





Review

# Advancements and Challenges in Handwritten Text Recognition: A Comprehensive Survey

Wissam AlKendi <sup>1,\*</sup> , Franck Gechter <sup>1</sup> , Laurent Heyberger <sup>2</sup>  and Christophe Guyeux <sup>3</sup> 

<sup>1</sup> CIAD, UMR 7533, UTBM, F-90010 Belfort, France; franck.gechter@utbm.fr

<sup>2</sup> FEMTO-ST Institute/RECITS, UMR 6174 CNRS, UTBM, F-90010 Belfort, France; laurent.heyberger@utbm.fr

<sup>3</sup> FEMTO-ST Institute/DISC, UMR 6174 CNRS, Université de Franche-Comté, F-90016 Belfort, France; christophe.guyeux@univ-fcomte.fr

\* Correspondence: wissam.al-kendi@utbm.fr; Tel.: +964-7702666363

**Abstract:** Handwritten Text Recognition (HTR) is essential for digitizing historical documents in different kinds of archives. In this study, we introduce a hybrid form archive written in French: the Belfort civil registers of births. The digitization of these historical documents is challenging due to their unique characteristics such as writing style variations, overlapped characters and words, and marginal annotations. The objective of this survey paper is to summarize research on handwritten text documents and provide research directions toward effectively transcribing this French dataset. To achieve this goal, we presented a brief survey of several modern and historical HTR offline systems of different international languages, and the top state-of-the-art contributions reported of the French language specifically. The survey classifies the HTR systems based on techniques employed, datasets used, publication years, and the level of recognition. Furthermore, an analysis of the systems' accuracies is presented, highlighting the best-performing approach. We have also showcased the performance of some HTR commercial systems. In addition, this paper presents a summarization of the HTR datasets that publicly available, especially those identified as benchmark datasets in the International Conference on Document Analysis and Recognition (ICDAR) and the International Conference on Frontiers in Handwriting Recognition (ICFHR) competitions. This paper, therefore, presents updated state-of-the-art research in HTR and highlights new directions in the research field.

**Keywords:** handwritten text recognition (HTR); HTR datasets; machine learning; Belfort civil registers of births; ICFHR; ICDAR



**Citation:** AlKendi, W.; Gechter, F.; Heyberger, L.; Guyeux, C.

Advancements and Challenges in Handwritten Text Recognition: A Comprehensive Survey. *J. Imaging* **2024**, *10*, 18. <https://doi.org/10.3390/jimaging10010018>

Academic Editor: Cosimo Distanto

Received: 4 December 2023

Accepted: 3 January 2024

Published: 8 January 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In recent years, handwritten text recognition has become one of the most critical research fields of pattern recognition. Many researchers proposed techniques to facilitate the possibility of transcribing historical archives [1], medical prescriptions [2], general forms, and any modern documents, through spatial (offline) or temporal (online) process [3]. Figure 1 shows the classification of the text recognition systems. Transcription involves automatically transforming a source's handwritten text within a digital image into its machine text representation. Optical Character Recognition (OCR) [4] represents the cornerstone technique of this field. It consists of two main phases: firstly, detecting the text by segmenting it into small patches. Secondly, recognizing the contents of the patches in order to be transcribed into machine-coded text, the first and simplest OCR developed in [5] for the recognition of Latin numerals.

Fortunately, HTR systems have incredibly improved since utilizing the Hidden Markov Model (HMM) for text recognition and handcrafted features [6–8]. However, the recognition results of HMMs are still poor due to some drawbacks in the model, such as memorylessness [9] and the manual feature selection process. Researchers overcome these problems by proposing hybrid systems that combine additional architecture with HMM, for instance,

HMM with Gaussian mixture emission distributions (HMM-GMM) [10], HMM with Convolutional Neural Network (CNN) [11], or HMM with Recurrent Neural Network (RNN) [12] which have significantly improved the outcomes.

Nowadays, systems can analyze document layouts and recognize letters, text lines, paragraphs, and whole documents. Arguably, these modern systems can recognize different handwritten styles in French, Arabic, Urdu, Chinese, and other languages. It includes the utilization of machine learning techniques such as Convolutional Neural Networks (CNN) [13], Recurrent Neural Networks (RNN) [14], Convolutional Recurrent Neural Networks (CRNN) [15], Gated-CNN [16], Multi-Dimensional Long Short-Term Memory Recurrent Neural Networks (MDLSTM-RNNs) [17]. Despite the many significant advancements in the past few years, there are still many challenges that need to be addressed.

Civil registers of births, or vital records, were created during the French Revolution. Since 1792, for all French men and women, regardless of their religion, all births that occur in each commune have been recorded, whether or not the parents reside in the commune concerned.

The study of the records of the commune of Belfort is very important for historical reasons: Belfort is a *hapax* in the French demographic landscape of the 19th century because, following the defeat of France against Prussia in 1871, this city experienced a mushrooming growth due to the influx of a large number of Alsatians, especially Mulhousians, who either decided to remain French after the annexation of Alsace–Lorraine by the newly created Germany or migrated to Belfort to continue working for the textile and mechanical industries that massively settled in the city after 1871. This city thus constitutes a privileged observatory of three major changes in characteristics of the European 19th century, in terms of urban growth, migration, and the sexualization of social relations. The historical issues that can be grasped from birth records alone are therefore numerous and important (study of midwives, concubinage, age at first child, inter-birth period and contraception, births outside of marriage, recognition of children, choice of witnesses for the declaration, choice of the child's first name and home or hospital births). To effectively analyze and address these historical questions, it is essential to establish a comprehensive knowledge database that will facilitate their exploration and resolution. This process involves transcribing the archive contents through two approaches. The first is manual transcription by humans. However, this approach is expensive and time-consuming. The second is automatic transcription utilizing computer science methods.

### *1.1. Purpose and Contributions*

Our intention is to provide an overview of existing handwritten text recognition models and highlight the commercially usable systems to acquire insights into the achievements in both modern and historical document recognition on a global scale, with a specific focus on the French language. This background will allow researchers to determine which novel techniques or combinations of techniques should be explored in order to transcribe the hybrid (printed and handwritten) text, as exemplified in the Belfort civil registers of births.

This paper also provides researchers with an up-to-date survey of state-of-the-art models. The survey encompasses research papers at four HTR levels: word, text line, paragraph, and page. Furthermore, a list of 14 publicly available datasets has been summarized for French and other languages, including the datasets used in the International Conference on Frontiers of Handwriting Recognition (ICFHR) and the International Conference on Document Analysis and Recognition (ICDAR) competitions.

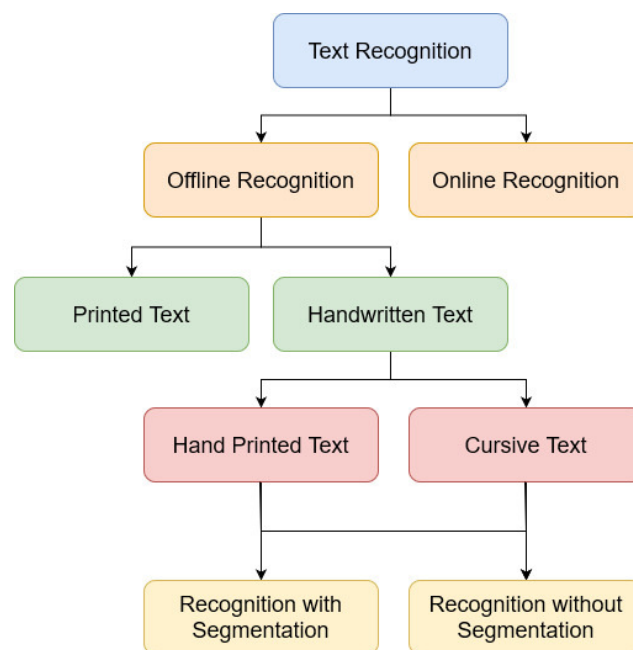
### *1.2. Selection Methodology*

We followed specific criteria in selecting the publications, including those that introduce novel techniques, are accessible, and widely cited. Our focus was primarily on recent studies published in reputable scientific journals and conference proceedings to ensure the reliability and credibility of the information. Additionally, we also took into consideration

the publication year for both international and French-language HTR systems over a time span ranging from 2012 to 2023.

Furthermore, we have selected datasets presented at the International Conference on Document Analysis and Recognition (ICDAR) and the International Conference on Frontiers of Handwriting Recognition (ICFHR). Additionally, our aim was to present the most well-known datasets for various languages, focusing on historical document image datasets and those specific to the French language. This selection spans from 1999 to 2021. Also, we selected and presented some HTR commercial systems based on their offering of free trials for text recognition.

These approaches allowed for providing valuable insights into the advancements and challenges within this domain. The publications and datasets were selected using Google Scholar and Web of Science.



**Figure 1.** Classification of handwritten text recognition systems.

The rest of this paper is organized as follows: Section 2 provides an overview of the literature on HTR models and recent achievements in recognizing handwritten text, specifically focusing on the French language. In Section 3, a summary of the publicly available datasets is presented. Section 4 is devoted to providing a summary of the state-of-the-art experimental results. Then in Section 5, a discussion on the current work related to the French language is proposed. In Section 6, we present suggestions for future research directions. Finally, the conclusion is drawn in Section 7.

## 2. Literature Review

### 2.1. State-of-the-Art Recent Surveys

Several surveys have been published to advance the field and address its challenges. Authors of [18–20] presented a systematic literature review that summarized and analyzed the research articles conducted on character recognition of handwritten text documents across six languages in order to highlight research directions. Others conducted a survey on automated approaches in character recognition of historical documents [21,22]. The studies covered historical manuscripts in various languages, including Indian, Kurdish-Latin, Arabic, ancient Chinese, and others. They also summarized the techniques used for data pre-processing and the types of datasets utilized. Additionally, in [23], authors conducted a survey on state-of-the-art applications, techniques, and challenges in Arabic language character recognition. In [24], authors surveyed the challenges of the recognition

and classification of named entities in various historical resources and languages, including the French language. Additionally, they discussed the employed approaches in the named entity recognition (NER) field and highlighted directions for future developments.

In [25], authors introduced an open database of historical handwritten documents that are fully annotated in the Norwegian language. To assess the performance of state-of-the-art HTR models on their dataset, they conducted a systematic survey of open-source HTR models, including twelve models with different characteristics. Their study highlighted the best performing technique and suggested a combination of different models to further improve performance. In [26], authors present a systematic literature review of image datasets for historical document image analysis at two scales: document classification and layout structure or content analysis. The research aims to assist researchers in identifying the most suitable datasets for their techniques and to advance the field of historical document image analysis. Similarly, others focused on the databases and benchmarks of this field [27].

On the other hand, authors of [28] surveyed the major phases of historical document digitization process, focusing on the standard algorithms, tools, and datasets within the field. Their research highlighted the critical importance of transcription accuracy as a prerequisite for meaningful information retrieval in archival documents. In contrast, authors of [29] focused on the feature extraction phase in handwritten Arabic text recognition.

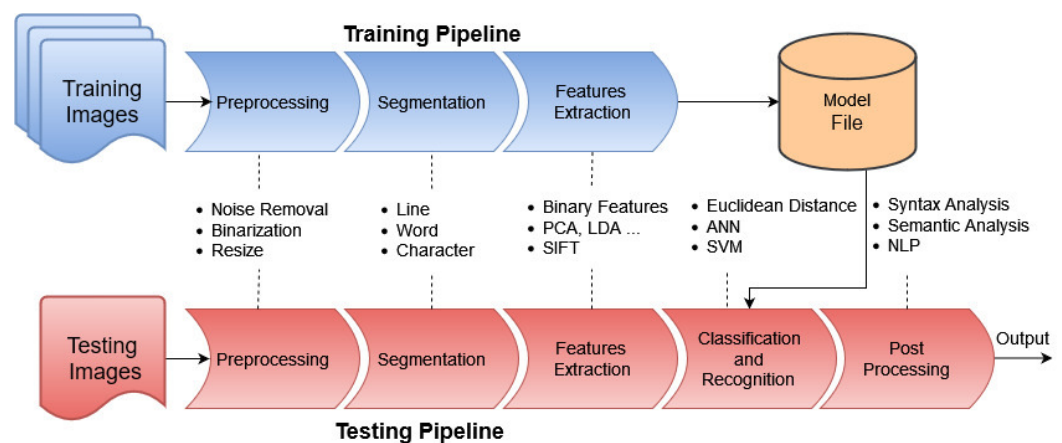
Additionally, in [30], authors presented a critical study of various document layout analysis (DLA) techniques aimed at detecting and annotating the physical structure of documents. The survey highlighted the different phases of DLA algorithms, including preprocessing, layout analysis strategies, post-processing, and performance evaluation. This research serves as a base step toward achieving a universal algorithm suitable for all types of document layouts.

In the study [31], authors discussed the importance of separating machine-printed texts and handwritten texts in hybrid-form documents to enhance the overall system accuracy. The discussion involved techniques employed for the separation process based on feature extraction methods for three categories: structural and statistical features, gradient features, and geometric features.

This research establishes an updated state-of-the-art survey of HTR systems, different languages datasets, HTR competitions, and HTR commercial systems of the French language and other international ones.

## 2.2. Handwritten Text Recognition Workflow

The workflow of handwritten text recognition includes classical image processing approaches and deep learning approaches. Figure 2 illustrates the general pipeline of the HTR workflow.



**Figure 2.** Handwritten text recognition general pipeline.

### 2.2.1. Image Digitization

It is the process of transforming a handwritten text image into an electronic form using various devices such as scanners and digital cameras. This form of the image can be used as the input to the pre-processing stage.

### 2.2.2. Pre-Processing

Pre-processing is the initial stage in enhancing digital images, it involves several key processes such as:

- **Binarization:** This process involves converting digital images into binary images consisting of dual collections of pixels in black and white (0 and 1). Binarization is valuable for segmenting the image into foreground text and background.
- **Noise removal:** This process involves eliminating unwanted pixels from the digitized image that can affect the original information. This noise may originate from the image sensor and electronic components of a scanner or digital camera. Various methods have been proposed for noise removal or reduction, such as Non-local means [32] and Anisotropic diffusion [33], as well as filters like Gaussian, Mean, and Median filters.
- **Edges detection:** This process involves identifying the edges of the text within the digitized image using various methods such as Sobel, Laplacian, Canny, and Prewitt edge detection.
- **Skew detection and correction:** Skew refers to the misalignment of text within a digital image. In other words, it indicates the amount of rotation needed to align the text horizontally or vertically. Various methods for skew detection and correction have been proposed to address this objective, such as Hough transforms and clustering.
- **Normalization:** This process involves reducing the shape and size variation of digital images. Additionally, it scales the input image features to a fixed range (e.g., between 0 and 1), while maintaining the relationship between these features. This process plays a valuable role in the training stage of deep learning models.

### 2.2.3. Segmentation

This process involves dividing the handwritten text image into characters, words, lines, and paragraphs, often utilizing the pixel characteristics within the image. Several methods have been developed to conduct the segmentation process, including threshold methods, region-based methods, edge-based methods, watershed-based methods, and clustering-based methods. The segmentation stage is considered as one of the most crucial steps that can significantly improve the accuracy of HTR models [34].

### 2.2.4. Feature Extraction

This process involves extracting specific information from the images that precisely represents the image, with the goal of reducing the size of high-dimensional data. The feature extraction process aims for higher discriminating power and control overfitting problems within HTR models. However, it may lead to a loss of data interpretability [35]. The most popular techniques for feature extraction include Principal Component Analysis (PCA), Linear Discriminant Analysis, and Independent Component Analysis (ICA). The accuracy of HTR systems is highly dependent on the choice of feature extraction techniques.

### 2.2.5. Classification

This process involves deciding the class membership in the HTR system. The decision-making criteria depend on comparing the input feature with a predefined pattern to identify the most suitable matching class for the input. Two key methods can be employed in this stage. First, the template-based method [36] calculates the correlation between the input and the predefined pattern. Second, the feature-based method [37] utilizes the extracted features from the input for classification.



### 2.2.6. Post-Processing

This stage aims to improve the results of the classification stage and enhance the overall accuracy of the HTR models. It involves output error correction using various techniques such as Dictionary lookup and statistical approaches. However, this stage is not compulsory in the HTR models development process.

### 2.3. *Advancements in Handwritten Text Recognition: A State-of-the-Art Overview*

Languages, by nature, vary in their letter shapes and word connections. The process of recognizing handwritten text is complicated due to many impediments such as the variety of writing style, documents poor quality, noise, spots on paper, and text alignment.

Researchers have proposed several models for recognizing language scripts using different architectures, including Convolutional Neural Network (CNN), Convolutional Recurrent Neural Network, sequence-to-sequence Transformer, Bidirectional Long Short-Term Memory (BLSTM), and others.

In a recent study, a Deep Learning (DL) system using two CNN architectures, named HMB1 and HMB2, was employed to recognize handwritten Arabic characters [38]. The models were trained using a complex Arabic handwritten characters dataset (HMBD), as well as CMATER [39,40] and AIA9k [41] datasets. The model demonstrated a significant accuracy rate when testing HMB1 on HMBD, which further improved when a data augmentation process was applied. Additionally, CMATER and AIA9k datasets were utilized to validate the generalization of the model. In the same context, two different architectures, namely the transformer transducer and sequence-to-sequence transformer, have been established in [42]. These architectures were evaluated using the KFUPM Handwritten Arabic Text (KHATT) dataset [43–45]. Several pre-processing steps, such as text lines rotation and elimination of empty spaces, were performed. An attention mechanism was then applied to reduce the model complexity and achieve satisfactory performance in terms of accuracy and latency, surpassing the previous literature on the KHATT benchmark. Similarly, in [46], a light encoder–decoder transformer-based architecture was introduced for handwriting text recognition, while in [47], an end-to-end approach with pre-trained image transformer and text transformer models for text recognition at the word level was proposed. More survey articles and recent research works for recognizing handwritten text can be found in [17,48–52].

Other researchers introduced Attention-Based Fully Gated CNN supported by multiple bidirectional gated recurrent units (BGRU), with a Connectionist Temporal Classification (CTC) model to predict the sequence of characters [53]. They evaluated this method using five well-known datasets within the HTR community, which encompassed Institut für Informatik und Angewandte Mathematik (IAM) [54], Saint Gall [55], Bentham [56], and Washington [57] for English, as well as the Russian–Kazakh dataset (HKR) [58]. The method achieved a remarkably high recognition rate with minimal parameter usage when applied to the latter dataset.

Furthermore, a novel approach was introduced that combines depthwise convolution with a gated-convolutional neural network and bidirectional gated recurrent unit [59]. This technique effectively reduces the total number of parameters while simultaneously enhancing the overall performance of the model. Authors of [15,60] presented Convolutional Recurrent Neural Network (CRNN) architecture; the latter utilized CRNN for handwriting recognition as an encoder for the input text lines while utilizing a Bidirectional Long Short-Term Memory (BLSTM) network followed by a fully CNN as a decoder to predict the sequence of characters. IAM and Reconnaissance et Indexation de données Manuscrites et de fac similés (RIMES) [61] with the newly created dataset (EPARCHOS) [60] that includes historical Greek manuscripts have been used in the evaluation process of the proposed architecture. Experiments produced improved results compared with the other state-of-the-art methods, especially on the RIMES dataset. Table 1 summarizes the articles based on the architectures used at different HTR levels for different international languages.

**Table 1.** State-of-the-art architectures utilized for handwritten text recognition of different international Languages datasets such as English, Arabic, Russian, and others at different prediction levels

Reference	Architecture	Dataset	HTR Level
[38]	AHCR-DLS (2-CNN)	HMBD, CMATER and AIA9k	Character level
[42]	Transformer-T and Transformer with Cross-Attention	KHATT	Character (Subword) level
[46]	Light Transformer	IAM	Character level
[53]	Attention-Gated-CNN-BGRU	Kazakh	Character level
[59]	CRNN-MDLSTM	IAM and George Washington	Line level
[60]	OctCNN-BGRU	EPARCHOS, IAM and RIMES	Line level
[15]	CRNN-FCNN	EPARCHOS, IAM and RIMES	Line level
[62]	OrigamiNet	IAM ICDAR 2017	Page level

Conversely, many researchers evaluated various models on the French handwritten text. In the competition organized at ICDAR2011 [63] using the RIMES dataset, authors of [64] presented a combination of three techniques of multi-word recognition. Firstly, Grapheme-based Multilayer perceptron (MLP)-HMM was used to decompose the words into letters. Secondly, sliding window Gaussian mixture HMM was used to model the letters with a HMM model using a Gaussian distribution mixture for the observation probabilities. Finally, training a MDLSTM-RNN model using raw values of pixels represents words as inputs. The system recorded an advanced recognition rate on the RIMES dataset for both word and multi-word recognition tasks. Similarly, an MDLSTM-RNN-CTC model using Graphics processing unit (GPU)-based implementation was proposed in [65] to decrease the model training time by processing the input text lines in a diagonal-wise mode, while in [66], authors applied the concept of curriculum learning to the MDLSTM-Convolution Layers-Fully Connected Network (FCN)-CTC model in order to enhance the learning process speed. Additionally, an attention-based RNN-LSTM architecture was proposed in [67]. This architecture was evaluated using the RIMES 2011 dataset. Another research study [66] on datasets such as IAM, RIMES, and OpenHaRT demonstrated significant improvement rates as a result of applying the curriculum learning concept to their model. However, due to alignment issues commonly observed in attention models caused by the recurrence alignment operation, authors of [68] introduced a decoupled attention network (DAN). The DAN is an end-to-end text recognizer that comprises three components. Firstly, a feature encoder is utilized to extract visual features from the source image. Secondly, a convolutional alignment module is employed. Finally, a decoupled text decoder is used for the prediction stage. The model has undergone numerous experiments using IAM and RIMES datasets, achieving effectiveness and merit recognition rates.

Furthermore, in [69], authors presented a system based on recurrent neural networks with weighted finite state transducers and an automatic mechanism for preparing annotated text lines to facilitate the model training process. The model used to decode sequences of characters or words on Maurdor [70] dataset. In the same context, the work was extended to text line recognition; the approach depends on segmenting the text line into words classified based on the confidence score into an anchor and non-anchor words (AWs and NAWs), AWs were equated to the BLSTM outputs, while dynamic dictionaries were created for NAWs by exploiting web resources for their character sequence. Finally, text lines were decoded using dynamic dictionaries [71].

Additionally, authors of [72] introduced a combination of a deep convolutional network with a recurrent encoder–decoder network to predict the sequence of characters at the word level within IAM and RIMES datasets images. Furthermore, in [73], authors combined CTC approaches with Sequence-To-Sequence (S2S) model to improve the recog-

recognition rate on the text line level, they developed the model based on a CNN as a visual backbone, BLSTM as encoder, and a Transformer used for character-wise S2S decoding. The evaluation process using IAM, RIMES, and Staatsarchiv des Kantons Zürich (StAZH) datasets shows competitive recognition results with 10–20 times fewer parameters.

Authors of [74] claimed that Multidimensional Long Short-Term Memory networks might not be necessary to attain good accuracy for HTR due to expensive computational costs. Instead, they suggested an alternative model that relies only on convolutional and one-dimensional recurrent layers. Experiments were carried out using IAM and RIMES 2006 datasets and achieved faster performance with equivalent or better results than MDLSTM models. Similarly, a multilingual handwriting recognition model that leverages a convolutional encoder for input images and a bidirectional LSTM decoder was presented to predict character sequences [75].

Additionally, as neural networks require large data amount in the training stage in order to improve the accuracy of recognition, and because the process of creating transcription data is expensive and time-consuming, authors of [76] presented a model architecture that aims to automatically transcribe Latin and French medieval documentary manuscripts produced between the 12th and 15th centuries based on a CRNN network with a CTC loss; the model has been trained depending on The Alcar-HOME database, the e-NDP corpus, and the Himanis project [77].

Recently, some research studies proposed an end-to-end architecture to recognize handwriting text at the paragraph level [78–80], the latter study introduced an end-to-end transformer-based approach for text recognition and named entities from multi-line historical marriage records images in the competition ICDAR 2017 (Esposalles [81] and French Handwritten Marriage Records (FHMR) [82]), while in [78], an end-to-end recurrence-free fully convolutional network named Simple Predict & Align Network (SPAN) was presented, performing OCR on RIMES, IAM, and READ 2016 at the paragraph level. Additionally, in [79], Vertical Attention Network (VAN), a novel end-to-end encoder–decoder segmentation-free architecture using hybrid attention was introduced.

Furthermore, research in [62,83,84] introduced segmentation-free document level recognition. Authors of [62] proposed a simple neural network module (OrigamiNet) that can augment with any fully convolutional single line text recognizer to convert it into a multi-line/full page recognizer. They conducted model experiments on the ICDAR2017 [85] competition and IAM datasets to demonstrate the applicability and generality of the proposed module, while authors of [83,84] presented Document Attention Network (DAN), an end-to-end architecture for full document text and layout recognition. DAN trained using a pre-defined labeling module for transcribing pages by tags style equivalent to Extensible Markup Language (XML) style aims to process the physical and geometrical information with language supervision only and reduce the annotation costs. The model predicts the document's text lines in parallel after determining the first character of each line. It has undergone numerous experiments using RIMES 2009 and READ 2016 and showed highly beneficial recognition rates on a text line, paragraph, and document level. Table 2 summarizes articles based on the architectures used at different HTR levels for the French language.

More methodologies on the recognition of French and other languages on different prediction levels can be found in: characters [86,87], words [88–90], lines [91,92], paragraphs [93,94], and pages [95,96].



**Table 2.** State-of-the-art architectures utilized for handwritten text recognition of French language datasets at different prediction levels.

Reference	Architecture	Dataset	HTR Level
[64]	Grapheme-based MLP-HMM + Gaussian Mixture HMM + MDLSTM-RNN	RIMES	Word and multi-word level
[68]	Decoupled Attention Network (DAN)	IAM and RIMES	Word level
[72]	Deep Convolutional Network + Recurrent Encoder-Decoder Network	IAM and RIMES	Word level
[65]	MDLSTM + RNN + CTC	IAM and RIMES	Line level
[74]	CNN + 1D-LSTM + CTC	IAM and RIMES	Line level
[66]	MDLSTM + Covolution Layers + FCN + CTC	IAM, RIMES 2011 and OpenHaRT	Line level
[91]	MDLSTM + CTC	IAM, RIMES and OpenHaRT	Line level
[67]	Attention-based RNN + LSTM	RIMES	Line level
[73]	CNN + BLSTM + S2S + CTC	IAM, RIMES and StAZH	Line level
[75]	Gated-CRNN	IAM and RIMES	Paragraph level
[80]	Transformer joint	ICDAR 2017 Esposalles and FHMR	Paragraph level
[78]	Simple Predict & Align Network (SPAN)	RIMES, IAM and READ 2016	Paragraph level
[79]	Vertical Attention Network (VAN)	RIMES, IAM and READ 2016	Paragraph level
[83,84]	Document Attention Network (DAN)	RIMES 2009 and READ 2016	Page level

#### 2.4. Commercial Systems in Handwritten Text Recognition

There are some online HTR commercial systems available for transcribing both modern and historical text such as Transkribus [97], Ocelus, Konfuzio, and DOCSUMO. All of these systems use artificial intelligence technology to recognize the target texts in a variety of languages, including English, French, Spanish, and more. These systems can be extremely beneficial in the archives transcribing process. It is important to acknowledge that the majority of these systems demand a cost for their services. Nevertheless, this cost is frequently lower than manual transcription services cost and greatly lowers the time required for the task. Some of these systems have offered free trials of their services, where users have the opportunity to test it on a variety of handwritten text images with diverse writing styles. These HTR commercial systems could be accessed through the links summarized in Table 3.

We provided an overview of the state-of-the-art in HTR systems, focusing on recent surveys, advancements in recognition methodologies, and commercial solutions. Various HTR architectures were employed for the recognition of different languages at various prediction levels. These levels range from character and word recognition to line, paragraph, and document recognition.

The selection of an appropriate architecture is frequently depends on the characteristics of the target language and dataset. The following section presents some freely available

modern and historical datasets that could be utilized to improve the performance of HTR systems.

**Table 3.** List of HTR commercial systems and their corresponding links.

Name	Link
Transkribus	<a href="https://readcoop.eu/transkribus/">https://readcoop.eu/transkribus/</a> (accessed on 21 November 2023)
Ocelus	<a href="https://ocelus.teklia.com/">https://ocelus.teklia.com/</a> (accessed on 21 November 2023)
Konfuzio	<a href="https://konfuzio.com/en/document-ocr/">https://konfuzio.com/en/document-ocr/</a> (accessed on 21 November 2023)
DOCSUMO	<a href="https://www.docsumo.com/free-tools/online-ocr-scanner">https://www.docsumo.com/free-tools/online-ocr-scanner</a> (accessed on 21 November 2023)

### 3. Datasets

In recent decades, various handwritten text dataset images have been utilized for both modern and historical documents in different languages to enhance the recognition rates of state-of-the-art systems, particularly those identified in the literature as benchmark datasets. Table 4 provides information about the language and year of the datasets, while Table 5 shows more detailed insights into these datasets.

**Table 4.** Languages and year information of the datasets.

1999	•	IAM (English) [54]
2001	•	Washington (English) [98,99]
2006	•	Saint Gall (Latin) [100]
	•	RIMES (French) [61]
2012	•	KHATT (Arabic) [43,44]
2013	•	Esposalles (Spanish) [81]
	•	MAURDOR (French, English and Arabic) [70]
2016	•	Bentham (English) [101]
	•	READ (German) [102,103]
2017	•	HIMANIS Guérin(French) [77]
2019	•	e-NDP (French)
	•	StAZH (Swiss-German)
2020	•	EPARCHOS (Greek) [60]
2021	•	HOME-Alcar (French)

**Table 5.** Datasets details: total number of pages, lines, and words.

Dataset	Language	Total Pages	Total Lines	Total Words
IAM	English	1539	13,353	115,320
Bentham	English	25,000	-	-
Washington	English	20	656	4894
KHATT	Arabic	2000 Paragraph	6742	-
EPARCHOS	Greek	120	9285	18,809
READ	German	30,000	-	-
Saint Gall	Latin	60	1410	11,597
Esposalles	Spanish	199	7063	-
MAURDOR (HT)	French, English, and Arabic	8129	49,412	-
RIMES	French	1500	12,723	-
HOME-Alcar	French	330 acts	-	-
e-NDP	French	500	33,735	-
HIMANIS Guérin	French	1500	30,000	-

### 3.1. IAM

In 1999, authors of [54] introduced Institut für Informatik und Angewandte Mathematik (IAM). This is a modern English gray-scaled handwriting benchmark dataset composed of 1539 pages of scanned text written by 657 writers with a resolution of 300 dpi, which is based on the Lancaster–Oslo/Bergen (LOB) corpus. The dataset includes about 13,353 isolated and labeled text lines with their corresponding transcriptions and consists of 79 different characters, including white space. IAM is able to train HTR models at page, paragraph, line, and word levels.

### 3.2. Washington

Washington dataset [98,99] (part of IAM Historical database) is constituted by English historical manuscripts created from George Washington's papers from the 18th century at the Library of Congress. It has data consisting of 20 pages only, 656 text lines binarized and normalized, 4894 words, and 82 unique characters, and annotated at the word level.

### 3.3. Saint Gall

Saint Gall dataset, its also part of IAM Historical database, contains handwritten historical manuscripts in Latin from the 9th century written by a single writer; the manuscript was originally from the Abbey library of Saint Gall, Switzerland. It comprises 60-page images at a resolution of 300 dpi, 1410 text lines binarized and normalized, 11,597 words, 4890-word labels, 5436-word spellings, and 49 letters. Ground truth was created using the semi-automatic proceeding proposed in [100], including text line locations, word locations, and transcription at line level.

### 3.4. Bentham

Bentham is a collection of historical manuscripts written by English philosopher Jeremy Bentham (1748–1832) and transcribed by volunteers. It consists of legal forms and drafts in English, with a few pages in French and Latin. Additionally, a public web platform has been utilized to transcribe approximately 25,000 pages [101]. The ground truth of the dataset possesses information about the layout and the transcription at line level in page format.

### 3.5. KHATT

In 2012, authors of [43,44] produced the KFUPM Handwritten Arabic Text (KHATT), a dataset of unconstrained handwritten Arabic text, developed by a research group from KFUPM, Dhahran, Saudi Arabia, and written by 1000 distinct writers. It includes 2000 similar text paragraph images and 2000 unique text paragraph images of 6742 extracted text line images with manually verified ground truth.

### 3.6. EPARCHOS

EPARCHOS is a Greek historical handwritten codex written by Antonius Eparchos and Camillos Zanettus [60]. It includes around 120 scanned handwritten text pages, 9285 text lines, and 18,809 words for 1500–1530, representing the most important abbreviations, logograms, and conjunctions in the Greek codex. The dataset is divided into training, validation, and testing text line sets, 1435 (63%), 381 (17%), and 456 (20%), respectively, consisting of 312 characters.

### 3.7. READ

The READ dataset is a part of the European Union's Horizon 2020 project and was first proposed at ICFHR 2016 competition on HTR [102]. The dataset is a subset of the Ratsprotokolle collection, containing early modern German handwriting. It encompasses the minutes of council meetings held between 1470 and 1805, comprising approximately 30,000 pages. It is also available with annotations in color images at page, paragraph, and line levels.

### 3.8. *Esposalles*

Esposalles dataset [81] consists of colored historical handwritten records at a resolution of 300 dpi representing the marriage licenses books in the archives of the cathedral of Barcelona. It comprises 291 books with information about 600,000 unions between 1451 and 1905. Each book contains an index of approximately 29 pages (1563 text lines) with information about the husbands' surnames, the page number, and information on the transcription text file. Approximately 170 pages include 1747 marriage records and 5500 text lines written by one writer; the records contain information about the husband, wife, parents' names, locations, occupations, and civil statuses.

### 3.9. *StAZH*

The Staatsarchiv des Kantons Zürich (StAZH) is a Swiss–German dataset part of the European Union's Horizon 2020 READ project. It comprises documents of resolutions and enactments from the cabinet and the parliament of the canton of Zurich, spanning the period from 1803 to 1882.

### 3.10. *MAURDOR*

The MAURDOR dataset [70] comprises 10,000 heterogeneous printed, typewritten, or manuscripts documents in French, English, or Arabic, scanned or collected from different sources and annotated manually (5000 in French, 2500 in English, and 2500 in Arabic). The current dataset contains 8129 documents out of the 10,000 in TIFF format and the annotations in XML format.

### 3.11. *RIMES*

Reconnaissance et Indexation de données Manuscrites et de fac similés (RIMES) [61] consists of French mail handwritten gray-scale images at a resolution of 300 dpi. It includes 12,723 pages of 5605 letters written by 1300 volunteers representing the mail sent to companies by fax or postal mail; there are 99 different characters in the dataset. The regions of the pages are classified into several classes: sender, recipient, subject, date, location, opening, body, and attachment. Segmentation and transcription are also available at paragraph, line, and word levels.

### 3.12. *e-NDP*

In 2019, the e-NDP corpus was created, consisting of the registers of Notre-Dame-de-Paris and its cloister from the late medieval period. These registers represent 14,000 pages of the institution minutes held by the canons of the cloister weekly. The registers were written in Latin using a family of cursive scripts. In order to digitize these pages, historians transcribed a total of 500 pages from 26 registers dating between 1326 and 1504 to train HTR models for transcription purposes. Additionally, a specific page layout with stereotypical patterns (e.g., titles, names, and margin notes) assists in the transcription process.

### 3.13. *HOME-Alcar*

In 2021, Institut de Recherche et d'Histoire des Textes (IRHT-CNRS) introduced the HOME-Alcar corpus by the European research project History Of Medieval Europe (HOME). This corpus includes manuscript images with their scholarly editions at line level in addition to the annotation of named entities (e.g., persons' names and places) to train HTR models. It consists of 17 French cartularies, such as the medieval documents of the 12th and 14th centuries, around 3090 acts, with 2760 in Latin and 330 in old French written using different script families Textualis, Cursiva Antiquior, and Semi-Hybrida.

### 3.14. *HIMANIS Guérin*

The training dataset (HIMANIS Guérin) was released and made available on Zenodo. It consists of 1500 images and 30,000 text lines of ground truth written in Latin and old French using a Cursiva script. These materials represent the cartulary registers produced

by the French royal chancery between 1302 and 1483, and contains letters of remission, mandates, ennoblements, and amortizations.

### 3.15. FHMR

French Handwritten Marriage Records (FHMR) [82] is a private dataset containing information concerning the husband, wife, parent's names, occupations, and locations. It provides text images at line and paragraph levels. Figure 3 shows image examples of the datasets. Also, the presented datasets could be accessed through the links in Table 6.



**Figure 3.** Samples of handwritten text datasets written in French and other languages: (a) Bentham (English). (b) Saint Gall (Latin). (c) READ (German). (d) EPARCHOS (Greek). (e) KHATT (Arabic). (f) Esposalles (Spanish). (g) RIMES (French). (h) HOME-Alcar (French). (i) HIMANIS Guérin (French).

**Table 6.** List of the datasets and their corresponding links.

Dataset	Link
IAM	<a href="https://fki.tic.heia-fr.ch/databases/iam-handwriting-database">https://fki.tic.heia-fr.ch/databases/iam-handwriting-database</a> (accessed on 14 October 2023)
Bentham	<a href="https://zenodo.org/record/44519#.Y_h6fx9ByMo">https://zenodo.org/record/44519#.Y_h6fx9ByMo</a> (accessed on 14 October 2023)
Washington	<a href="https://paperswithcode.com/dataset/george-washington">https://paperswithcode.com/dataset/george-washington</a> (accessed on 14 October 2023)
KHATT	<a href="http://khatt.ideas2serve.net/KHATTDownload.php">http://khatt.ideas2serve.net/KHATTDownload.php</a> (accessed on 14 October 2023)
EPARCHOS	<a href="https://zenodo.org/record/4095301#.Y_eylx9ByMo">https://zenodo.org/record/4095301#.Y_eylx9ByMo</a> (accessed on 17 October 2023)
READ	<a href="https://zenodo.org/record/218236#.WnLhaCHhBGF">https://zenodo.org/record/218236#.WnLhaCHhBGF</a> (accessed on 17 October 2023)
Saint Gall	<a href="https://fki.tic.heia-fr.ch/databases/saint-gall-database">https://fki.tic.heia-fr.ch/databases/saint-gall-database</a> (accessed on 17 October 2023)
Esposalles	<a href="http://dag.cvc.uab.es/the-esposalles-database/">http://dag.cvc.uab.es/the-esposalles-database/</a> (accessed on 17 October 2023)
StAZH	<a href="https://archives-quickaccess.ch/stazh">https://archives-quickaccess.ch/stazh</a> (accessed on 18 October 2023)
MAURDOR	<a href="https://catalog.elra.info/en-us/repository/browse/ELRA-E0045/">https://catalog.elra.info/en-us/repository/browse/ELRA-E0045/</a> (accessed on 18 October 2023)
RIMES	<a href="https://paperswithcode.com/dataset/rimes">https://paperswithcode.com/dataset/rimes</a> (accessed on 14 October 2023)
e-NDP	<a href="https://digital.library.unt.edu/ark:/67531/metadc1477117/">https://digital.library.unt.edu/ark:/67531/metadc1477117/</a> (accessed on 19 October 2023)
HOME-Alcar	<a href="https://zenodo.org/record/5600884#.Y_eqoB9ByMo">https://zenodo.org/record/5600884#.Y_eqoB9ByMo</a> (accessed on 19 October 2023)
HIMANIS Guérin	<a href="https://zenodo.org/record/5535306#.Y_emnx9ByMo">https://zenodo.org/record/5535306#.Y_emnx9ByMo</a> (accessed on 19 October 2023)
FHMR	Private



### 3.16. ICDAR and ICFHR Competitions

International Conference on Document Analysis and Recognition (ICDAR) and International Conference on Frontiers in Handwriting Recognition (ICFHR) are held each year, respectively. The latter ICFHR [102,104] was conducted for the first time as an international Workshop on Frontiers of Handwriting Recognition (IWFHR) in Montreal in 1990. Then, the name changed to ICFHR in the year 2008. To date, seventeen conference editions have been held every two years. Similarly, ICDAR [103,105] has been held every two years since 1991.

Both conferences held competitions on historical document recognition using subsets of different languages dataset such as Bentham, READ, RASM, KHATT, and more. For example, ICDAR 2017 competition was for a full-page HTR. The dataset collected from the Alfred Escher Letter Collection (AEC), written in German, French, and Italian, consists of two training sets. The first includes 50 fully annotated images with their corresponding transcription ground truth. In contrast, the second set includes 10,000 images with annotated line breaks. Table 7 exhibits examples of these two competitions and the winner architecture. The competitions datasets could be accessed through the links in Table 8.

**Table 7.** Examples of ICDAR and ICFHR competitions: years, datasets, and winning techniques.

Competition	Dataset	Winner	Winner Approach
ICFHR 2014	Bentham	[106]	CRNN+lexicon
ICDAR 2015	TranScriptorium	[105,106]	CRNN B1&B2
ICFHR 2016	READ	[102]	CRNN+char10-gram
ICDAR 2017	READ	[103]	CRNN+char10-gram
ICFHR 2018	RASM, READ and others	[107]	STPP-PHOCNet

**Table 8.** Links to ICDAR and ICFHR competitions datasets.

Competition	Link
ICFHR 2014	<a href="http://doi.org/10.5281/zenodo.44519">http://doi.org/10.5281/zenodo.44519</a> (accessed on 24 October 2023)
ICDAR 2015	<a href="http://doi.org/10.5281/zenodo.248733">http://doi.org/10.5281/zenodo.248733</a> (accessed on 24 October 2023)
ICFHR 2016	<a href="http://doi.org/10.5281/zenodo.1164045">http://doi.org/10.5281/zenodo.1164045</a> (accessed on 24 October 2023)
ICDAR 2017	<a href="http://doi.org/10.5281/zenodo.835489">http://doi.org/10.5281/zenodo.835489</a> (accessed on 24 October 2023)
ICFHR 2018	<a href="http://doi.org/10.5281/zenodo.1442182">http://doi.org/10.5281/zenodo.1442182</a> (accessed on 24 October 2023)

### 3.17. HTR-United

HTR-United is an online ecosystem hosted on Github that comprises a catalog of datasets and, eventually, models for handwritten text recognition. The developers of a wide range of datasets have made their work available in this catalog, together with the corresponding ground truth data. The portal allows users to download these datasets or contact their owners. The catalog aggregates metadata detailing the content of the datasets, including temporal factors, the number of contributors, and distinguishing qualities, in addition to enabling access to the data. It also contains information about the circumstances behind the datasets' creation, such as the controlling authority, annotation criteria, and data format. The catalog is available online at the following link: <https://htr-united.github.io/> (accessed on 15 November 2023)

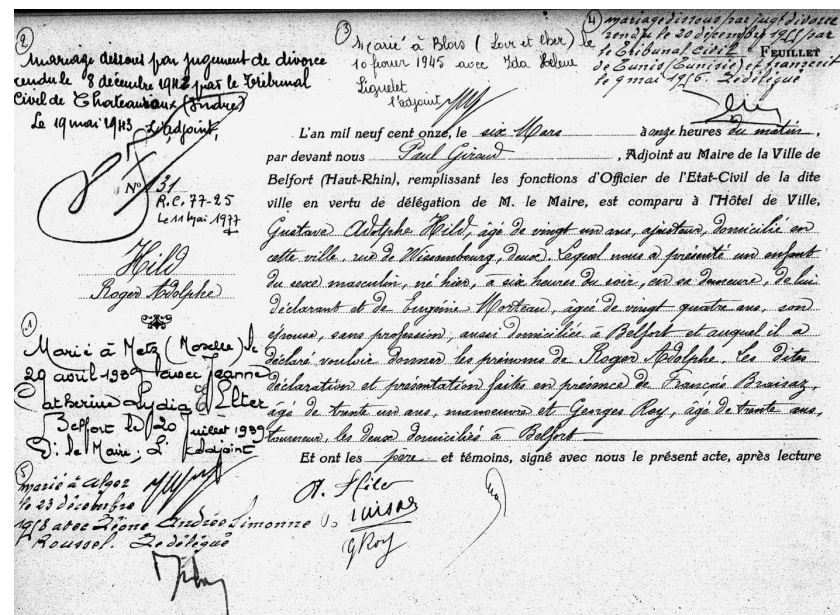
### 3.18. The Belfort Civil Registers of Births

The proposed civil registers of births in the commune of Belfort, spanning from 1807 to 1919, comprises 39,627 birth records at a resolution of 300 dpi. These records were chosen for their homogeneity, as they feature Gregorian dates of birth starting from 1807 and are available until 1919 due to legal reasons.

The registers initially consist of completely handwritten entries, later transitioning to a partially printed format with spaces left free for the individual information concerning

the declaration of the newborn. The transition to this hybrid preprint/manuscript format varied from one commune to another. In Belfort, it occurred in 1885 and concerns 57.5% of the 39,627 declarations. The record contain crucial information, including the child's name, parent' names, witnesses, among other relevant data. Figure 4 provides a visual representation of a sample page from the civil registers, while Table 9 outlines the structure and content of an entry within the archive.

The archive is publicly accessible online until the year 1902 via the following link: <https://archives.belfort.fr/search/form/e5a0c07e-9607-42b0-9772-f19d7bfa180e> (accessed on 12 November 2023). Additionally, we have obtained permission from the municipal archives to access data up to the year 1919.



**Figure 4.** Sample of Belfort civil registers of births, featuring a hybrid mix of printed and handwritten text, along with marginal annotations.

**Table 9.** The structure of an entry in the Belfort civil registers of births

Structure	Content
Head margin	Registration number. First and last name of the person born.
Main text	Time and date of declaration. Surname, first name and position of the official registering. Surname, first name, age, profession and address of declarant. Sex of the newborn. Time and date of birth. First and last name of the father (if different of the declarant). Surname, first name, status (married or other), profession (sometimes) and address (sometimes) of the mother. Surnames of the newborn. surnames, first names, ages, professions and addresses (city) of the 2 witnesses. Mention of absence of signature or illiteracy of the declarant (very rarely).
Margins (annotations)	Mention of official recognition of paternity/maternity (by father or/and mother): surname, name of the declarant, date of recognition (by marriage or declaration). Mention of marriage: date of marriage, wedding location, surname and name of spouse. Mention of divorce: date of divorce, divorce location. Mention of death: date and place of death, date of the declaration of death.

### 3.18.1. Belfort Records Transcription Challenges

Belfort records pose several challenges that complicate the transcription process of its entries, categorized into seven main areas:

- **Document layout:** The Belfort registers of birth exhibit two document layouts. The first type consists of double pages with only one entire entry on each page, while the second type comprises double pages with two entire entries per page. Each entry within these layouts contains the information outlined in Table 9. However, there are some documents where entries begin on the first page and extend to the second page.
- **Reading order:** It is important to identify the reading order of text regions, including the main text and marginal annotation text within the entry.
- **Hybrid format:** Some of the registers consists of entries that includes printed and handwritten text, as shown in Figure 4.
- **Marginal mentions:** These mentions pertain to the individual born but are added after the birth, often in different writing styles and by means of scriptural tools that can be quite distinct. Moreover, they are placed in variable positions compared to the main text of the declaration.
- **Text styles:** The registers are written in different handwritten styles that consist of angular, spiky letters, varying character sizes, and ornate flourishes, resulting in overlapped word and text lines within the script.
- **Skewness:** Skewness refers to the misalignment of handwritten text caused by human writing. Many handwritten text lines in the main paragraphs and margins exhibit variations in text skew, including vertical text (90 degrees of rotation). Effective processes are needed to correct the skewness of the images for any degree of rotation.
- **Degradation:** The images exhibit text degradation caused by fading handwriting and page smudging (ink stains and yellowing of pages).

In this section, we have presented several modern and historical HTR datasets from various languages and time periods, especially those identified as benchmark datasets in ICFHR and ICDAR competitions. The datasets exhibit variations in writing styles and document layouts, offering a rich set of challenges for future research directions.

Additionally, we highlighted the importance of Belfort civil registers of births in advancing the field of handwritten text recognition due to its impediments in document layouts, reading orders, hybrid formats, marginal mentions, diverse text styles, skewness, and degradation. However, overcoming these challenges requires innovative approaches in image processing, text recognition, and layout analysis.

The following section presents state-of-the-art architecture stages and performance accuracy reported on the French dataset (RIMES), along with an accuracy comparison of HTR commercial systems in recognizing English and French languages.

## 4. Results

We presents a comparison of state-of-the-art methods based on French RIMES dataset using Character Error Rate (CER) and word Error Rate (WER) metrics as reported in the publications. This dataset has emerged as a benchmark in the field of handwritten text recognition, many models have been evaluated using this dataset, making it a widely accepted and recognized standard for assessing the performance of such systems. Utilizing the RIMES dataset allows for meaningful and relevant comparisons, ensuring that our research facilitates more accurate assessments of systems performance and highlights the best approaches in this field.

Additionally, we evaluated some of the HTR commercial systems using our proposed dataset and compared their performance with that on an English dataset.

### 4.1. Evaluation Metrics

The standard and common evaluation metrics for measuring the performance of HTR systems are the Word Error Rate (WER) and the Character Error Rate (CER). WER is derived

from the Levenshtein distance [108] and is defined simply as the number of errors in the recognized words divided by the total original words. Calculating the recognized words' error includes adding substitutions, deletions, or insertions to match the recognized words and the ground truth word, as shown in the following formula:

$$WER = \frac{S + D + I}{N} = \frac{S + D + I}{S + D + C} \quad (1)$$

where  $S$  is the number of substitutions,  $D$  is the number of deletions,  $I$  is the number of insertions,  $C$  is the number of correct words, and  $N$  is the number of words in the reference ( $N = S + D + C$ ). For example, a substitution occurs when the word "French" is transcribed as "Erench", and an insertion occurs when a new word(s) is added instead of the original, such as "work" added as "wow rock". Finally, deletion occurs when a word is deleted from the original text, such as "bring it here" becomes "bring here". Furthermore, CER is defined similarly but at the character level instead of the word level. A lower CER/WER result means better accuracy in the text recognition process.

#### 4.2. Advances in Model Performance: State-of-the-Art Results

Experiments on multi-word recognition were carried out in [64] in three stages: word segmentation, word recognition, and language modeling. The first stage relies on the spaces between words to perform word segmentation. The second stage involves calculating the weight of each segmented word to provide N-Best recognition candidates. Finally, a scaling factor is utilized to compensate for the variation between the recognition log probabilities and model log probabilities.

A dataset consisting of 51,739 images of ICDAR2011 (a subset of the RIMES database) was employed as the training set, and 7464 images were devoted to the validation stage. The best performance on isolated word recognition processes based on recurrent neural networks recorded an error rate of nearly 10%, whereas when combining the recognizers, the system achieved an error rate of 5%. Additionally, in the multi-word recognition process, the model was trained on 1300 images, including 6039 different words, while the last 200 images of the training set were used in the validation process. The system achieved a word error rate on the evaluation set 15.2% and 11.8% on the validation set.

In [72], authors evaluated their model using the RIMES database; 11,333 text lines were devoted for training and 778 for testing. However, 10% random samples of the training set have been dedicated to the validation set. The base model recorded a word error rate of 19.1%. Additionally, integrating Layer Normalization (LN), Focal Loss, and Beam Search with the base yielded an improved recognition rate, achieving 9.6% on the same dataset. Furthermore, authors of [73] evaluated CNN-BLSTM-S2S-CTC model and achieved CER and WER of 3.13% and 8.94%, respectively.

According to [74], convolutional and one-dimensional recurrent layers perform better or at least equivalent to MDLSTM models. Experiments were conducted on the RIMES database at the line level, with 10,203 text lines used for training, 1130 for validation, and 778 for testing. The outcomes recorded significant recognition rates, and further improvements were observed when synthetic data augmentation was used during the training stage, resulting in a CER of 2.3% and a WER of 9.6% with a word 4-gram language model, and a CER of 2.5% and a WER of 9.0% without a language model.

Authors of [75] conducted experiments demonstrating that convolutions are faster than LSTMs, and 1D LSTMs are faster than MDLSTMs. They proposed the Gated-CRNN with BLSTM model, which utilizes GPU training to enhance model speed, size, and accuracy. The experiments were conducted using the IAM and RIMES datasets at the line and paragraph levels. The results achieved a CER of 1.9%, and a WER of 7.9% at line level, and a CER of 2.2%, and a WER of 7.9% at paragraph level, without line segmentation process.

Additionally, authors of [80] conducted experiments at both the line and paragraph levels. For their research, 997 FHMR marriage records were allocated to the training set, 103 records for the validation set, and 132 records for the evaluation set. To determine the



best hyper-parameters for the proposed model, they employed a joint tagging encoding technique of the ground truth and conducted an ablation study. The ablation study revealed that the experimental outcomes improved by excluding experiments on images larger than  $384 \times 1024$  pixels. Authors reported that the best performance was achieved using images of size  $256 \times 1024$  pixels at both the paragraph and line levels. They also concluded that 2D positional encoding performs better than 1D positional encoding at both levels. Furthermore, the integration of the “< eol >” label in the ground-truth transcription resulted in improved accuracy rate. The model’s outcomes demonstrated high recognition quality when utilizing Transformer and joint two-stage mixed-level training. In fact, they outperformed other architectures applied to FHMR images at the paragraph level (marriage record).

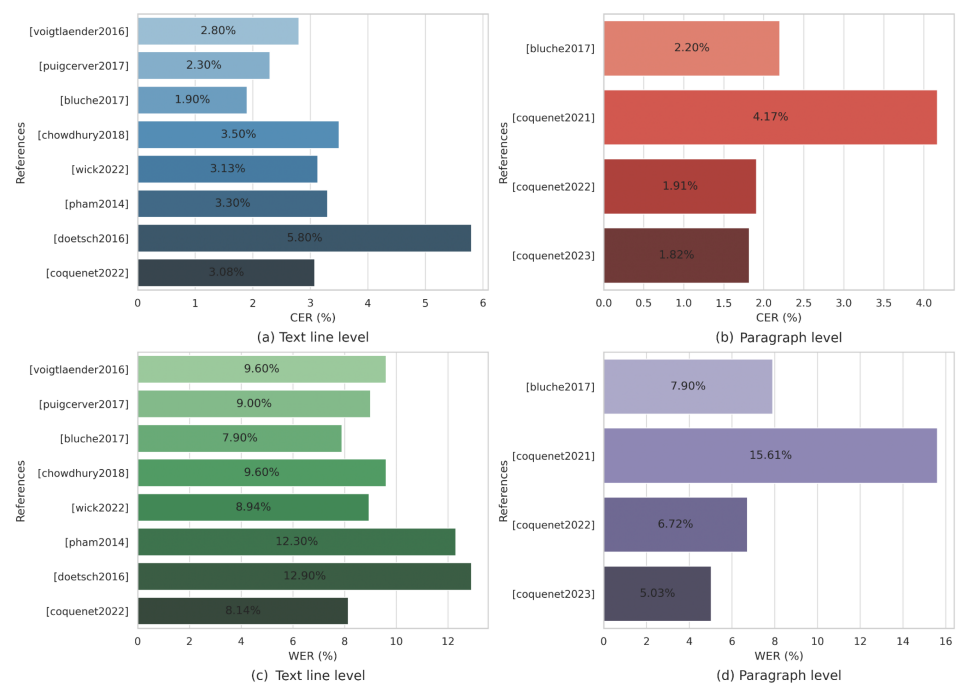
The Simple Predict & Align Network (SPAN) proposed in [78] was tested on the RIMES dataset at both the text line and paragraph levels, in which 1500 paragraphs were devoted to the training set, and 100 paragraphs assigned to the evaluation set. The last 100 training images were utilized as the validation set. In order to mitigate model overfitting, data augmentation techniques were employed during the model training process. These techniques included image resolution modification, perspective transformation, elastic distortion, dilation and erosion, and brightness adjustments. The experimental results were compared with other methods and demonstrated competitive performance in terms of CER and WER metrics. Furthermore, it was observed that pre-training the model at the line level significantly improved the recognition task. They also designated Vertical Attention Network (VAN) for handwritten paragraph text recognition [79]. Experiments conducted on the RIMES dataset using the same training, evaluation, and validation sets as mentioned previously. It involved segmentation and transcription at both the paragraph and line levels. Moreover, data augmentation techniques and post-processing steps were applied. This involved employing CTC best path decoding to identify the last character within the white space character probabilities lattice and subsequently removing it if it was found to be redundant or localized either at the beginning or the end of the text line. VAN was pre-trained on text line images using the learned-stop strategy. The results achieved a CER of 1.91%, and a WER of 6.72%.

Recently, the Document Attention Network (DAN) [83] was evaluated for text recognition and document layout analysis using RIMES 2009 dataset at page and paragraph levels, where 1050 pages were allocated for the training set, 100 pages for the validation set, and 100 pages for the testing set. Additionally, the model was tested using the RIMES 2011 dataset (1400 for training, 100 for validation, and 100 for testing) at paragraph and line levels to compare the performance with other pre-segmented input architectures.

Furthermore, to address the lack of training data, synthetic documents were generated to facilitate the learning of document reading. An ablation study was conducted on the test set of the dataset at single-page and double-page levels with a two-day training period. The model was subjected to various experiments wherein each training component was independently removed. The results obtained from these experiments demonstrated significantly higher accuracy rates compared to other handwritten text recognition approaches. However, the CER percentage may be inaccurate due to misreading order when compared with the ground truth.

Additionally, two metrics have been proposed to evaluate the layout recognition process in handwritten document recognition. First, the Layout Ordering Error Rate (LOER) assesses the accuracy of layout recognition. Second, the mean Average Precision Character Error Rate (mAPCER) is employed to evaluate the recognition of the document layout with respect to the text content. Although the model has shown improvements over the state-of-the-art outcomes, the prediction time is still high, recording linear growth with the number of tokens to be predicted. Figure 5 depicts the CER and WER of state-of-the-art methods at the line and paragraph levels. Table 10 presents a comparison of state-of-the-art methods’ CER and WER accuracies based on the French RIMES dataset.





**Figure 5.** State-of-the-art Character Error Rate (CER) and Word Error Rate (WER) across various studies applied to the French language at two different levels: text line and paragraph. (a) Shows the CER at the text line level, based on studies by Voigtlaender et al. [65], Puigcerver et al. [74], Bluche et al. [75], Chowdhury et al. [72], Wick et al. [73], Pham et al. [91], Doetsch et al. [67], and Coquenot et al. [79]. (b) Depicts the CER at the paragraph level, as reported in studies by Bluche et al. [75], Coquenot et al. [78,79,83]. (c,d) Present the WER at the text line and paragraph levels, respectively, from the same studies.

**Table 10.** Accuracy comparison (%) of state-of-the-art HTR methods on a French-language dataset, based on dataset used and recognition level. Improvement: adding classifier, feature extraction, and segmentation columns to discuss relevant previous research in greater detail.

Dataset	Ref.	Classifier	Feature Extraction	Segmentation	CER (%)	WER (%)	Level
RIMES 2006	[65]	MDLSTM	Automatic	No	2.8	9.6	Line
RIMES 2006	[74]	CNN + 1D-LSTM	Automatic	No	2.3	9.0	Line
RIMES 2006	[75]	GCRNN	Convolutional gates encoder	Yes	1.9	7.9	Line
RIMES 2006	[72]	Deep CN + Recurrent Encoder-Decoder Network	CNN + Sequence Learning	Yes	3.5	9.6	Word
RIMES 2006	[73]	CNN + BLSTM + S2S + CTC	CNN	No	3.13	8.94	Line
RIMES 2006	[75]	GCRNN	Convolutional gates encoder	No	2.2	7.9	Paragraph

Table 10. Cont.

Dataset	Ref.	Classifier	Feature Extraction	Segmentation	CER (%)	WER (%)	Level
RIMES 2009	[68]	Decoupled text decoder	CNN encoder	No	2.7	8.9	Word
RIMES 2009	[64]	Grapheme-based MLP-HMM + Gaussian Mixture HMM + RNN	Multiple methods	Yes	-	4.82	Multi Word
RIMES 2009	[83]	Transformer decoder	FCN encoder	No	5.46	13.04	Paragraph
RIMES 2009	[83]	Transformer decoder	FCN encoder	No	4.54	11.85	Page
RIMES 2009	[84]	Multi-target transformer	FCN encoder	No	6.38	13.69	Page
RIMES 2011	[67]	Attention-based LSTM	Automatic	No	2.9	6.8	Word
RIMES 2011	[64]	Grapheme-based MLP-HMM + Gaussian Mixture HMM + RNN	Multiple methods	Yes	-	4.75	Multi Word
RIMES 2011	[91]	MDLSTM + CTC	Automatic	No	3.3	12.3	Line
RIMES 2011	[67]	Attention-based LSTM	Automatic	No	5.8	12.9	Line
RIMES 2011	[79]	Hybrid attention mechanism	FCN encoder	No	3.08	8.14	Line
RIMES 2011	[83]	Transformer decoder	FCN encoder	No	2.63	6.78	Line
RIMES 2011	[78]	Single convolutional layer + CTC	FCN encoder	No	4.17	15.61	Paragraph
RIMES 2011	[79]	Hybrid attention mechanism	FCN encoder	No	1.91	6.72	Paragraph
RIMES 2011	[83]	Transformer decoder	FCN encoder	No	1.82	5.03	Paragraph

The research was extended in [84] to improve the prediction time of the model through two-step strategy. First, the model predicts the first character of each text line within the page. Second, it predicts all the text lines in parallel using multi-target queries and a document positional encoding scheme. Three document-level datasets have been used to evaluate the speed of the new model, and the findings indicate that the new DAN [84] achieved a faster prediction time, reaching a speed four times faster (1.4 s compared to 5.6 s for the initial DAN). Additionally, the approach demonstrated competitive recognition rates compared to DAN [83] and other state-of-the-art approaches. It is worth mentioning that the models [83,84] were also tested using other language datasets such as READ 2016 and MAURDOR dataset.

Table 11 illustrates the development years of the models. Furthermore, Table 12 displays the training, validation, and testing sets of the French datasets utilized by the state-of-the-art techniques.

**Table 11.** State-of-the-art HTR techniques applied to the French language from 2012 to 2023

Year	2012	2014	2016	2017	2018	2020	2021	2022	2023
Reference	[64]	[66] [91]	[65] [67]	[74] [75]	[72]	[68] [16]	[78]	[73] [80] [79]	[83] [84]

**Table 12.** Details of training, validation, and testing sets in French datasets used by state-of-the-art HTR techniques.

Reference	Dataset	Training Set	Validation Set	Testing Set	Level
[65]	RIMES 2006	11,279	-	778	Line
[74]	RIMES 2006	10,203	1130	778	Line
[72]	RIMES 2006	10,203	1130	778	Line
[73]	RIMES 2006	10,171	1162	778	Line
[83]	RIMES 2009	5875	540	559	Paragraph
[84]	RIMES 2009	1050	100	100	Paragraph
[83]	RIMES 2009	1050	100	100	Page
[67]	RIMES 2011	11,275	1,128	778	Line
[83]	RIMES 2011	10,530	801	778	Line
[91]	RIMES 2011	1400	100	100	Line
[64]	RIMES 2011	1300	200	100	Paragraph
[78]	RIMES 2011	1500	100	100	Paragraph
[79]	RIMES 2011	1500	100	100	Paragraph
[83]	RIMES 2011	1400	100	100	Paragraph
[80]	FHMR	997	103	132	Paragraph

#### State-of-the-Art Techniques Limitations

It is worth noting that limited resources are available for the French historical datasets with their transcriptions. Hence, minimal research has been conducted on such documents. The lack of common benchmarks for historical French texts adds impediments to objectively comparing the performance of various handwritten text recognition systems.

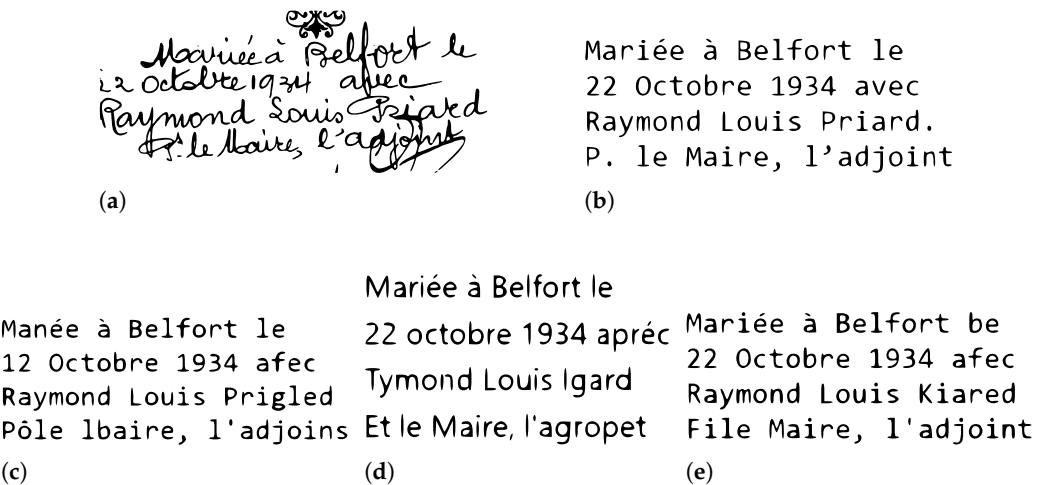
Consequently, most of the studies have been conducted on modern French dataset, such as RIMES, which are specified as standard benchmark dataset in ICFHR and ICDAR competitions. These datasets are commonly used to validate state-of-the-art approaches in the recognition of handwritten French text.

#### 4.3. Commercial HTR systems

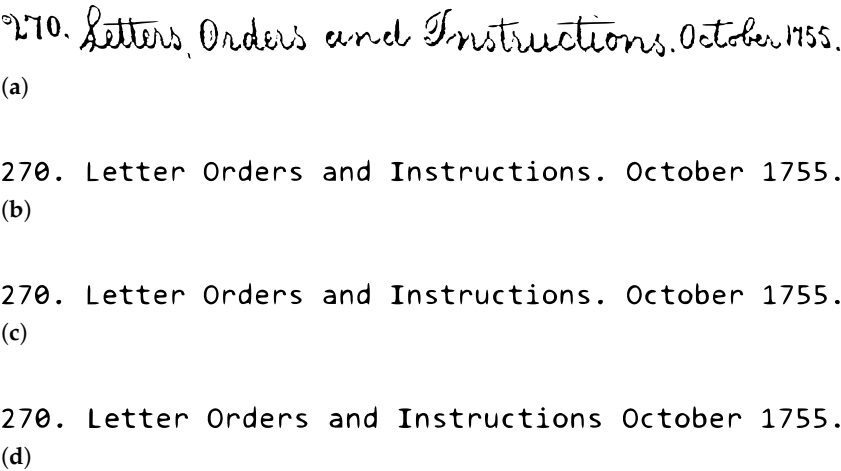
To evaluate the effectiveness of commercial systems in recognizing handwritten text in both English and French languages, three systems: Ocelus, Transkribus, and DOCSUMO were chosen to conduct this experiment, as they are among the most well-known and offer free trials for text recognition. For the evaluation, a text line image of Washington dataset was utilized for the English language, and a margin segment from the proposed Belfort civil registers of births was used for the French language. These experiments allowed us to compare and demonstrate the performance of these systems on the recognition of different international languages and various handwriting styles. Figures 6 and 7 display examples of the results obtained from Ocelus, Transkribus, and DOCSUMO. Table 13 provides a detailed comparison of their performance.

**Table 13.** Accuracy comparison (%) of HTR commercial systems on French- and English-language datasets.

System	RIMES		Washington	
	CER (%)	WER (%)	CER (%)	WER (%)
Ocelus	15	53	2	14
Transkribus	18	33	4	29
DOCSUMO	11	33	2	14



**Figure 6.** Examples of using Ocelus, Transkribus, and DOCSUMO on a margin segment from Belfort civil registers of births (historical French handwritten text): (a) The original text image. (b) The original text ground truth. (c) The result of utilizing Ocelus. (d) The result of utilizing Transkribus. (e) The result of utilizing DOCSUMO.



**Figure 7.** Examples of using Ocelus, Transkribus, and DOCSUMO on Washington dataset (historical English handwritten text): (a) The original text image. (b) The original text ground truth. (c) The result of utilizing Ocelus and DOCSUMO. (d) The result of utilizing Transkribus.

In this section, we have presented an overview of the evaluation stage of HTR models and commercial systems, utilizing metrics such as Character Error Rate (CER) and Word Error Rate (WER). We have also highlighted advances in model design and performance, providing detailed insights into experiments and outcomes from various studies conducted on the RIMES dataset. Additionally, the evaluation of commercial systems in both English and French has revealed varying degrees of accuracy, with the most beneficial performance observed in the context of the English language.

The subsequent section provides a discussion of the optimal architecture for transcribing historical documents, encompassing diverse challenges, including those posed by the French Belfort civil registers of births.

## 5. Discussion

This study surveys several interesting state-of-the-art techniques, and that mainly focus on recognizing French handwritten text. Table 2 illustrates the techniques utilized at each recognition level (word, text line, paragraph, page), where most of the studies concentrated on word and text line levels. Notably, the majority of researchers relayed on LSTMs with CTC systems to fulfill the objectives since the LSTMs networks appear to provide the highest performances in solving high-dimensional dataset situations as compared to other techniques, particularly during the period from 2012 to 2020. This trend is depicted in Table 11, which summarizes the year-wise distribution of these techniques. However, it is worth noting that this addressed architecture still holds expensive computational costs.

Moreover, it is shown that minimal recent studies conducted between 2021 and 2023 have focused on paragraph and page levels using the attention mechanism. As illustrated in Figure 5, the best-performing architecture at the text line level achieved CER of 1.9%, authors of [74] verified that convolutional and one-dimensional recurrent layers perform better than MDLSTMs models in terms of CER metrics. Additionally, the attention mechanism proposed in [83] yielded the highest performance for the WER metric, achieving an impressive 6.78%. The latter also outperformed the state-of-the-art at the paragraph and page levels, as shown in Table 10.

According to the historical datasets detailed in Table 5, it is noticed that limited resources are available for the French historical datasets with their transcriptions. Thus, researchers attempt to devote a data augmentation process to overcoming this limitation. Additionally, they encountered critical challenges within the historical dataset images, such as angular, spiky letters, and ornate flourishes. These challenges result in overlapped word and text lines within the script of different handwritten text styles, leading to highest time consumption for the pre-processing phase, and strengthen the challenges over researchers attempts. Therefore, the demand arises for new considerable transcribed datasets that capture the specificities of such historical sources.

Furthermore, recognizing and digitizing such French historical datasets should be preceded by transcribing and cross-checking processes in order to enhance model development and evaluation processes. Two options can be employed for transcribing historical documents.

Firstly, manual transcription can be performed by teams or by a call to the community. However, it is time-consuming and costly. Additionally, the transcriptions produced by different transcribers may not be uniform, particularly concerning abbreviations, which may result in poor model performance.

Secondly, automatic transcription using commercial systems, as summarized in Table 3. The performance outcomes of these systems, displayed in Figure 6 and Table 13, indicate better performance on historical English language samples in comparison with French language samples. Nevertheless, these systems facilitates an increase in transcription accuracy by permitting instant manual corrections on the transcribed text, which assist in terms of vocabulary (the lexicon). It is worth mentioning that evaluating such commercial systems at a documents level has resulted in improved accuracy rates due to differences in character and word counts.

Additionally, it is also important for researcher to possess language skills when working with historical French documents. This includes a solid understanding of French grammar, vocabulary, spelling, and an awareness of writing style variations.

In response to our research question, the automatic transcription of the Belfort civil registers of births shown in Figure 4 and described in Table 9, as well as other historical documents with similar characteristics, appears to be applicable using a novel attention mechanism technique as employed in [83,84], along with additional pre-processing stages for segmenting the dataset into paragraphs and text line levels. nonetheless, manual tran-



scription of some registers should be performed to train the model before initiating the automatic transcription process. Additionally, we might also consider employing commercial systems for this purpose, with manual corrections and associated costs. However, this archive constitutes many challenges, as in Section 3.18.1, offering rich opportunities for enhancing researchers' models and historical research.

## 6. Suggestions for Future Directions

The field of historical handwritten text recognition poses challenging research problems due to various complications that still need to be addressed. The following are some suggestions for future research directions that could enhance the accuracy of the model recognition:

- **Pre-processing:** Developing an effective approach based on machine learning techniques for historical text skew detection and correction, text degradation, and document layout analysis could highly improve the recognition accuracy of the models.
- **Segmentation:** Implement an additional process after the segmentation stage to refine the segments and preserve the shape of historical handwritten text.
- **Classification models:** Combining various classifiers to handle the recognition of hybrid-form documents, in other words, documents that include printed and historical handwritten text. This also involves utilizing different optimizers within the model to improve the accuracy rate.
- **Post-processing:** Extend the existing post-processing techniques to capture semantic relationships between words and reduce errors by excluding less reliable predictions.

## 7. Conclusions

Handwritten text recognition systems have made significant progress in recent years, becoming increasingly accurate and reliable. In this study, we have presented several state-of-the-art models and achievements in offline handwritten text recognition across various international language documents. Additionally, we presented a comprehensive survey of French handwritten text recognition models specifically. The research papers were reviewed at four HTR levels: word, text line, paragraph, and page. Furthermore, we provided a summary of available public datasets for both French and other languages.

Despite significant achievements in recognizing modern handwritten text, there is still a need to extend these capabilities to historical text documents. Historical handwritten text recognition poses unique challenges, such as transcription cost, a variety of writing styles, abbreviations, symbols, and reproduction quality of historical documents.

We also observed that some commercial handwritten text recognition systems are performing exceptionally on handwritten text in English. In contrast, they are inaccurate in recognizing the French historical cursive handwritten text. Nevertheless, these systems could be a promise tool that can assist in automatically transcribing large volumes of historical documents with manual corrections. This is attributed to their advantages in automatic segmentation and dictionary support. Hence, decreasing time and cost.

Finally, we facilitate researchers in identifying the appropriate technique or datasets for further research on both modern and historical handwritten text documents. Furthermore, we conclude that there is a compelling need to design a new technique specifically tailored for transcribing the French Belfort civil registers of births.

**Author Contributions:** Conceptualization, W.A., F.G., L.H. and C.G.; Formal analysis, W.A., F.G. and C.G.; Methodology, W.A., F.G., L.H. and C.G.; Project administration, F.G. and L.H.; Validation, W.A., F.G., L.H. and C.G.; Writing-review and editing, W.A., F.G., L.H. and C.G.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The datasets (Belfort civil registers of births) analyzed during the current study are available online as mentioned, and from the corresponding author on reasonable request.

**Conflicts of Interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Abbreviations

The following abbreviations are used in this manuscript:

DL	Deep Learning
HTR	Handwritten Text Recognition
OCR	Optical Character Recognition
HMM	Hidden Markov Model
CNN	Convolutional Neural Networks
FCN	Fully Convolutional Network
RNN	Recurrent Neural Networks
CRNN	Convolutional Recurrent Neural Networks
BGRU	Bidirectional Gated Recurrent Units
LSTM	Long Short Term Memory
MDLSTM	Multi Dimensional Long Short Term Memory
BLSTM	Bidirectional Long Short Term Memory
S2S	Sequence-To-Sequence
CTC	Connectionist Temporal Classification
GPU	Graphics Processing Unit
LOER	Layout Ordering Error Rate
mAPCER	mean Average Precision Character Error Rate
mAP	mean Average Precision
CER	Character Error Rate
WER	Word Error Rate
XML	Extensible Markup Language
MLP	Multi-Layer Perceptron

## References

- Fontanella, F.; Colace, F.; Molinara, M.; Di Freca, A.S.; Stanco, F. Pattern recognition and artificial intelligence techniques for cultural heritage. *Pattern Recognit. Lett.* **2020**, *138*, 23–29. [[CrossRef](#)]
- Kamalanaban, E.; Gopinath, M.; Premkumar, S. Medicine box: Doctor’s prescription recognition using deep machine learning. *Int. J. Eng. Technol. (UAE)* **2018**, *7*, 114–117. [[CrossRef](#)]
- Bezerra, B.L.D.; Zanchettin, C.; Toselli, A.H.; Pirlo, G. *Handwriting: Recognition, Development and Analysis*; Nova Science Publishers, Inc.: Hauppauge, NY, USA, 2017.
- Lu, Z.; Schwartz, R.; Natarajan, P.; Bazzi, I.; Makhoul, J. Advances in the bbn byblos ocr system. In Proceedings of the Fifth International Conference on Document Analysis and Recognition. ICDAR’99 (Cat. No. PR00318), Bangalore, India, 22–22 September 1999; IEEE: Piscataway, NJ, USA, 1999; pp. 337–340.
- Schantz, H.F. *History of OCR, Optical Character Recognition*; Recognition Technologies Users Association: Manchester, VT, USA, 1982.
- Saritha, B.; Hemanth, S. An efficient hidden markov model for offline handwritten numeral recognition. *arXiv* **2010**, arXiv:1001.5334.
- Rodríguez-Serrano, J.A.; Perronnin, F. A model-based sequence similarity with application to handwritten word spotting. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2108–2120. [[CrossRef](#)] [[PubMed](#)]
- Sánchez, J.A.; Romero, V.; Toselli, A.H.; Villegas, M.; Vidal, E. A set of benchmarks for handwritten text recognition on historical documents. *Pattern Recognit.* **2019**, *94*, 122–134. [[CrossRef](#)]
- Schuster-Böckler, B.; Bateman, A. An introduction to hidden Markov models. *Curr. Protoc. Bioinform.* **2007**, *18*, A.3A.1–A.3A.9. [[CrossRef](#)] [[PubMed](#)]
- Toselli, A.H.; Juan, A.; González, J.; Salvador, I.; Vidal, E.; Casacuberta, F.; Keysers, D.; Ney, H. Integrated handwriting recognition and interpretation using finite-state models. *Int. J. Pattern Recognit. Artif. Intell.* **2004**, *18*, 519–539. [[CrossRef](#)]
- Guo, Q.; Wang, F.; Lei, J.; Tu, D.; Li, G. Convolutional feature learning and Hybrid CNN-HMM for scene number recognition. *Neurocomputing* **2016**, *184*, 78–90. [[CrossRef](#)]
- Cheikhrouhou, A.; Kessentini, Y.; Kanoun, S. Hybrid HMM/BLSTM system for multi-script keyword spotting in printed and handwritten documents with identification stage. *Neural Comput. Appl.* **2020**, *32*, 9201–9215. [[CrossRef](#)]

13. Wu, J. *Introduction to Convolutional Neural Networks*; National Key Lab for Novel Software Technology, Nanjing University: Nanjing, China 2017; Volume 5, 495.
14. Sherstinsky, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Phys. D Nonlinear Phenom.* **2020**, *404*, 132306. [[CrossRef](#)]
15. Markou, K.; Tsochatzidis, L.; Zagoris, K.; Papazoglou, A.; Karagiannis, X.; Symeonidis, S.; Pratikakis, I. A convolutional recurrent neural network for the handwritten text recognition of historical Greek manuscripts. In *Proceedings of the Pattern Recognition. ICPR International Workshops and Challenges, Virtual Event, 10–15 January 2021; Proceedings, Part VII*; Springer: Berlin, Germany, 2021; pp. 249–262.
16. de Sousa Neto, A.F.; Bezerra, B.L.D.; Toselli, A.H.; Lima, E.B. HTR-Flor: A deep learning system for offline handwritten text recognition. In *Proceedings of the 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, Porto de Galinhas, Brazil, 7–10 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 54–61.
17. Zayene, O.; Touj, S.M.; Hennebert, J.; Ingold, R.; Ben Amara, N.E. Multi-dimensional long short-term memory networks for artificial Arabic text recognition in news video. *IET Comput. Vis.* **2018**, *12*, 710–719. [[CrossRef](#)]
18. Memon, J.; Sami, M.; Khan, R.A.; Uddin, M. Handwritten optical character recognition (OCR): A comprehensive systematic literature review (SLR). *IEEE Access* **2020**, *8*, 142642–142668. [[CrossRef](#)]
19. Purohit, A.; Chauhan, S.S. A literature survey on handwritten character recognition. *Int. J. Comput. Sci. Inf. Technol. (IJCSIT)* **2016**, *7*, 1–5.
20. Nisha; Singh, L.; Sandhu, J.K.; Sahu, R. A Literature Survey on Handwritten Character Recognition. In *Proceedings of the 2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)*, Kollam, India, 10–11 August 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1755–1760.
21. Narang, S.R.; Jindal, M.K.; Kumar, M. Ancient text recognition: A review. *Artif. Intell. Rev.* **2020**, *53*, 5517–5558. [[CrossRef](#)]
22. Dhivya, S.; Devi, U.G. Study on automated approach to recognize characters for handwritten and historical document. *ACM Trans. Asian Low Resour. Lang. Inf. Process.* **2021**, *20*.
23. Faizullah, S.; Ayub, M.S.; Hussain, S.; Khan, M.A. A Survey of OCR in Arabic Language: Applications, Techniques, and Challenges. *Appl. Sci.* **2023**, *13*, 4584. [[CrossRef](#)]
24. Ehrmann, M.; Hamdi, A.; Pontes, E.L.; Romanello, M.; Doucet, A. Named entity recognition and classification in historical documents: A survey. *ACM Comput. Surv.* **2023**, *56*, 1–47. [[CrossRef](#)]
25. Maarand, M.; Beyer, Y.; Kåsen, A.; Fosseide, K.T.; Kermorvant, C. A comprehensive comparison of open-source libraries for handwritten text recognition in norwegian. In *Proceedings of the International Workshop on Document Analysis Systems*, La Rochelle, France, 22–25 May 2022; Springer: Berlin, Germany, 2022; pp. 399–413.
26. Nikolaidou, K.; Seuret, M.; Mokayed, H.; Liwicki, M. A survey of historical document image datasets. *Int. J. Doc. Anal. Recognit. (IJDAR)* **2022**, *25*, 305–338. [[CrossRef](#)]
27. Hussain, R.; Raza, A.; Siddiqi, I.; Khurshid, K.; Djeddi, C. A comprehensive survey of handwritten document benchmarks: structure, usage and evaluation. *EURASIP J. Image Video Process.* **2015**, *2015*, 1–24. [[CrossRef](#)]
28. Philips, J.; Tabrizi, N. Historical Document Processing: A Survey of Techniques, Tools, and Trends. In *Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2020)*, Virtual Event, 2–4 November 2020; pp. 341–349.
29. El Qacimy, B.; Hammouch, A.; Kerroum, M.A. A review of feature extraction techniques for handwritten Arabic text recognition. In *Proceedings of the 2015 International Conference on Electrical and Information Technologies (ICEIT)*, Marrakech, Morocco, 25–27 March 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 241–245.
30. Binmakhashen, G.M.; Mahmoud, S.A. Document layout analysis: A comprehensive survey. *ACM Comput. Surv. (CSUR)* **2019**, *52*, 1–36. [[CrossRef](#)]
31. Srivastva, R.; Raj, A.; Patnaik, T.; Kumar, B. A survey on techniques of separation of machine printed text and handwritten text. *Int. J. Eng. Adv. Technol.* **2013**, *2*, 552–555.
32. Buades, A.; Coll, B.; Morel, J.M. A non-local algorithm for image denoising. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, 20–25 June 2005; IEEE: Piscataway, NJ, USA, 2005; Volume 2, pp. 60–65.
33. Perona, P.; Malik, J. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 629–639. [[CrossRef](#)]
34. Abliz, A.; Simayi, W.; Moydin, K.; Hamdulla, A. A survey on methods for basic unit segmentation in off-line handwritten text recognition. *Int. J. Future Gener. Commun. Netw.* **2016**, *9*, 137–152. [[CrossRef](#)]
35. Naik, G.R.; Kumar, D.K. An overview of independent component analysis and its applications. *Informatica* **2011**, *35*, 63–81.
36. Qu, X.; Swanson, R.; Day, R.; Tsai, J. A guide to template based structure prediction. *Curr. Protein Pept. Sci.* **2009**, *10*, 270–285. [[CrossRef](#)] [[PubMed](#)]
37. Hazza, A.; Shoaib, M.; Alshebeili, S.A.; Fahad, A. An overview of feature-based methods for digital modulation classification. In *Proceedings of the 2013 1st International Conference on Communications, Signal Processing, and Their Applications (ICCSIPA)*, Sharjah, United Arab Emirates, 12–14 February 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 1–6.
38. Balaha, H.M.; Ali, H.A.; Saraya, M.; Badawy, M. A new Arabic handwritten character recognition deep learning system (AHCR-DLS). *Neural Comput. Appl.* **2021**, *33*, 6325–6367. [[CrossRef](#)]

39. Das, N.; Sarkar, R.; Basu, S.; Kundu, M.; Nasipuri, M.; Basu, D.K. A Genetic Algorithm Based Region Sampling for Selection of Local Features in Handwritten Digit Recognition Application. *Appl. Soft Comput.* **2012**, *12*, 1592–1606. [CrossRef]
40. Das, N.; Reddy, J.M.; Sarkar, R.; Basu, S.; Kundu, M.; Nasipuri, M.; Basu, D.K. A Statistical-topological Feature Combination for Recognition of Handwritten Numerals. *Appl. Soft Comput.* **2012**, *12*, 2486–2495. [CrossRef]
41. Torki, M.; Hussein, M.E.; Elsallamy, A.; Fayyaz, M.; Yaser, S. Window-based descriptors for Arabic handwritten alphabet recognition: A comparative study on a novel dataset. *arXiv* **2014**, arXiv:1411.3519.
42. Momeni, S.; Babaali, B. Arabic Offline Handwritten Text Recognition with Transformers. [CrossRef]
43. Mahmoud, S.A.; Ahmad, I.; Alshayeb, M.; Al-Khatib, W.G.; Parvez, M.T.; Fink, G.A.; Märgner, V.; El Abed, H. Khatt: Arabic offline handwritten text database. In Proceedings of the 2012 International Conference on Frontiers in Handwriting Recognition, Bari, Italy, 18–20 September 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 449–454.
44. Mahmoud, S.A.; Ahmad, I.; Al-Khatib, W.G.; Alshayeb, M.; Parvez, M.T.; Märgner, V.; Fink, G.A. KHATT: An open Arabic offline handwritten text database. *Pattern Recognit.* **2014**, *47*, 1096–1112. [CrossRef]
45. Ahmad, R.; Naz, S.; Afzal, M.Z.; Rashid, S.F.; Liwicki, M.; Dengel, A. Khatt: A deep learning benchmark on arabic script. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; IEEE: Piscataway, NJ, USA, 2017; Volume 7, pp. 10–14.
46. Barrere, K.; Soullard, Y.; Lemaitre, A.; Coüasnon, B. A light transformer-based architecture for handwritten text recognition. In Proceedings of the Document Analysis Systems: 15th IAPR International Workshop, DAS 2022, La Rochelle, France, 22–25 May 2022; Springer: Berlin, Germany, 2022; pp. 275–290.
47. Li, M.; Lv, T.; Chen, J.; Cui, L.; Lu, Y.; Florencio, D.; Zhang, C.; Li, Z.; Wei, F. Trocr: Transformer-based optical character recognition with pre-trained models. *arXiv* **2021**, arXiv:2109.10282.
48. Parvez, M.T.; Mahmoud, S.A. Offline Arabic handwritten text recognition: A survey. *ACM Comput. Surv. (CSUR)* **2013**, *45*, 1–35. [CrossRef]
49. Jannoud, I.A. Automatic Arabic handwritten text recognition system. *Am. J. Appl. Sci.* **2007**, *4*, 857–864.
50. Alrobah, N.; Albahli, S. Arabic handwritten recognition using deep learning: A Survey. *Arab. J. Sci. Eng.* **2022**, *47*, 9943–9963. [CrossRef]
51. Balaha, H.M.; Ali, H.A.; Youssef, E.K.; Elsayed, A.E.; Samak, R.A.; Abdelhaleem, M.S.; Tolba, M.M.; Shehata, M.R.; Mahmoud, M.R.; Abdelhameed, M.M.; et al. Recognizing arabic handwritten characters using deep learning and genetic algorithms. *Multimed. Tools Appl.* **2021**, *80*, 32473–32509. [CrossRef]
52. Kass, D.; Vats, E. AttentionHTR: Handwritten text recognition based on attention encoder-decoder networks. In Proceedings of the Document Analysis Systems: 15th IAPR International Workshop, DAS 2022, La Rochelle, France, 22–25 May 2022; Springer: Berlin, Germany, 2022; pp. 507–522.
53. Abdallah, A.; Hamada, M.; Nurseitov, D. Attention-based fully gated CNN-BGRU for Russian handwritten text. *J. Imaging* **2020**, *6*, 141. [CrossRef] [PubMed]
54. Marti, U.V.; Bunke, H. A full English sentence database for off-line handwriting recognition. In Proceedings of the Fifth International Conference on Document Analysis and Recognition. ICDAR'99 (Cat. No. PR00318), Bangalore, India, 22 September 1999; IEEE: Piscataway, NJ, USA, 1999; pp. 705–708.
55. Fischer, A.; Frinken, V.; Fornés, A.; Bunke, H. Transcription alignment of Latin manuscripts using hidden Markov models. In Proceedings of the 2011 Workshop on Historical Document Imaging and Processing, Beijing, China, 16–17 September 2011; pp. 29–36.
56. Causser, T.; Wallace, V. Building a volunteer community: Results and findings from Transcribe Bentham. *Digit. Humanit. Q.* **2012**, *6*, 2. Available online: <https://www.digitalhumanities.org/dhq/vol/6/2/000125/000125.html> (accessed on 1 January 2024).
57. Fischer, A.; Keller, A.; Frinken, V.; Bunke, H. Lexicon-free handwritten word spotting using character HMMs. *Pattern Recognit. Lett.* **2012**, *33*, 934–942. [CrossRef]
58. Nurseitov, D.; Bostanbekov, K.; Kurmankhojayev, D.; Alimova, A.; Abdallah, A.; Tolegenov, R. Handwritten Kazakh and Russian (HKR) database for text recognition. *Multimed. Tools Appl.* **2021**, *80*, 33075–33097. [CrossRef]
59. Kumari, L.; Singh, S.; Rathore, V.; Sharma, A. A Lexicon and Depth-Wise Separable Convolution Based Handwritten Text Recognition System. In Proceedings of the Image and Vision Computing: 37th International Conference, IVCNZ 2022, Auckland, New Zealand, 24–25 November 2022; Revised Selected Papers; Springer: Berlin, Germany, 2023; pp. 442–456.
60. Tsochatzidis, L.; Symeonidis, S.; Papazoglou, A.; Pratikakis, I. HTR for greek historical handwritten documents. *J. Imaging* **2021**, *7*, 260. [CrossRef]
61. Augustin, E.; Carré, M.; Grosicki, E.; Brodin, J.M.; Geoffrois, E.; Prêteux, F. RIMES evaluation campaign for handwritten mail processing. In Proceedings of the International Workshop on Frontiers in Handwriting Recognition (IWFHR'06), La Baule, France, 23–26 October 2006; pp. 231–235.
62. Yousef, M.; Bishop, T.E. OrigamiNet: Weakly-supervised, segmentation-free, one-step, full page text recognition by learning to unfold. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 14710–14719.
63. Grosicki, E.; El-Abed, H. Icdar 2011-french handwriting recognition competition. In Proceedings of the 2011 International Conference on Document Analysis and Recognition, Beijing, China, 18–21 September 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 1459–1463.



64. Menasri, F.; Louradour, J.; Bianne-Bernard, A.L.; Kermorvant, C. The A2iA French handwriting recognition system at the Rimes-ICDAR2011 competition. In Proceedings of the Document Recognition and Retrieval XIX, SPIE, Burlingame, CA, USA, 25–26 January 2012; Volume 8297, pp. 263–270.
65. Voigtlaender, P.; Doetsch, P.; Ney, H. Handwriting recognition with large multidimensional long short-term memory recurrent neural networks. In Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 23–26 October 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 228–233.
66. Louradour, J.; Kermorvant, C. Curriculum learning for handwritten text line recognition. In Proceedings of the 2014 11th IAPR International Workshop on Document Analysis Systems, Tours, France, 7–10 April 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 56–60.
67. Doetsch, P.; Zeyer, A.; Ney, H. Bidirectional decoder networks for attention-based end-to-end offline handwriting recognition. In Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 23–26 October 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 361–366.
68. Wang, T.; Zhu, Y.; Jin, L.; Luo, C.; Chen, X.; Wu, Y.; Wang, Q.; Cai, M. Decoupled attention network for text recognition. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 12216–12224.
69. Moyssset, B.; Bluche, T.; Knibbe, M.; Benzeghiba, M.F.; Messina, R.; Louradour, J.; Kermorvant, C. The A2iA multi-lingual text recognition system at the second Maurdor evaluation. In Proceedings of the 2014 14th International Conference on Frontiers in Handwriting Recognition, Heronissos, Greece, 1–4 September 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 297–302.
70. Brunessaux, S.; Giroux, P.; Grilheres, B.; Manta, M.; Bodin, M.; Choukri, K.; Galibert, O.; Kahn, J. The maudor project: Improving automatic processing of digital documents. In Proceedings of the 2014 11th IAPR International Workshop on Document Analysis Systems, Tours, France, 7–10 April 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 349–354.
71. Oprean, C.; Likforman-Sulem, L.; Mokbel, C.; Popescu, A. BLSTM-based handwritten text recognition using Web resources. In Proceedings of the 2015 13th International Conference on Document Analysis and Recognition (ICDAR), Tunis, Tunisia, 23–26 August 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 466–470.
72. Chowdhury, A.; Vig, L. An efficient end-to-end neural model for handwritten text recognition. *arXiv* **2018**, arXiv:1807.07965.
73. Wick, C.; Zöllner, J.; Grüning, T. Rescoring sequence-to-sequence models for text line recognition with ctc-prefixes. In Proceedings of the Document Analysis Systems: 15th IAPR International Workshop, DAS 2022, La Rochelle, France, 22–25 May 2022; Springer: Berlin, Germany, 2022; pp. 260–274.
74. Puigserver, J. Are multidimensional recurrent layers really necessary for handwritten text recognition? In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; IEEE: Piscataway, NJ, USA, 2017; Volume 1, pp. 67–72.
75. Bluche, T.; Messina, R. Gated convolutional recurrent neural networks for multilingual handwriting recognition. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; IEEE: Piscataway, NJ, USA, 2017; Volume 1, pp. 646–651.
76. Torres Aguilar, S.O.; Jolivet, V. Handwritten Text Recognition for Documentary Medieval Manuscripts. *J. Data Min. Digit. Humanit.* **2023**. Available online: <https://hal.science/hal-03892163v2> (accessed on 1 January 2024). [[CrossRef](#)]
77. Stutzmann, D.; Moufflet, J.F.; Hamel, S. La recherche en plein texte dans les sources manuscrites médiévales: Enjeux et perspectives du projet HIMANIS pour l'édition électronique. *Médiévales* **2017**, 67–96, [[CrossRef](#)]
78. Coquenot, D.; Chatelain, C.; Paquet, T. SPAN: A simple predict & align network for handwritten paragraph recognition. In Proceedings of the Document Analysis and Recognition-ICDAR 2021: 16th International Conference, Lausanne, Switzerland, 5–10 September 2021; Proceedings, Part III 16; Springer: Berlin, Germany, 2021; pp. 70–84.
79. Coquenot, D.; Chatelain, C.; Paquet, T. End-to-end handwritten paragraph text recognition using a vertical attention network. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, 45, 508–524. [[CrossRef](#)] [[PubMed](#)]
80. Rouhou, A.C.; Dhiaf, M.; Kessentini, Y.; Salem, S.B. Transformer-based approach for joint handwriting and named entity recognition in historical document. *Pattern Recognit. Lett.* **2022**, 155, 128–134. [[CrossRef](#)]
81. Romero, V.; Fornés, A.; Serrano, N.; Sánchez, J.A.; Toselli, A.H.; Frinken, V.; Vidal, E.; Lladós, J. The ESPOSALLES database: An ancient marriage license corpus for off-line handwriting recognition. *Pattern Recognit.* **2013**, 46, 1658–1669. [[CrossRef](#)]
82. Dhiaf, M.; Jemni, S.K.; Kessentini, Y. DocNER: A deep learning system for named entity recognition in handwritten document images. In Proceedings of the Neural Information Processing: 28th International Conference, ICONIP 2021, Sanur, Bali, Indonesia, 8–12 December 2021; Proceedings, Part VI 28; Springer: Berlin, Germany, 2021; pp. 239–246.
83. Coquenot, D.; Chatelain, C.; Paquet, T. DAN: A segmentation-free document attention network for handwritten document recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, 45, 8227–8243. [[CrossRef](#)] [[PubMed](#)]
84. Coquenot, D.; Chatelain, C.; Paquet, T. Faster DAN: Multi-target Queries with Document Positional Encoding for End-to-end Handwritten Document Recognition. *arXiv* **2023**, arXiv:2301.10593.
85. Cloppet, F.; Eglin, V.; Helias-Baron, M.; Kieu, C.; Vincent, N.; Stutzmann, D. Icdar2017 competition on the classification of medieval handwritings in latin script. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; IEEE: Piscataway, NJ, USA, 2017; Volume 1, pp. 1371–1376.



86. Cilia, N.D.; De Stefano, C.; Fontanella, F.; di Freca, A.S. A ranking-based feature selection approach for handwritten character recognition. *Pattern Recognit. Lett.* **2019**, *121*, 77–86. [\[CrossRef\]](#)
87. Clanuwat, T.; Lamb, A.; Kitamoto, A. KuroNet: Pre-modern Japanese Kuzushiji character recognition with deep learning. In Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, NSW, Australia, 20–25 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 607–614.
88. Shi, B.; Bai, X.; Yao, C. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 2298–2304. [\[CrossRef\]](#)
89. Swaileh, W.; Lerouge, J.; Paquet, T. A unified French/English syllabic model for handwriting recognition. In Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 23–26 October 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 536–541.
90. Bhunia, A.K.; Das, A.; Bhunia, A.K.; Kishore, P.S.R.; Roy, P.P. Handwriting recognition in low-resource scripts using adversarial learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 4767–4776.
91. Pham, V.; Bluche, T.; Kermorvant, C.; Louradour, J. Dropout improves recurrent neural networks for handwriting recognition. In Proceedings of the 2014 14th International Conference on Frontiers in Handwriting Recognition, Hersonissos, Greece, 1–4 September 2014; pp. 285–290.
92. Ingle, R.R.; Fujii, Y.; Deselaers, T.; Baccash, J.; Popat, A.C. A scalable handwritten text recognition system. In Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, NSW, Australia, 20–25 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 17–24.
93. Bluche, T. Joint line segmentation and transcription for end-to-end handwritten paragraph recognition. *Adv. Neural Inf. Process. Syst.* **2016**, *29*. [\[CrossRef\]](#)
94. Bluche, T.; Louradour, J.; Messina, R. Scan, attend and read: End-to-end handwritten paragraph recognition with mdlstm attention. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; IEEE: Piscataway, NJ, USA, 2017; Volume 1, pp. 1050–1055.
95. Moysset, B.; Kermorvant, C.; Wolf, C. Full-page text recognition: Learning where to start and when to stop. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; IEEE: Piscataway, NJ, USA, 2017; Volume 1, pp. 871–876.
96. Wigington, C.; Tensmeyer, C.; Davis, B.; Barrett, W.; Price, B.; Cohen, S. Start, follow, read: End-to-end full-page handwriting recognition. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 367–383.
97. Nockels, J.; Gooding, P.; Ames, S.; Terras, M. Understanding the application of handwritten text recognition technology in heritage contexts: A systematic review of Transkribus in published research. *Arch. Sci.* **2022**, *22*, 367–392. [\[CrossRef\]](#) [\[PubMed\]](#)
98. Rath, T.M.; Manmatha, R.; Lavrenko, V. A search engine for historical manuscript images. In Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Sheffield, UK, 25–29 July 2004; pp. 369–376.
99. Wieprecht, C.; Rothacker, L.; Fink, G.A. Word spotting in historical document collections with online-handwritten queries. In Proceedings of the 2016 12th IAPR Workshop on Document Analysis Systems (DAS), Santorini, Greece, 11–14 April 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 162–167.
100. Fischer, A.; Indermühle, E.; Bunke, H.; Viehhauser, G.; Stolz, M. Ground truth creation for handwriting recognition in historical documents. In Proceedings of the 9th IAPR International Workshop on Document Analysis Systems, Boston, MA, USA, 9–11 June 2010; pp. 3–10.
101. Moyle, M.; Tonra, J.; Wallace, V. Manuscript transcription by crowdsourcing: Transcribe Bentham. *Liber Q.* **2010**, *20*, 347–356. [\[CrossRef\]](#)
102. Sanchez, J.A.; Romero, V.; Toselli, A.H.; Vidal, E. ICFHR2016 competition on handwritten text recognition on the READ dataset. In Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 23–26 October 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 630–635.
103. Sanchez, J.A.; Romero, V.; Toselli, A.H.; Villegas, M.; Vidal, E. ICDAR2017 competition on handwritten text recognition on the READ dataset. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; IEEE: Piscataway, NJ, USA, 2017; Volume 1, pp. 1383–1388.
104. Sánchez, J.A.; Romero, V.; Toselli, A.H.; Vidal, E. ICFHR2014 competition on handwritten text recognition on transcriptorium datasets (HTRtS). In Proceedings of the 2014 14th International Conference on Frontiers in Handwriting Recognition, Hersonissos, Greece, 1–4 September 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 785–790.
105. Sanchez, J.A.; Toselli, A.H.; Romero, V.; Vidal, E. ICDAR 2015 competition HTRtS: Handwritten Text Recognition on the tranScriptorium dataset. In Proceedings of the 2015 13th International Conference on Document Analysis and Recognition (ICDAR), Tunis, Tunisia, 23–26 August 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1166–1170.
106. Strauß, T.; Grüning, T.; Leifert, G.; Labahn, R. Citlab ARGUS for historical handwritten documents. *arXiv* **2014**, arXiv:1412.3949.

107. Rusakov, E.; Rothacker, L.; Mo, H.; Fink, G.A. A probabilistic retrieval model for word spotting based on direct attribute prediction. In Proceedings of the 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), Niagara Falls, NY, USA, 5–8 August 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 38–43.
108. Haldar, R.; Mukhopadhyay, D. Levenshtein distance technique in dictionary lookup methods: An improved approach. *arXiv* **2011**, arXiv:1101.1232.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.