in Kannada scripts are deemed, which offered a way ahead for upcoming research. The necessity for man along with machine interaction has raised this field from the viewpoint of upcoming work. For assisting society by improving the interface betwixt humans along with computers, the researchers have hugely helped the progression of CR systems that could well be applied practically. Thus, in the handwritten Kannada script identification field, this review is executed for compiling important contributions within the field that will promote and offer in-depth information to the researchers.

5. References

1. Sahu VL, Kubde B. Offline handwritten character recognition techniques using neural network: A review. International Journal of Science and Research. 2013;2(1):87–94.
2. Nasien D, Haron H, Yuhaniz SS. Support vector machine (SVM) for English handwritten character recognition. In: Second International Conference on Computer Engineering and Applications. IEEE; c2010 .p. 249–52. DOI:10.1109/ICCEA.2010.56.
3. Nooraliei A. Persian handwritten digits recognition by using zoning and histogram projection. In: 3rd Joint Conference of AI & Robotics and 5th RoboCup Iran Open International Symposium. IEEE; c2013. p. 1–5.
4. Soora NR, Deshpande PS. Review of feature extraction techniques for character recognition. IETE Journal of Research. 2018;64(2):280–95. DOI:10.1080/03772063.2017.1351323.
5. John J, Pramod KV, Kannan B. Handwritten character recognition of South Indian scripts: A review. Arxiv Preprint Arxiv. 2011:1106.0107.
6. Naveena C, Aradhya VM. An impact of ridgelet transform in handwritten recognition: A study on a very large dataset of Kannada script. In: World Congress on Information and Communication Technologies. IEEE; c2011. p. 618–21.
7. Hamad KA, Kaya M. A detailed analysis of optical character recognition technology. International Journal of Applied Mathematics, Electronics and Computers. 2016;Special Issue-1:244–9.
8. Kumar M, Sharma RK, Jindal MK. Efficient feature extraction techniques for offline handwritten Gurmukhi character recognition. National Academy Science Letters. 2014;37(4):381–91. DOI:10.1007/s40009-014-0253-4.
9. Dongre VJ, Mankar VH. A review of research on Devanagari character recognition. Arxiv Preprint Arxiv. 2011:1101.2491.
10. Zhong C, Ding Y, Fu J. Handwritten character recognition based on 13-point feature of skeleton and self-organizing competition network. In: International Conference on Intelligent Computation Technology and Automation. IEEE. 2010;2:414–7. DOI:10.1109/ICICTA.2010.837.
11. Mamatha HR, Srikantamurthy K. Morphological operations and projection profiles based segmentation of handwritten Kannada documents. International Journal of Applied Information Systems. 2012;4(5):13–9.
12. Patil MM, Hanni AR. Handwritten Kannada document image processing using optical character recognition. IOSR Journal of Computer Engineering. 2016;18(4):39–47. DOI:10.9790/0661-1804063947.
13. Ramappa MH, Krishnamurthy S. Skew detection, correction, and segmentation of handwritten Kannada documents. International Journal of Advanced Science and Technology. 2012;48:31.
14. Niranjan SK, Kumar V, Kumar H. FLD-based unconstrained handwritten Kannada character recognition. In: Second International Conference on Future Generation Communication and Networking Symposia. IEEE; c2008. vol. 3, p. 7–10.
15. Bombade BR. Handwritten Kannada character recognition with border features. Journal of Applied Science and Computations. 2018:1076–5131.
16. Karthik S, Murthy SK. Segmentation and recognition of handwritten Kannada text using relevance feedback and histogram of oriented gradients: A novel approach. International Journal of Advanced Computer Science and Applications. 2016;7(1).
17. Chen Z, Yin F, Zhang XY, Yang Q, Liu CL. MuLTReNets: Multilingual text recognition networks for simultaneous script identification and handwriting recognition. Pattern Recognition. 2020;108:107555. DOI:10.1016/j.patcog.2020.107555.
18. Sushma A, Veena GS. Kannada handwritten word conversion to electronic textual format using HMM model. In: International Conference on Computation System and Information Technology for Sustainable Solutions. IEEE; c2016. p. 330–5.
19. Patel MS, Reddy SL. An impact of grid-based approach in offline handwritten Kannada word recognition. In: International Conference on Contemporary Computing and Informatics. IEEE; c2014. p. 630–3.
20. Ragha LR, Sasikumar M. Feature analysis for handwritten Kannada Kagunita recognition. International Journal of Computer Theory and Engineering. 2014;3(1):94.
21. Aradhya VM, Naveena C. Text line segmentation of unconstrained handwritten Kannada script. In: Proceedings of the International Conference on Communication, Computing & Security; c2011. p. 231–4.
22. Sah RK, Indira K. Online Kannada character recognition using SVM classifier. In: IEEE International Conference on Computational Intelligence and Computing Research. IEEE; c2017. p. 1–6.
23. Bannigidad P, Gudada C. Identification and classification of historical Kannada handwritten document images using LBP features. International Journal of Intelligent Systems Design and Computing. 2018;2(2):176–88.
24. Prasad M, Sukumar M, Ramakrishnan AG. Divide and conquer technique in online handwritten Kannada character recognition. In: Proceedings of the International Workshop on Multilingual OCR; c2009. p. 1–7.
25. Padma MC, Pasha S. Quadtree-based feature extraction technique for recognizing handwritten Kannada characters. In: Emerging Research in Electronics, Computer Science and Technology. Springer, New Delhi; c2014. p. 725–35. DOI:10.1007/978-81-322-1157-0_74.

26. Fernandes R, Rodrigues AP. Kannada handwritten script recognition using machine learning techniques. In: IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics. IEEE; c2019. p. 1–6.
27. Ragha LR, Sasikumar M. Using moments features from Gabor directional images for Kannada handwriting character recognition. In: Proceedings of the International Conference and Workshop on Emerging Trends in Technology; c2010. p. 53–8.
28. Rajput GG, Horakeri R. Handwritten Kannada vowel character recognition using crack codes and Fourier descriptors. In: International Workshop on Multi-disciplinary Trends in Artificial Intelligence. Springer, Berlin, Heidelberg; c2011. p. 169–80.
29. Obaidullah SM, Halder C, Santosh KC, Das N, Roy K. PHDIndic_11: Page-level handwritten document image dataset of 11 official Indic scripts for script identification. Multimedia Tools and Applications. 2018;77(2):1643–78.
30. Rumma S, Vishweshwarayya CH, Bhuvaneshwari BD. Handwritten Kannada numeral recognition using radial basis function. International Journal of Computer Applications. 2014;98(8).
31. Vishweshwrayya C Hallur, Hegadi RS. Handwritten Kannada numerals recognition using deep learning convolution neural network (DCNN) classifier. CSI Transactions on ICT. 2020;8:295-309. DOI:10.1007/s40012-020-00273-9.
32. Mallikarjun C Hangarge, Veershetty, Rajmohan Pardeshi, Dhandra BV. Gabor wavelets based word retrieval from Kannada documents. Procedia Computer Science. 2016;79:441-448. DOI:10.1016/j.procs.2016.03.057.
33. Keerthi Prasad G, Imran Khan, Naveen R Chanukotimath, Firoz Khan. On-line handwritten character recognition system for Kannada using principal component analysis approach: For handheld devices. In: World Congress on Information and Communication Technologies. IEEE; c2012:675-678.
34. Manjunath VN Aradhya, Niranjan SK, Hemantha Kumar G. Probabilistic neural network based approach for handwritten character recognition. Special Issue of IJCCT. 2010;1:9-13.
35. Rajashekararadhya SV, Vanaja Ranjan P. Zone-based hybrid feature extraction algorithm for handwritten numeral recognition of two popular Indian scripts. In: World Congress on Nature & Biologically Inspired Computing (NaBIC). IEEE; c2009:526-530.
36. Rakesh Rampalli, Angarai Ganesan Ramakrishnan. Fusion of complementary online and online strategies for recognition of handwritten Kannada characters. Journal of Universal Computer Science. 2011;17(1):81-93. DOI:10.3217/jucs-017-01-0081.
37. Pawan Kumar Singh, Ram Sarkar, Vikrant Bhateja, Mita Nasipuri. A comprehensive handwritten Indic script recognition system: a tree-based approach. Journal of Ambient Intelligence and Humanized Computing; c2018 .p. 1-18. DOI:10.1007/s12652-018-1052-4.
38. Abhishek S Rao, Sandhya S, Anusha K, Chandana Nayak Arpitha, Sneha Nayak Meghana. Exploring deep learning techniques for Kannada handwritten character recognition: a boon for digitization; c2020.
39. Asha K, Krishnappa HK. Kannada handwritten document recognition using convolutional neural network. In: 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS). IEEE; c2018 .p. 299-301.
40. Soumyadeep Kundu, Sayantan Paul, Pawan Kumar Singh, Ram Sarkar, Mita Nasipuri. Understanding NFC-Net: A deep learning approach to word-level handwritten Indic script recognition. Neural Computing and Applications. 2020;32(12):7879-7895. DOI:10.1007/s00521-019-04235-4.
41. Karthik S, Srikanta Murthy K. Deep belief network based approach to recognize handwritten Kannada characters using distributed average of gradients. Cluster Computing. 2019;22(2):4673-4681. DOI:10.1007/s10586-018-2274-0.
42. Swapnil A Vaidya, Balaji R Bombade. A novel approach of handwritten character recognition using positional feature extraction. International Journal of Computer Science and Mobile Computing. 2013;2(6):179-186.
43. Ganapatsingh G Rajput, Rajeswari Horakeri. Handwritten Kannada vowel character recognition using crack codes and Fourier descriptors. In: International Workshop on Multi-disciplinary Trends in Artificial Intelligence. Springer, Berlin, Heidelberg; c2011 .p. 169-180.
44. Venkateswara Rao N, Srikrishna A, Raveendra Babu B, Rama Mohan Babu. An efficient feature extraction and classification of handwritten digits using neural networks. International Journal of Computer Science, Engineering and Applications. 2011;1(5):47. DOI:10.5121/ijcsea.2011.1505.
45. Gururaj Mukarambi, Dhandra BV, Mallikarjun Hangarge. A zone-based character recognition engine for Kannada and English scripts. Procedia Engineering. 2012;38:3292-3299. DOI:10.1016/j.proeng.2012.06.381.
46. Deepika Gupta, Soumen Bag. CNN-based multilingual handwritten numeral recognition: A fusion-free approach. Expert Systems with Applications. 2011;165:113784. DOI:10.1016/j.eswa.2020.113784.
47. Parikshith H, Naga Rajath SM, Shwetha, Sindhu CM, Ravi P. Handwritten character recognition of Kannada language using convolutional neural networks and transfer learning. IOP Conference Series: Materials Science and Engineering. 2021;1110(1):012003. DOI:10.1088/1757-899X/1110/1/012003.
48. Pasha S, Padma MC. Recognition of handwritten Kannada characters using hybrid features; c2013 .p. 59-65.
49. Pasha S, Padma MC. Handwritten Kannada character recognition using wavelet transform and structural features. In: International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT). IEEE; c2015. p. 346-351.
50. Rajput GG, Horakeri R. Unconstrained Kannada handwritten character recognition using multi-level SVM classifier. In: International Conference on Pattern Recognition and Machine Intelligence. Springer, Berlin, Heidelberg; c2013. p. 204-212.
51. Shobha Rani N, Subramani AC, Kumar A, Pushpa BR. Deep learning network architecture based Kannada

handwritten character recognition. In: Second International Conference on Inventive Research in Computing Applications (ICIRCA); c2020. p. 213-220. IEEE.
52. Ramesh G, Manoj Balaji J, Ganesh N. Sharma, Champa HN. Recognition of off-line Kannada handwritten characters by deep learning using capsule network. International Journal of Engineering and Advanced Technology. 2019;8(6):4767-4777. DOI:10.35940/ijeat.F8726.088619.
53. Shobha Rani N, Chandan N, Sajan Jain A, Kiran HR. Deformed character recognition using convolutional neural networks. International Journal of Engineering & Technology. 2018;7(3):1599-1604. DOI:10.14419/ijet.v7i3.14053.
54. Sampath AK, Gomathi N. Fuzzy-based multi-kernel spherical support vector machine for effective handwritten character recognition. Sādhanā. 2017;42(9):1513-1525. DOI:10.1007/s12046-017-0706-9.
55. Angadi S. A, Sharanabasavaraj H. Angadi. Structural features for recognition of handwritten Kannada character based on SVM. International Journal of Computer Science, Engineering and Information Technology. 2015;5(2):25-32. DOI:10.5121/ijcseit.2015.5203.
56. Aravinda CV, Prakash HN. Template matching method for Kannada handwritten recognition based on correlation analysis. In: International Conference on Contemporary Computing and Informatics (IC3I) IEEE; c2014. p. 857-861.
57. Vishwaas M, Arjun MM, Dinesh R. Handwritten Kannada character recognition based on Kohonen Neural Network. In: International Conference on Recent Advances in Computing and Software Systems. IEEE; c2012. p. 91-97.
58. Yadav DP, Kumar M. Kannada character recognition in images using histogram of oriented gradients and machine learning. In: Proceedings of 2nd International Conference on Computer Vision & Image Processing; c2018. p. 265-277. Springer, Singapore. DOI:10.1007/978-981-10-7898-9_22.
59. Kunwar R, Mohan P, Shashikiran K, Ramakrishnan A. G. Unrestricted Kannada online handwritten akshara recognition using SDTW. In: International Conference on Signal Processing and Communications (SPCOM); c2010. p. 1-5. IEEE. DOI:10.1109/SPCOM.2010.5560515.
60. Dhandra BV, Benne RG, Hangarge M. Kannada, Telugu and Devanagari handwritten numeral recognition with probabilistic neural network: a novel approach. International Journal of Computer Applications. 2010;26(9):83-88.
61. Alireza Alaei, Nagabhushan P, Pal U. A benchmark Kannada handwritten document dataset and its segmentation. In: International Conference on Document Analysis and Recognition; c2011. p. 141-145. IEEE. DOI:10.1109/ICDAR.2011.37.
62. Prasanna Kumar KS. Algorithm to identify Kannada vowels using minimum features extraction method; c2013.
63. Kevin George Joe, Meghna Savit, Chandrasekaran K. Offline character recognition on segmented handwritten Kannada characters. In: Global Conference for Advancement in Technology (GCAT); c2019. p. 1-5. IEEE.
64. Ramesh G, Ganesh N. Sharma, Manoj Balaji J, Champa H. N. Offline Kannada handwritten character recognition using convolutional neural networks. In: IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE); c2019. p. 1-5. IEEE.
65. Dhandra BV, Hangarge M, Mukarambi G. Spatial features for handwritten Kannada and English character recognition. International Journal of Computer Applications; c2010 .p. 146-150.
66. Ramakrishna Murty, Murthy MJVR, Prasad Reddy PVGD. Text document classification based on least square support vector machines with singular value decomposition. International Journal of Computer Applications. 2011;27(7):21-26.
67. Roopananda MK. Weighted zone based Kannada character recognition using crack feature extraction technique and support vector machine classification. International Journal of Science, Technology and Engineering (IJSTE). 2015;1:180.
68. Patel MS, Reddy, Sandyal KS. An improved handwritten word recognition rate of South Indian Kannada words using better feature extraction approach. In: Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA); c2015. p. 553-561. Springer, Cham. DOI:10.1007/978-3-319-12012-6_61.
69. Phangtriastu MR, Harefa J, Tanoto DF. Comparison between neural network and support vector machine in optical character recognition. Procedia Computer Science. 2017;116:351-357. DOI:10.1016/j.procs.2017.10.061.
70. Kumar P, Sharma A. Segmentation-free writer identification based on convolutional neural network. Computers & Electrical Engineering. 2020;85:106707. DOI:10.1016/j.compeleceng.2020.106707.
71. Liu Y, He T, Chen H, *et al*. Exploring the capacity of an orderless box discretization network for multi-orientation scene text detection. International Journal of Computer Vision. 2021;129(9):1972–1992. DOI:10.1007/s11263-021-01459-7.
72. Shabir M, Islam N, Jan Z, *et al*. Real-time Pashto handwritten character recognition using salient geometric and spectral features. IEEE Access. 2021;160238-160248. DOI:10.1109/ACCESS.2021.3123726.
73. Xudong Li, Wang J, Zhang H, *et al*. SwordNet: Chinese character font style recognition network. IEEE Access; c2022 .p. 8388-8398. DOI:10.1109/ACCESS.2022.3143795.
74. Lee S, Yun JS, Yoo SB. Alternative collaborative learning for character recognition in low-resolution images. IEEE Access. 2022;22003-22017. DOI:10.1109/ACCESS.2022.3153116.
75. Gonwirat S, Surinta O. DeblurGAN-CNN: Effective image denoising and recognition for noisy handwritten characters. IEEE Access. 2022;90133-90148. DOI:10.1109/ACCESS.2022.3201560.
76. Rasheed A, Ali N, Zafar B, *et al*. Handwritten Urdu characters and digits recognition using transfer learning and augmentation with AlexNet. IEEE Access. 2022;102629-102645.

DOI:10.1109/ACCESS.2022.3208959.
77. Minli T, Shaomin X, Xiangrong L. Ancient character recognition: A novel image dataset of Shui manuscript characters and classification model. Chinese Journal of Electronics. 2023;32(1):64-75. DOI:10.23919/cje.2022.00.077.
78. Vinotheni C, Lakshmana Pandian S. End-to-end deep-learning-based Tamil handwritten document recognition and classification model. IEEE Access. 2023;11:43195-43204. DOI:10.1109/ACCESS.2023.3270895.