# A Systematic Review of Multilingual Numeral Recognition Using Machine and Deep Learning Methodology

Meenal Jabde[1], Chandrashekhar Patil*,[1] Amol D. Vibhute[2], Jatinderkumar R. Saini[2]
[1]School of Computer Science, Dr. Vishwanath Karad MIT World Peace University, Pune, 411038, MH, India.
[2]Symbiosis Institute of Computer Studies and Research (SICSR), Symbiosis International (Deemed University), Pune-411016, MH, India.
meenal82@gmail.com (Meenal Jabde), chpatil.mca@gmail.com (Chandrashekhar Patil), amolvibhute2011@gmail.com (Amol D. Vibhute), saini_expert@yahoo.com (Jatinderkumar R. Saini)

**Abstract**

The COVID-19 epidemic has altered how people live and interact with one another and tangible objects. The offline optical character recognition (OCR) system has been affected in this regard. Due to its practical uses, the OCR has long been an intriguing subject for researchers. Numerous research reported on offline and online multilingual numeral recognition. The conventional way of air-writing with sensors is costly and time-consuming. This has led to research in air-writing without using sensors for multilingual numeral recognition. However, air-writing without sensors is an emerging topic that needs to be appropriately addressed efficiently and effectively. Air-writing without sensors has many advantages, such as being contactless, cost-effective, having real-time and fast processing, and ease of use. Air-writing without sensors can be effectively used in hospital operation theatres, top-secret agencies, online education, reservation counters, banks, airports, and post offices. Therefore, touchless technology is needed in OCR.

The paper thoroughly reviews the recent studies executed for offline and online multilingual numeral recognition. We pay particular attention to available datasets and machine and deep learning models used on various datasets for multilingual number identification. This review analyzed work done in datasets using various segmentation, feature extraction, and classification methods. It also focuses on several classification algorithms used and the accuracy obtained. Finally, the paper also elaborates on the applications and challenges of multilingual numeral recognition. This review will benefit numerous researchers working offline and online in multilingual numeral recognition and understanding the systematic approach.

**Keywords**: Air-writing, multilingual, offline numeral recognition, online numeral recognition, Air-writing without sensors, Real-time numeral recognition, Convolutional Neural Network.

## 1. Introduction

### 1.1 Optical Character Recognition (OCR) :

India is a multilingual, multiscript country with 22 official languages (Obaidullah et al., 2014). The Indo-Aryan and Dravidian language groups are India's most widely spoken and adopted languages. Each language category is used by over 97% of the Indian population (Hamad & Kaya, 2016). A document containing the text, such as handwritten or, printed or scanned text images, can be converted using OCR into an editable digital version. Thus, the OCR makes it possible for a machine to automatically detect text in these documents (Hamad & Kaya, 2016).

Nevertheless, the OCR system is used to identify the characters. Then, the outcomes are compared with the records in the database, for instance, to develop specific information (Kodgire & Sable, 2016). On the other hand, OCR categorizes optical structures in digital images that correspond to alphanumeric letters. The stages of OCR include data collection, preprocessing, segmentation, extraction of essential features, classifying, and post-processing (Kodgire & Sable, 2016).

Two categories under which handwritten character recognition systems may be categorized are online and offline handwriting recognition schemes. Using an online character recognition system, the user writes on an electronic surface with a specific pen, and the data is stored as (x, y) coordinates (Mehdi et al., 2013). The processed document containing the scanned images of handwritten text is a task for the offline character recognition system. Figure 1 shows the types of handwritten character recognition.

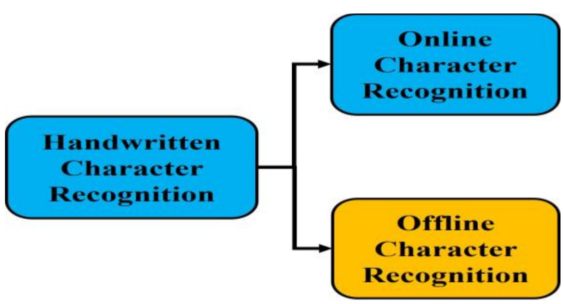


**Figure 1**. Handwritten Character Recognition System Classification.

Several researchers reviewed the literature on handwriting character recognition for text written on paper. However, there is a strong need to review the online-offline multilingual numeral recognition systems to identify and focus on the current state-of-art research, the datasets, methods, applications, and challenges with opportunities. Therefore, in the present paper, we proposed and conducted a systematic literature review on online multilingual numeral recognition systems for various languages. This review also shows the importance of online numeral recognition systems, presents the available data, finds out the limitations and challenges in earlier studies, focuses on an in-depth analysis of online multilingual numeral recognition systems and provides applications and future directions.

### 1.2 Motivation and Objectives

The diversity in culture and languages makes it difficult to recognize the regional numerals written through online and offline methods. Therefore, accurate numerical recognition systems are needed to analyze the data written in regional languages. Historical texts, papers, and documents have been written in different languages and need to be digitized. In this regard, accurate numeral recognition systems help digitize the historical information. In international businesses, several numerals are frequently used. Thus, an effective solution is also essential to recognize the different numerals automatically.

Moreover, the numeral recognition system is strongly recommended in translating the different languages into regional words. Furthermore, the numeral recognition systems are primarily required in the educational and literacy sectors to develop educational materials and methods accordingly. Therefore, we are motivated to conduct the review on online and offline numeral recognitions.

### 1.3 Major contributions of the article:

- The present review's significant contribution is exploring the current scenario for offline and online multilingual numeral recognition systems by examining several recent articles from reputable publications and notable conferences.
- The present article studies and critically analyses a variety of datasets used and available, as well as reported work and methodologies used on offline and online multilingual numeral recognition.
- A critical analysis of results was reported, and a system for multilingual numeral recognition was made available.
- Applications, challenges, and future perspectives have been identified and presented by reviewing the literature on offline and online multilingual numeral string recognition systems.

### 1.4 Outline of the Article:

The rest of this manuscript is prepared into eight sections. Section two focuses on the review methodology. Section three discusses the essentials of offline and online multilingual numeral recognition systems. This section reports and critically analyses all the work on offline and online multilingual numeral recognition. This section also discusses the available system for multilingual numeral recognition, variations, and applications in online-offline multilingual numeral recognition systems. The datasets used, and available datasets are critically analyzed and summarized in section four. Sections five and six are dedicated to machine and deep learning, respectively. A brief discussion on the challenges in online-offline multilingual numeral recognition is presented in section seven. The conclusions, investigation findings, and future scope are discussed in section eight..

## 2. The methodology of reviewing the literature:

The present review studies several manuscripts done on handwritten online-offline recognition systems. In reviewing the literature, the research articles were considered from reputed journals and conferences and books published by standard publishers such as IEEE, Springer, ACM, Elsevier, Taylor and Francis, Wiley-online, etc. Other sources like Scopus, Web of Science, ResearchGate, and pertinent websites were also used to select the relevant articles. Several studies have been done on handwritten recognition with some limits. In this regard, we followed the major papers by searching and screening them with essential elements. The identified papers on handwriting recognition with keywords were 1200. Then, we shortlisted 857 papers on title and abstract during the initial screening. Subsequently, the eligibility criteria were applied to include the relevant papers and finalize 489 papers for further inclusion. Finally, we applied the inclusion criteria to select only 64 quality papers from standard journals and conferences. Finally, the papers were selected from IEEE journals (21), Elsevier journals (8), Springer journals (15), and standard conferences (21) for review purposes.

## 3. Essentials of offline and online multilingual numeral recognition system

This section reviews offline and online multilingual numeral recognition systems written in the air with and without sensors. A comparative study has also been discussed based on the dataset, classification methods, and obtained accuracy reported in the literature.

### 3.1 Offline Multilingual Numeral Recognition

Offline multilingual numeral recognition includes processing documents containing scanned text images written by the user on paper. It automatically transforms text into a picture that may be used in text processing and computer programs.

Offline multilingual numeral recognition has been the subject of several studies. For example, the study (Bhattacharya & Chaudhuri, 2009) suggested a wavelet-based multi-resolution and multi-stage cascaded identification approach for recognizing English-Devanagari and English-Bangla mixed numbers. Authors (Basu et al., 2009) have also created a technique for locating postal codes in situations with several scripts. They developed 25 classes by merging Latin, Devanagari, Bangla, and Arabic numerals. The handwritten numeral is then recognized using a Support Vector Machine (SVM).

Authors (Akhand & Ahmed, 2016) developed a CNN-based approach for recognizing Bangla and Bangla-English combined handwritten digits in their research. This research used 20 classes to recognize numerals in Bangla-English mixed scripts. The experimentation was performed for 16 classes of multilingual numerals. The authors (Md, Shopon, Nabeel, 2016) employed a deep CNN to identify handwritten numerals in English and Bangla. They used a data augmentation strategy to increase recognition accuracy by creating a blocky effect on the training photos. They considered two languages for the experiment but only looked at one at a time (Md, Shopon; Nabeel, 2016).

Handwritten number recognition in four Indian scripts was achieved using CNN and pre-trained networks, according to the authors (Pramanik et al., 2019). The hybrid scripts of Bangla and Odia languages were selected separately to recognize the scripts. They also combined Bangla and Odia since both languages have the same structure, which is comparable to zero. In addition, the fusion of numbers with similar shapes has been done (Pramanik et al., 2019), and it has been handled as a 19-class issue. In contrast, the research (Dash et al., 2015) proposes a new handwritten numeral identification approach based on human perception. The digit's form is split into four sections in their work. The classification rules were then applied to extract the Odiya, Bangla, Arabic, and English texts. The authors (Ashiquzzaman et al., 2019) developed a novel CNN technique to read Arabic numerals via offline handwriting. The proposed model was trained with pre-processed data, and their accuracy (97.4%) was assessed using Arabic handwriting digits. The handwritten Arabic numerals have been classified in the research (Mahmoud & Abu-Amara, 2010) using the two-stage classification approach with Nearest Mean Classifier (NMC) and K-Nearest Neighbor (KNN) and obtained 98% accuracy for numeral recognition. Similarly, the researchers (Choudhary et al., 2011)applied the Artificial Neural Network (ANN) model with the single-layer and multi-layered perceptron in recognizing the offline handwritten numerals.

Table 1 focuses on the summary of offline multilingual numeral recognition with user data and applied methods for classification. It is seen from Table 1 that the CNN-based methods have mainly been used due to their high performance.

**Table 1**. Summary of offline multilingual numeral recognition.

| Data used | Methods used | Success rate | Refernces |
|---|---|---|---|
| CMATERDB Arabic handwritten digit dataset | MLP-CNN | 97.4% | (Ashiquzzaman et al., 2019) |
| Bangla, Devanagari, Latin, and Urdu language postal codes. | Support vector machine | Bangla- 97.15% Devanagari- 95.63%, Latin- 95.55%, Urdu- 96.20% | (Basu et al., 2009) |
| Postal mails, job application forms, handwritten papers with other circumstances | MLP | 92.14% | (Bhattacharya et al., 2006) |
| self-generated Arabic/Indian numerals | MLP | 99.1% | (Choudhary et al., 2011) |
| Odia and Bangla data collected by IITBBS, ISI Kolkata, ISI Kolkata, and CMATERdb Arabic, MNIST dataset | Perceptual shape decomposition | IITBBS Odia- 99.25%, ISI Kolkata Bangla -99.66%, ISI Kolkata Odia- 97.96% CMATERdb Arabic- 99.02%, MNIST English- 99.11% | (Dash et al., 2015) |
| ISI and CVPR Bengali numerals, MNIST database | CNN | 98.71% | (H. Akhand & Ahmed, 2016) |
| Self-generated Arabic/Indian numerals | NMC, KNN, and | HMM - 97.1%, | (Mahmoud & Abu- |

| | | | |
|---|---|---|---|
| | HMM | K-NNC - 98.33%, NMC - 98.66% | Amara, 2010) |
| ISI, MNIST, and CMATERDB3.1.1 datasets | CNN | CMATERdb - 99.83%, ISI: 99.35%, MNIST- 99.56%, | (Md, Shopon; Nabeel, 2016) |
| Bangla, Oriya, Telugu, and Devanagari datasets | CNN | 93.33% | (Pramanik et al., 2019) |

### 3.2 Online Multilingual Numeral Recognition

Online multilingual numeral recognition automatically converts text written on a unique digitizer or PDA. The present review uses several studies on recognizing online multilingual numerals. For instance, the authors (Azeem et al., 2012-a)suggested three phases for recognizing the Arabic digits using preprocessing, feature extraction, and classification. The study verified the applied methodology on a tiny dataset containing 100 samples from 100 users with a 98% recognition rate due to the small size of the dataset. Likewise, authors (Ramakrishnan, 2013) proposed an Arabic digit recognition system and obtained a 98.73% recognition rate using spatial and temporal information with a training set of 30,000 picture samples. Several feature extraction methods were used for numeral recognition with 95% recognition accuracy (Ghods & Sohrabi, 2016).

On the other hand, Persian characters were recognized using 18 letters based on a structure and HMM method. Their technique attained 94.2–95.9% accuracy for various groups (Ghods & Sohrabi, 2016). Similarly, the robust feature set and KNN classification were used in the Persian-Arabic-numeral recognition system proposed in the study (Salouan et al., 2014). In this case, the block-based methods with three significant features were used and achieved an accuracy of 99.90% (Salouan et al., 2014). The novel dataset of 45,000 samples was introduced by authors (Al-wzwazy et al., 2016) for Arabic numeral recognition. The CNN-enabled model was applied to the developed dataset and achieved 95.7% accuracy (Al-wzwazy et al., 2016). In research (Boukharouba & Bennia, 2015), the authors have developed the horizontal and vertical feature sets for Arabic and Persian numbers and applied the SVM method to recognize the numerals with sufficient results. The researchers have also applied deep learning methods to recognize the Arabic and Persian numbers (Boukharouba & Bennia, 2015). Similarly, several numerals such as English, Arabic, Persian, Urdu, and Devanagari were recognized using the CNN method (Latif, 2018) with 99.26%-99.29% accuracy.

Some studies also focus on fuzzy-based systems to recognize the numerals. For instance, the study (Musleh et al., 2021) suggested a fuzzy system for identifying Arabic numerals. In this regard, they created the distinct segments using fuzzy methods with directional properties and obtained 99.55% accuracy (Musleh et al., 2021). The research by (Hanmandlu et al., 2007) focused on the Hindi and English numerals using the fuzzy-based exponential membership functions (EMF) with 98.4% and 95% accuracies for the English ad Hindi numerals, respectively (Hanmandlu et al., 2007). The study (Jou & Lee, 2009) proposes the digit identification method using the condensed structural classification and fuzzy method to generate the primitive types of every image. The digit classification was done on the NIST dataset using the primitives, features, and fuzzy membership function to classify the digits, with 87.33%–88.72 % (Jou & Lee, 2009).

Conversely, the English and Devanagari characters have been recognized using fuzzy methods with modifying membership functions(Murthy, 2011). In addition, the modified membership function was also used in the study (Hanmandlu et al., 2007) to recognize Hindi handwritten digits. The authors (Majhi et al., 2014) recognized the Odia handwritten digits using HMM and a fuzzy inference with 96.3% accuracy. Similarly, a feature selection method based on an axiomatic fuzzy set was suggested to extract the handwritten numbers in Arabic, Telugu, Roman, Devanagari, and Bangla (Roy, 2014). However, the SVM method has been applied to classify the numeral features. The study (Takruri et al., 2015) uses hybrid classification techniques such as fuzzy C-means, neural networks, and SVM for handwritten Arabic numerals recognition. Table 2 highlights SVM, HMM, CNN, and fuzzy methods with more than 90% accuracy for online handwriting recognition.

**Table 2**. Summary of online multilingual numeral recognition.

| Datasets | Classification Method | Accuracy (%) | Reference |
|---|---|---|---|
| ICDAR 2009 | SVM | 98.73 | (Azeem et al., 2012a) |
| IWFHR 2006 | SVM | 95 | (Ramakrishnan, 2013) |
| TMU dataset | HMM | 94.2-95.9 | (Ghods & Sohrabi, 2016) |
| self-generated numeral dataset | CNN | 95.7 | (Al-wzwazy et al., 2016) |

| | | | |
|---|---|---|---|
| “Hoda” Farsi dataset | SVM | 100 | (Boukharouba & Bennia, 2015) |
| MADBase, MNIST, HODA, DHCD | CNN | 99.32 | (Latif, 2018) |
| AOD | SVM | 98.01 | (Musleh et al., 2021) |
| self-generated Hindi digit, CEDAR | fuzzy | 90.65-96 | (Hanmandlu et al., 2007) |
| handwritten numerals in NIST database | fuzzy | 87.33 - 88.72 | (Jou & Lee, 2009) |
| self-generated Hindi digit | fuzzy | 90.65 | (Murthy, 2011) |
| — | HMM | 96.3 | (Majhi et al., 2014) |
| CMATERdb, MNIST | SVM | 95.15 | (Roy, 2014) |
| CENPARNI Indian (Arabic) handwritten numerals dataset | fuzzy C-means | 88- 89 | (Takruri et al., 2015) |

### 3.3 The primary objectives of offline and online multilingual numeral recognition systems are:

1. To develop a constant, effective, and trustworthy model to recognize the regional language numerals,
2. To provide quick and accurate online-offline numeral recognition systems to support the input sources for several applications like virtual whiteboards, digital signatures, interactive forms, etc.,
3. To design and develop a flexible system that uses the transfer learning techniques to enhance the recognition performance with the limited datasets,
4. To develop a user-friendly interface that can be used in several applications,
5. To provide an open-source platform that supports developers, researchers, and educators to access and operate efficiently,
6. Enhance the accuracy of recognition and model performance to handle the latest issues.

Thus, the online-offline multilingual recognition system can be effectively used in the various sectors to help and guide the language translator, communication, information sharing, etc.

### 3.4 Variances in online and offline multilingual numeral recognition systems:

Five significant variations in the online and offline multilingual numeral recognition systems are demonstrated in Table 3.

**Table 3**. Differences in online and offline multilingual numeral recognition systems.

| **Resources** | **Offline Multilingual Numerical Recognition** (Bhattacharya et al., 2006)(Mahmoud & Abu-Amara, 2010)(Salouan et al., 2014)(Musleh et al., 2021)(Jou & Lee, 2009)(Ramakrishnan, 2013) | **Online Multilingual Numeral Recognition** (Mehdi et al., 2013) (Azeem et al., 2012a)(Ghods & Sohrabi, 2016) |
|---|---|---|
| Input source | The static pages or documents could be prepared in digital format, such as scanned docs, digital images, etc. | Dynamic inputs like the stylus, digital pens, and touchscreens are used. |
| The processing timings | The processing can only be done after digitization. However, there are no restrictions. | Instantly, when the user types or draws digits. |
| Use cases | The offline systems are used in OCR, document digitization, and non-changing materials applications. | These systems are used in interactive forms, virtual whiteboards, and digital signatures. |
| Challenges and difficulties | The change in font sizes, font types, orientations, and noises are challenging to recognize in the scanned documents. | The recognition of online sites is challenging due to the dynamic handwriting. |

### 3.5 Applications on online and offline multilingual numeral recognition systems:

The online-offline multilingual numeral recognition systems could be beneficial in several applications. For instance, offline multilingual numeral recognition systems are used in digitizing documents to record handwritten texts, optical character recognition to automate the data input, language conversion to translate into regional languages, analyzing scanned images, storing historical documents, extracting the numerals from the transactions, and developing the educational software, etc. Alternatively, online multilingual numeral recognition systems are used in digital signatures, interactive forms, virtual whiteboards, online banking, gesture recognition, and number or symbolic programs.

Thus, multilingual numeral recognition has numerous and significant applications in fields including communication, education, business, finance, technology, and culture. These applications may be found both offline and online. These programs help number recognition be more effective, accessible, and seamlessly integrated into all facets of our digital lives.

## 4. Datasets Available:

Offline and online handwriting multilingual numeral recognition systems require considerable data with a high success rate for training and testing purposes. The primary objective is to compile the data for the multilingual handwritten numerical recognition research. The primary datasets available (Hussain et al., 2015) for online-offline multilingual handwritten numerical recognition are demonstrated in the present section.

### 4.1 Datasets available for offline multilingual numeral recognition:

**4.1.1 CEDAR** (Hull, 1994)**:** A variety of handwritten datasets have been produced at the State University of New York at Buffalo's Centre of Excellence for Document Analysis and Recognition (CEDAR), including handwritten words, ZIP codes, numbers, and alphanumeric characters. These datasets were created primarily to support a study into the automated handling of postal addresses found on envelopes. The samples comprise 5632 city words, 4938 state keywords, and 9454 ZIP codes. As a result, address blocks and ZIP codes are separated into sets of 21,179 digits and 27,835 alphanumeric characters, respectively.

**4.1.2 NIST** (Wilkinson & Creecy, n.d.) : To help with tasks like character segmentation and identification, character isolation, and the detection and removal of boxes from forms, the National Institute of Standards and Technology, or NIST, established a variety of databases containing handwritten characters and numbers. There are 28 boxes on the form for numbers, two for alphabets, one for a paragraph of text, and boxes for the writer's information. Examples were produced for the NIST Special Database-1 by 2100 writers. The most recent edition of the database, Special Database 19, has 810,000 isolated character photos, 3600 writers' handwritten forms, and ground truth data.

**4.1.3 MNIST** (HAFFNER, 1998) **:** 60,000 training examples and 10,000 test samples comprise the extensive collection of handwritten numbers known as MNIST. The MNIST, a subset of the

NIST database, consists of samples from the NIST Special Database 3 (SD-3) and Special Database 1 (SD-1). The first recommendation was to use SD-3 as the training set and SD-1 as the test set. Census Bureau employees provided SD-3 samples, whereas high school students authored SD-1 examples. As a result, SD-1 posed more significant recognition-related challenges than SD-3. In order to obtain a uniform distribution of samples from SD-1 and SD-3 in training and test sets, the MNIST database was built using a training set of 30,000 pictures from SD-1 and 30,000 photographs from SD-3. Five thousand samples, each from the SD-1 and SD-3 databases, were included in the test set.

**4.1.4 Al-Isra** (Kharma, 1999)**:** The College of British Columbia's professors put together the Al-Isra Database, a significant collection of handwritten samples that contain words, numbers, signatures, and phrases. Five hundred students from the Jordanian university Al-Isra provided samples. Each student devised a preset list of phrases, words, and numbers. Five hundred unrestricted records are in the database: sentences, 37,000 words, and 10,000 numerals in Arabic with 2500 signatures. The database may be used. To identify writers and recognize handwriting tasks.

**4.1.5 Check DB** (Al-Ohali et al., 2003)**:** This database aims to promote research on the automated processing and identification of Arabic checks. The database comprises 7000 pictures of checks with over 15,000 numbers and around 30,000 subwords.

**4.1.6 AHDB** (Al-Ma'adeed et al., 2002)**:** The Arabic Handwriting Database (AHDB) is an offline database with several pre-processing stages. On checks written in Arabic by 100 different authors, phrases, paragraphs, and words are employed to represent numbers. Although testing various Arabic handwriting recognition algorithms was not the database's primary purpose, it does contain pages of unconstrained text that may be utilized in that capacity. The database might be utilized for things like handwriting recognition and writer identification.

**4.1.7 ARABASE** (Amara et al., 2005): A comprehensive database for online and offline handwriting recognition is ARABASE. The database also enables reading Arabic text printed by an offline machine. The database contains whole sentences, words, isolated numbers, letters, and signatures. The data source is a supporting tool for conventional documents, including jobs and database analysis. This database can assess (online or offline) handwriting systems for signature and recognition verification.

**4.1.8 Numerals DB** (Bhattacharya & Chaudhuri, 2005) : The authors provide a substantial collection of numbers written by hand in two popular Hindi scripts. The data was compiled using employment application forms written in Devanagari and Bangla, postal mail and other sources. It contains 368 postal mail and 274 job application forms, yielding 22,556 Devanagari numbers, 268 job applications, and 465 various types of correspondence, yielding 23,392 Bangla numerals. All images were converted into digital format at a resolution of 300 dpi and saved as grayscale "tif" images. Numerous digit recognition system technologies have been evaluated using the repository.

**4.1.9 CENPARMI-A** (Alamri et al., 2008): The CENPARMI Arabic database contains isolated digits, characters, numerical sequences, and sentences for offline Arabic handwriting recognition. A two-page form was produced to collect the data from 100 participants. The samples were of Arabic data, such as 38 numerical strings, 35 single numbers, and 20 digits for each of the 70 Arabic words and letters in these documents. A database was divided into three sets. The first set includes the initial forms for 100 authors, while the second batch includes the completed forms from 228 authors. The third group consists of several samples derived from sets 1 and 2.

**4.1.10 CENPARMI-U** (Sagheer et al., 2009): The CENPARMI Urdu handwritten database contains strings of numerals, phrases, and characters written in Urdu. Several native Urdu speakers from various regions worldwide participated in the data-gathering procedure. 57 Urdu words and 44 Urdu characters comprise the language, primarily of financial terminology, to facilitate recognition of words, characters, and numbers in offline Urdu.

**4.1.11 CENPARMI-F** (Haghighi et al., 2009): The Centre for Pattern Recognition and Machine Intelligence (CENPARMI) Farsi database and Farsi text word spots have been developed to support research on handwriting recognition. The archive has 432,357 images with dates, words, single characters, digits, sequences of numbers, and writings from 400 native Farsi speakers. Grayscale and binarized versions of each image are available. The repository has been used to assess symbol and digit recognition Along with Farsi handwriting recognition.

**4.1.12 Devanagari DB** (Dongre & Mankar, 2012): This presents a database of Devanagari numbers and characters. Writing examples from 750 people of all ages, educational levels, and occupations gathered. The database includes the sum of 20,305 isolated letters and 5137 isolated numbers. 'Tif' binary images are saved. A database has been created to make it publicly visible and usable for recognizing Devanagari symbols. The list of available offline multilingual datasets has been provided in Table 4.

**Table 4**: Offline multilingual numeral datasets

| **Database** | **Language / Script** | **Number of writers** | **Dataset Size** | **Reference** |
|---|---|---|---|---|
| CEDAR | English | Not specified | 21,179 digits | (Hull, 1994) |
| NIST | English | 3600 | 8,10,000 digits | (Wilkinson & Creecy, n.d.) |
| MNIST | English | Not specified | 70,000 digits | (HAFFNER, 1998) |
| Al-Isra | Arabic | 500 | 10,000 digits | (Kharma, 1999) |
| Check DB | Arabic | Not specified | 15,000 digits | (Al-Ohali et al., 2003) |
| AHDB | Arabic | 100 | 105 forms | (Al-Ma'adeed et al., 2002) |
| ARABASE | Arabic | 400 | 400 forms | (Amara et al., 2005) |
| Numerals DB | Bangla, Devanagari | Not specified | 45,948 numerals | (Bhattacharya & Chaudhuri, 2005) |
| CENPARMI-A | Arabic | 328 | 13,439 digits | (Alamri et al., 2008) |
| CENPARMI-U | Urdu | Not specified | 18,000 | (Sagheer et al., 2009) |
| CENPARMI-F | Farsi | 400 | 4,32,357 | (Haghighi et al., 2009) |
| Devanagari DB | Devanagari | 750 | 5,137 | (Dongre & Mankar, 2012) |
| BanglaLekha | Bangla | Not Specified | 1,66,105 | (Biswas et al., 2017) |

**4.2 Datasets available for online multilingual numeral recognition:**

**4.2.1 IRONOFF** (Viard-Gaudin et al., 1999) : The Institute de Recherche et d'Enseignement Supérieur aux Techniques de l'Electronique, or IRESTE, developed a dual on/off database known as IRONOFF (Viard-Gaudin et al., 1999). Examples of French writers' handwriting, including words, figures, and letters, are added to the database. The contributors were required to complete forms with the appropriate boxes, and human operators then examined both the completed forms and the actual data. Each donor filled out one of the three forms: B, C, or D. The information on form B consists of the lowercase and uppercase letters of the alphabet, numerals, the euro symbol, and strings frequently found in French checks. Forms C and D of the French language use cursive writing. Fifty thousand cursive words are stored in the database, 32,000 isolated characters, and 1000 total forms.

**4.2.2 LMCA** (Baati & Alimi, n.d.): The On/Off LMCA ("Lettres, Mots et Chiffres Arabe" in French) is a dual Arabic database comprising letters, words, and numbers. There are 55 samples total in the database. Participants produced 500 words and 30,000 numbers in total. The UNIPEN format, similar to the IRONOFF database, generates the database. Arabic words and numerals may be recognized using the database both online and offline.

**4.2.3 AOD** (Azeem et al., 2012b): The authors developed a significant number of Arabic online digits (AOD). The database was created using 300 authors, who ranged in age from different age groups, with about 75% belonging to the older group of 20 to 35 persons. This was done to ensure that various writing styles were covered. The youngest author is 11 years old, the oldest member is 70 years old, 60% of the authors are female, and more than 90% are right-handed. There were no numerical constraints on each number or the writing style utilized in oriental size or type, and each writer was expected to submit an average of 10 examples for each digit. There were 30,000 samples collected or 3000 samples for each digit.

**4.2.4 hpl-tamil-iwfhr06-train** (India, 2010): Over 300 isolated samples of each of the 156 Tamil "characters" (details) written by regional Tamil writers from Salem, Tamil Nadu, and Bangalore, Karnataka, India, are included in this database. Adults, university graduates, and students all contributed as authors. The data is in the usual UNIPEN format and was collected using HP TabletPCs. An offline version of the data was also used to produce bi-level TIFF images from the online data using simple piecewise linear interpolation and a constant thickening factor. The information can only be used for study.

**4.2.5 Online handwritten Assamese characters** (Baruah & Hazarika, 2015) : The online handwritten Assamese characters were added with 8,235 characters for 183 classes. The conjunct consonants, basic alphabetic letters, and Assamese numbers were present in this data.

**4.2.6 Ekush** (Rabby et al., 2019): This dataset's 367,018 isolated handwritten characters were written by 3086 individual writers from Bangladesh. In the Bangla language, these symbols comprise modifiers, vowels, consonants, complicated letters, and numerical digits. This dataset may be used for a range of tasks, such as gender, age, and district-based handwriting-related study, given that the samples gathered contain a diversity of the district, age group, and an equal number of male and female participants. The available datasets on online multilingual numerals are provided in Table 5.

**Table 5**. Online multilingual numeral datasets

| **Database** | **Language / Script** | **Number of writers** | **Dataset Size** | **Reference** |
|---|---|---|---|---|
| IRONOFF | French | Not specified | 50,000 words, 32,000 characters | (Viard-Gaudin et al., 1999) |
| LMCA | Arabic | 55 | 30,000 digits | (Baati & Alimi, n.d.) |
| AOD | Arabic | 300 | 30,000 | (Azeem et al., 2012b) |
| hpl-tamil-iwfhr06-train | Tamil | Not specified | 46,800 | (India, 2010) |
| online handwritten Assamese characters | Assamese | 183 | 8,235 | (Baruah & Hazarika, 2015) |
| Ekush | Bangla | 3086 | 3,67,018 | (Rabby et al., 2019) |

**5. Machine learning for multilingual numeral recognition:**

The main goals of machine learning-based methods are categorization and feature extraction of essential properties from the data. These kinds of studies (Ashikur Rahman et al., 2022) take digit photos, extract handmade characteristics from them, and then put them into a classifier to provide a prediction. A traditional machine learning-based recognition system is required to categorize the chores from images. Usually, the handwritten numbers require the fulfilment of the three stages. The initial action is to take important information out of the provided image, and the information characteristics are extracted elements. The feature space is then shrunk to a smaller dimensional size. This can reduce the amount of processing complexity needed or remove skewed or unnecessary characteristics and confound a model. The selected qualities are sent to a classifier for training within the last phase. Prior studies have shown that several approaches may be used to successfully use numerous methods for tackling each of these procedures. Table 6 shows the performances received by the earlier studies with the machine learning methods.

**Table 6.** Performances of the traditional machine learning-based approaches

| **Feature Extraction** | **Classifier** | **Dataset** | **Number of images** | **Accuracy (%)** | **Reference** |
|---|---|---|---|---|---|
| the features of chaincode, gradient, profile structure, and peripheral direction contributively | the learning vector quantization classifier, the discriminative learning quadratic discriminant function (DLQDF) classifier, the k-nearest neighbour classifier, three neural classifiers, and two support vector classifiers (SVCs) | CEDAR | 21,179 digits | 99.32 and 99.46 | (Liu et al., 2003) |
| primitive segments for each image | a tree-like classifier | NIST Special Database 19 | Not specified | 87.33 and 88.72 | (Jou & Lee, 2009) |
| Not Specified | Multilayer Perceptron, Support Vector Machine, Naïve Bayes, Bayes Net, Random Forest, J48 and Random Tree | Austrian Research Institute for Artificial Intelligence, Austria | 3689 | MLP – 90.37 | (Shamim et al., 2018) |
| MOD Features | SVM | Own Dataset | 6000 | 95.02 | (Singh et al., 2017) |
| Structural features | Probabilistic Neural network (PNN) classifier | Kannada, Telugu, Devanagari | 2550 | 97.20 | (Dhandra et al., 2011) |

| | | | | | |
|---|---|---|---|---|---|
| Shape similarity | Fuzzy Rule | Urdu and Arabic – Own Dataset | 900 | 96.30 | (Razzak et al., 2009) |

## 6. Deep learning techniques for multilingual numeral recognition:

Conventional deep learning techniques (Ashikur Rahman et al., 2022) heavily relied on extensive feature engineering, which placed various restrictions on the model's performance. However, to succeed in excellent results on massive and complex datasets, researchers have discovered the fantastic generalization ability of several deep learning-based algorithms. As the network grows more complicated, the approaches based on deep learning can give a more abstract picture of the data, enabling the model to uncover essential qualities that machine learning techniques typically overlook.

Because of the substantial inter-class similarities and variety in writing styles, creating highly efficient hand-crafted feature extractors for machine learning techniques is challenging—which deep learning-based methods can automatically learn. Due to these factors, a significant paradigm shift has been noted in recent publications, nearly all of which have concentrated on pipelines based on deep learning. The methods based on deep learning are mainly applied, such as Neural Convolutional CNNs and RNNs or recurrent neural networks. Table 7 shows the performances from the earlier studies with the deep learning methods.

**Table 7.** Performances of the Deep learning-based approaches

| **Dataset** | **Number of images** | **Classifier** | **Accuracy (%)** | **Reference** |
|---|---|---|---|---|
| MNIST | 70,000 | CNN | Max- 99.81% Mean- 99.7741% | (Byerly et al., 2021) |
| Hoda | 78002 | CNN | CNN – 99.34% | (Akhlaghi & Ghods, 2020) |
| MNIST | 70000 | Support Vector Machine (SVM), K- nearest Neighbor (K-NN), Convolutional Neural Network (CNN) | CNN - 99.06% | (Agrawal et al., 2021) |

| | | | | |
|---|---|---|---|---|
| CMATERDB 3.3.1 Arabic handwritten digit dataset | 3000 | Deep learning neural networks using appropriate activation function and regularization layer. | proposed model gives 97.4 percent accuracy | (Ashiquzzaman & Tushar, 2017) |
| Own Dataset | 9800 | CNN | With the Softmax classifier, CNN's accuracy is 98.41%, and with the SVM classifier, it is 99.0%. | (Bhatti et al., 2023) |
| Own Dataset | 1000 | CNN | CNN 99.78% | (Sufian et al., 2022) |
| Own Dataset English, Bengali and Devanagari numerals | 10,000 | CNN | English, Bengali, and Devanagari numbers, CNN has a recognition rate of 97.7%, 95.4%, and 93.7%, respectively | (Roy et al., 2018) |
| Own Data set – English Numerals | 10,000 | CNN | 98.45% and 82.89% for the recognition of single and multi-digit numerals, respectively | (B et al., 2020) |
| Own Dataset – English Character and Numeral | Not specified | DCNN | **98.33% classification accuracy similar to DCNN** | (Arsalan & Santra, 2019) |
| IIT Bhuwaneshwar Dataset – Odia | 5166 | LSTM CNN | 97.93% | (Das et al., 2020) |

## 7. Challenges in offline and online multilingual numeral recognition:

The multilingual numeral string recognition system has several factual business and functional applications, developing the researchers' interest in exploring this field. There are various reported challenges in the case of offline and online multilingual numeral recognition. The following are some of these challenges.

**7.2 Challenges in Offline Multilingual Numeral Recognition:**

1. Different languages utilize a variety of scripts, each with its unique character forms and styles (e.g., Latin, Arabic, Chinese). It may be challenging to distinguish between numbers since the structure of the digits differs across different scripts.
2. **Handwriting Variability**: Characters in printed documents are often standardized; nevertheless, there may be much variation in how numerals are written in handwriting, making identification more difficult.
3. Documents may have sophisticated layouts and a variety of languages and scripts. It might be challenging to distinguish and recognize digits inside such materials accurately.
4. **Historical characters**: When working with historical records, obsolete or archaic characters might make the identification process more difficult.
5. Sensor or other noises could be challenging to correct and identify the numerals. This noise could have occurred due to poor paper quality, blotches, fading ink, and wrinkles.
6. **Multilingual Dictionaries**: It is complicated to create and keep up-to-date, complete multilingual dictionaries that include all potential number variances between languages.

**7.3 Challenges in Online Multilingual Numeral Recognition:**

1. Due to dynamic input, a novel algorithm or system must adjust the various stroke sequences and speeds.
2. There could be an ambiguity in collected data via digital writing; hence, robust algorithms are needed.
3. There is difficulty in input data that must be processed fast to get a positive response in digit recognition and verification.
4. The collected data from various input devices could challenge the recognition of numerals.
5. There is a challenge in integrating the information collected via various multimodal sources like spoken languages, gestures, and handwriting.
6. In some instances, recognizing numbers in a user's handwriting involves adaptation and learning from user-specific patterns (for example, for personalized note-taking applications).

7. Handwriting styles may differ significantly between cultures, so it is essential to create models that consider these distinctions.

## 8. Conclusions and Future Scope:

The present review focuses on online-offline handwritten multilingual numeral recognition systems for various purposes. We provided the systematic approach to recognizing the multilingual numerals that contain the datasets used, the renovation of the datasets, segmentations, and feature extractions followed by the classifications. In addition, this review provides the available datasets for online-offline multilingual numeral recognition. It was found that researchers mostly prefer deep learning-enabled systems with online-offline datasets for recognizing several numerals. Multilingual numeral recognition still needs to be improved due to the complexity of acquired datasets. Thus, developing novel online datasets for regional languages for various applications is suggested. The work on online multilingual numeral recognition must be effectively done due to its several applications. In conclusion, this review finds challenges in detecting real-time texts and suggests standard regional databases, effective methodology, and robust algorithms.

Compliance with Ethical Standards:

* Disclosure of potential conflicts of interest – All authors declare that there are no potentila conflicts of interest.
* Research involving human participants and/or animals – No human and/or animal participants.
* Informed consent – All are agree to the consent.

## Bibliography

Agrawal, A. K., Shrivas, A. K., & Awasthi, V. K. (2021). A robust model for handwritten digit recognition using machine and deep learning technique. *2021 2nd International Conference for Emerging Technology, INCET 2021*, 1–4. https://doi.org/10.1109/INCET51464.2021.9456118

Akhlaghi, M., & Ghods, V. (2020). Farsi handwritten phone number recognition using deep learning. *SN Applied Sciences*, *2*(3), 1–10. https://doi.org/10.1007/s42452-020-2222-5

Al-Ma'adeed, S., Elliman, D., & Higgins, C. A. (2002). A data base for Arabic handwritten text recognition research. *Proceedings - International Workshop on Frontiers in Handwriting Recognition, IWFHR*, 485–489. https://doi.org/10.1109/IWFHR.2002.1030957

Al-Ohali, Y., Cheriet, M., & Suen, C. (2003). Databases for recognition of handwritten Arabic cheques. *Pattern Recognition*, *36*(1), 111–121. https://doi.org/10.1016/S0031-3203(02)00064-X

Al-wzwazy, H., Albehadili, H. M., Alwan, Y. S., & Islam, N. (2016). *Convolutional Neural Networks*. *May 2017*. https://doi.org/10.15680/IJIRCCE.2016.

Alamri, H., Sadri, J., Ching Y Suen, Nobile, N., & Suen, C. Y. (2008). A Novel Comprehensive Database for Arabic Off-Line Handwriting Recognition. *Eleventh International Conference on Frontiers in Handwriting Recognition*, *January*, 664--669, Montreal.

Amara, N. E. Ben, Mazhoud, O., Bouzrara, N., & Ellouze, N. (2005). ARABASE: A Relational Database for Arabic OCR Systems. *Int. Arab J. Inf. Technol.*, *2*(4), 259–266.

Arsalan, M., & Santra, A. (2019). Character Recognition in Air-Writing Based on Network of Radars for Human-Machine Interface. *IEEE Sensors Journal*, *19*(19), 8855–8864. https://doi.org/10.1109/JSEN.2019.2922395

Ashikur Rahman, A. B. M., Hasan, M. B., Ahmed, S., Ahmed, T., Ashmafee, M. H., Kabir, M. R., & Kabir, M. H. (2022). Two Decades of Bengali Handwritten Digit Recognition: A Survey. *IEEE Access*, *10*(July), 92597–92632. https://doi.org/10.1109/ACCESS.2022.3202893

Ashiquzzaman, A., & Tushar, A. K. (2017). Handwritten Arabic numeral recognition using deep learning neural networks. *2017 IEEE International Conference on Imaging, Vision and Pattern Recognition, IcIVPR 2017*, 1–4. https://doi.org/10.1109/ICIVPR.2017.7890866

Ashiquzzaman, A., Tushar, A. K., Rahman, A., & Mohsin, F. (2019). An efficient recognition method for handwritten arabic numerals using CNN with data augmentation and dropout. In *Advances in Intelligent Systems and Computing* (Vol. 808). Springer Singapore. https://doi.org/10.1007/978-981-13-1402-5_23

Azeem, S. A., El Meseery, M., & Ahmed, H. (2012a). Online Arabic handwritten digits recognition. *Proceedings - International Workshop on Frontiers in Handwriting Recognition, IWFHR*, 135–140. https://doi.org/10.1109/ICFHR.2012.249

Azeem, S. A., El Meseery, M., & Ahmed, H. (2012b). Online Arabic handwritten digits recognition. *Proceedings - International Workshop on Frontiers in Handwriting Recognition, IWFHR*, *September 2012*, 135–140. https://doi.org/10.1109/ICFHR.2012.249

B, A. R., Roy, P., & Pal, U. (2020). *Using RNN Architecture in Air-Writing*. *1*, 76–90.

Baati, A. El, & Alimi, A. M. (n.d.). *The On / Off ( LMCA ) Dual Arabic Handwriting Database Menu*.

Baruah, U., & Hazarika, S. M. (2015). A dataset of online handwritten Assamese characters. *Journal of Information Processing Systems*, *11*(3), 325–341. https://doi.org/10.3745/JIPS.02.0008

Basu, S., Das, N., Sarkar, R., Kundu, M., Nasipuri, M., & Basu, D. K. (2009). Recognition of numeric postal codes from multi-script postal address blocks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *5909 LNCS*, 381–386. https://doi.org/10.1007/978-3-642-11164-8_62

Bhattacharya, U., & Chaudhuri, B. B. (2005). *Databases for Research on Recognition of Handwritten Characters of Indian Scripts*. 5–9.

Bhattacharya, U., & Chaudhuri, B. B. (2009). Handwritten numeral databases of Indian scripts and multistage recognition of mixed numerals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(3), 444–457. https://doi.org/10.1109/TPAMI.2008.88

Bhattacharya, U., Shridhar, M., & Parui, S. K. (2006). *On Recognition of Handwritten Bangla Characters*. 817–828. https://doi.org/10.1007/11949619_73

Bhatti, A., Arif, A., Khalid, W., Khan, B., Ali, A., Khalid, S., & Rehman, A. ur. (2023). Recognition and Classification of Handwritten Urdu Numerals Using Deep Learning Techniques. *Applied Sciences (Switzerland)*, *13*(3). https://doi.org/10.3390/app13031624

Biswas, M., Islam, R., & Kumar, G. (2017). Data in Brief BanglaLekha-Isolated : A multi-purpose comprehensive dataset of Handwritten Bangla Isolated characters. *Data in Brief*, *12*, 103–107. https://doi.org/10.1016/j.dib.2017.03.035

Boukharouba, A., & Bennia, A. (2015). Novel feature extraction technique for the recognition of handwritten digits. *APPLIED COMPUTING AND INFORMATICS*. https://doi.org/10.1016/j.aci.2015.05.001

Byerly, A., Kalganova, T., & Dear, I. (2021). No routing needed between capsules. *Neurocomputing*, *463*, 545–553. https://doi.org/10.1016/j.neucom.2021.08.064

Choudhary, A., Rishi, R., Ahlawat, S., & Delhi, N. (2011). *ANN Structure*. 56–65.

Das, A., Patra, G. R., & Mohanty, M. N. (2020). LSTM based Odia Handwritten Numeral Recognition. *Proceedings of the 2020 IEEE International Conference on Communication and Signal Processing, ICCSP 2020*, 538–541. https://doi.org/10.1109/ICCSP48568.2020.9182218

Dash, K. S., Puhan, N. B., & Panda, G. (2015). Handwritten numeral recognition using non-redundant Stockwell transform and bio-inspired optimal zoning. *IET Image Processing*, *9*(10), 874–882. https://doi.org/10.1049/iet-ipr.2015.0146

Dhandra, B. V., Benne, R. G., & Hangarge, M. (2011). "Kannada, Telugu and Devanagari Handwritten Numeral Recognition with Probabilistic Neural Network: A Script Independent Approach." *International Journal of Computer Applications*, *26*(9), 11–16. https://doi.org/10.5120/3134-4319

Dongre, V. J., & Mankar, V. H. (2012). Development of Comprehensive Devnagari Numeral and Character Database for Offline Handwritten Character Recognition. *Applied Computational Intelligence and Soft Computing*, *2012*, 1–5. https://doi.org/10.1155/2012/871834

Ghods, V., & Sohrabi, M. K. (2016). Online Farsi Handwritten Character Recognition Using Hidden Markov Model. *Journal of Computers*, *11*(2), 169–175. https://doi.org/10.17706/jcp.11.2.169-175

H. Akhand, M. A., & Ahmed, M. (2016). Convolutional Neural Network based Handwritten Bengali and Bengali-English Mixed Numeral Recognition. *International Journal of Image, Graphics and Signal Processing*, *8*(9), 40–50. https://doi.org/10.5815/ijigsp.2016.09.06

HAFFNER, Y. L. L. B. Y. B. P. (1998). Gradient-Based Learning Applied to Document Recognition. *PROCEEDINGS OF THE IEEE, VOL. 86, NO. 11, NOVEMBER 1998*, *330*(4), 1299–1305. https://doi.org/10.1016/j.bbrc.2005.03.111

Haghighi, P. J., Nobile, N., He, C. L., & Suen, C. Y. (2009). A new large-scale multi-purpose handwritten farsi database. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5627 LNCS*, 278–286. https://doi.org/10.1007/978-3-642-02611-9_28

Hamad, K., & Kaya, M. (2016). A Detailed Analysis of Optical Character Recognition Technology. *International Journal of Applied Mathematics, Electronics and Computers*, *4*(Special Issue-1), 244–244. https://doi.org/10.18100/ijamec.270374

Hanmandlu, M., Nath, A. V, Mishra, A. C., & Madasu, V. K. (2007). *Fuzzy Model Based Recognition of Handwritten Hindi Numerals using Bacterial Foraging*. *Icis,* 5–10.

Hull, J. J. (1994). A Database for Handwritten Text Recognition Research. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *16*(5), 550–554. https://doi.org/10.1109/34.291440

Hussain, R., Raza, A., Siddiqi, I., Khurshid, K., & Djeddi, C. (2015). A comprehensive survey of handwritten document benchmarks: structure, usage and evaluation. *Eurasip Journal on Image and Video Processing*, *2015*(1), 1–24. https://doi.org/10.1186/s13640-015-0102-5

India, H. L. (2010). *HP Labs India*. Raining Dataset Hpl-Tamil-Iwfhr06-Train for IWFHR 2006 Online Tamil Handwritten Character Recognition Competition. http://shiftleft.com/mirrors/www.hpl.hp.com/india/research/penhw-resources/tamil-iwfhr06-train.html

Jou, C., & Lee, H. C. (2009). Handwritten numeral recognition based on simplified structural classification and fuzzy memberships. *Expert Systems with Applications, 36*(9), 11858–11863. https://doi.org/10.1016/j.eswa.2009.04.025

Kharma, N. (1999). A New Comprehensive Database of Hadritten Arabic Words, Numbers, and Signatures used for OCR Testing. *Proceedings of the 1999 IEEE Canadian Conference on Electrical and Computer Engineering Shaw Conference Center, Edmonton, Alberta, Canada May 9-12 1999, 37*(2), 766–768.

Kodgire, S. V., & Sable, G. S. (2016). *A Review on Optical Character Recognition and Text to Speech Conversion*. *5*(6), 1964–1970.

Latif, G. (2018). Deep Convolutional Neural Network for Recognition of Unified Multi-Language Handwritten Numerals. *2018 IEEE 2nd International Workshop on Arabic and*

*Derived Script Analysis and Recognition (ASAR)*, 90–95.

Liu, C. L., Nakashima, K., Sako, H., & Fujisawa, H. (2003). Handwritten digit recognition: Benchmarking of state-of-the-art techniques. *Pattern Recognition*, *36*(10), 2271–2285. https://doi.org/10.1016/S0031-3203(03)00085-2

Mahmoud, S., & Abu-Amara, M. (2010). Recognition of handwritten Arabic (Indian) numerals using Radon-Fourier-based features. *Proceedings of the 9th WSEAS International Conference on Signal Processing, Robotics and Automation, January 2010*, 158–163.

Majhi, B., Sa, P. K., & Panda, S. (2014). *Model based odia numeral recognition*. *8*(6), 916–922.

Md, Shopon; Nabeel, M. M. A. A. (2016). Bangla Handwritten Digit Recognition Using Autoencoder and Deep Convolutional Neural Network. *2016 International Workshop on Computational Intelligence (IWCI) 12-13 December 2016, Dhaka, Bangladesh, December*, 12–13. https://doi.org/10.1145/3377049.3377077

Mehdi, F., Srinivasa, G., Ashwini, A. J., & K, H. K. H. (2013). *Online Hand Written Character Recognition*. *11*(5), 30–36.

Murthy, O. V. R. (2011). *Interactive Fuzzy Model Based Recognition of Handwritten Characters*. *2*, 154–165.

Musleh, D., Halawani, K., & Mahmoud, S. (2021). *Fuzzy Modeling for Handwritten Arabic Numeral Recognition*. *January 2017*.

Obaidullah, S. M., Mondal, A., Das, N., & Roy, K. (2014). Script Identification from Printed Indian Document Images and Performance Evaluation Using Different Classifiers. *Applied Computational Intelligence and Soft Computing*, *2014*, 1–12. https://doi.org/10.1155/2014/896128

Pramanik, R., Dansena, P., & Bag, S. (2019). A Study on the Effect of CNN-Based Transfer Learning on Handwritten Indic and Mixed Numeral Recognition. In *Communications in Computer and Information Science* (Vol. 1020). Springer Singapore. https://doi.org/10.1007/978-981-13-9361-7_4

Rabby, A. K. M. S. A., Haque, S., Islam, M. S., Abujar, S., & Hossain, S. A. (2019). Ekush: A Multipurpose and Multitype Comprehensive Database for Online Off-Line Bangla Handwritten Characters. In *Communications in Computer and Information Science* (Vol.

1037). Springer Singapore. https://doi.org/10.1007/978-981-13-9187-3_14

Ramakrishnan, A. G. (2013). *Global and Local Features for Recognition of Online Handwritten Numerals and Tamil Characters*.

Razzak, M. I., Hussain, S. A., & Sher, M. (2009). Numeral recognition for Urdu script in unconstrained environment. *2009 International Conference on Emerging Technologies, ICET 2009*, 44–47. https://doi.org/10.1109/ICET.2009.5353203

Roy, P. (2014). *Image Segmentation Using Rough Set Theory : 1*(December), 62–74. https://doi.org/10.4018/ijrsda.2014070105

Roy, P., Ghosh, S., & Pal, U. (2018). A CNN Based framework for unistroke numeral recognition in air-writing. *Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR*, *2018-Augus*, 404–409. https://doi.org/10.1109/ICFHR-2018.2018.00077

Sagheer, M. W., He, C. L., Nobile, N., & Suen, C. Y. (2009). *A New Large Urdu Database for Off-Line Handwriting Recognition BT - Image Analysis and Processing – ICIAP 2009*. *37*, 538–546.

Salouan, R., Safi, S., & Bouikhalene, B. (2014). *Printed Eastern Arabic Noisy Numerals Recognition Using Hidden Markov Model and Support Vectors Machine*. *9*(3), 1032–1042.

Shamim, S. M., Miah, M. B. A., Sarker, A., Rana, M., & Jobair, A. Al. (2018). Handwritten digit recognition using machine learning algorithms. *Indonesian Journal of Science and Technology*, *3*(1), 29–39. https://doi.org/10.17509/ijost.v3i1.10795

Singh, P. K., Das, S., Sarkar, R., & Nasipuri, M. (2017). Recognition of offline handwriten Devanagari numerals using regional weighted run length features. *2016 International Conference on Computer, Electrical and Communication Engineering, ICCECE 2016*, 1–6. https://doi.org/10.1109/ICCECE.2016.8009567

Sufian, A., Ghosh, A., Naskar, A., Sultana, F., Sil, J., & Rahman, M. M. H. (2022). BDNet: Bengali Handwritten Numeral Digit Recognition based on Densely connected Convolutional Neural Networks. *Journal of King Saud University - Computer and Information Sciences*, *34*(6), 2610–2620. https://doi.org/10.1016/j.jksuci.2020.03.002

Takruri, M., Al-hmouz, R., Al-hmouz, A., & Momani, M. (2015). *Fuzzy C Means Based Hybrid*

*Classifiers For Offline Recognition Of Handwritten Indian ( Arabic ) Numerals*. *April 2016*.

Viard-Gaudin, C., Lallican, P. M., Knerr, S., & Binter, P. (1999). The IRESTE On/Off (IRONOFF) dual handwriting database. *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, 459–462. https://doi.org/10.1109/ICDAR.1999.791823

Wilkinson, R. A., & Creecy, R. (n.d.). *The First Census Optical Character Recognition System Conference*.