Writer Identification Using Deep Neural Networks: Impact of Patch Size and Number of Patches

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Akshay Punjabi Jose Ramón Prieto Enrique Vidal

Universitat Politècnica de València PRHLT Research Center

PRHLT

Motivation

Every person is used to writing in unique ways and characteristics. This clear variation allows the identification of a writer based on only on the writer's handwriting.
There are multiple applications for this concept: it can be used as evidence in criminal investigations, verification of financial processes or to identify the writer of ancient manuscripts.
Creating an automatic method of identifying the writer would facilitate, especially, in the investigation of large quantities of manuscripts, and, in turn, it would save a lot of time and resources. With the success of deep neural networks in all kinds of tasks, it would be interesting to use them in the field of writer identification. For this reason, in this paper we aspired to create a writer identification system based on a deep learning neural network.

State of the art

Writer identification methods can follow traditional methods that focus on extracting information with handmade algorithms. These, in turn, can focus on the of the ***structure*** (how are words connected) ***textures*** of the image (such as color, width, the thickness of the characters), or in the ***grapheme*** of the letters (particular details of how letters are written). On the other hand, recently there have been new methods based on deep learning. In this table we collect the state of the art of all these methods. Although deep learning methods are proving to be very effective, traditional methods are still the state of the art.

Structure-based approaches					
Year	Features	Classifier	Dataset	Writers	Top1 (%)
2014	MSDH + TDH	KNN	IAM	657	97.1
2014	SDS + SOH	Euclidean	IAM	657	**98.5**
			Firemaker	250	92.4
2012	Quill–Hinge	NN	IAM	657	97
			Firemaker	251	86
Texture-based approaches					
Year	Feature	Classifier	DB	Writers	Top1
2016	Chain code	KDA	IAM	650	82.7
2013	Texture LPQ	SVM	IAM	650	96.7
Grapheme-based approaches					
Year	Feature	Classifier	DB	Writers	Top1
2019	SIFT + RootSIFT	GMM	IAM	650	97.85
			Firemaker	250	**97.98**
2016	p(ls,lθ), p(IBOS)		IAM	657	86.9
2015	Graphemes	SR-KDA	IAM	657	92
2013	Connected	KNN, x2	IAM	650	94.8
			Firemaker	250	95.2
2012	SIFT	x2	IAM	650	93.1
2011	KAS	SVM	IAM	650	92.1
2011	Global and local	KNN, GMM, Bayes	IAM	93	98.76
Combination of structure and grapheme based methods					
Year	Feature	Classifier	DB	Writers	Top1
2010	Codebook and contour	KNN	IAM	650	91
2007	Contour PDFs and ink trace	PDFs	IAM	650	89
			Firemaker	250	83
Deep Learning-based approaches					
Year	Feature	Classifier	DB	Writers	Top1
2020	CNN with word fragments		IAM	657	96.3
			Firemaker	250	97.6
2019	CNN with tuples of images		IAM	650	93.14
			Firemaker	250	93.56
2016	Multi-stream CNN		IAM	657	97.3

Table 1:Summary of state of the art in writer identification methods on IAM and Firemaker datasets with respect to the type of approach. In bold best results in the respective dataset.

github.com/akpun/writer-identification

Model architecture

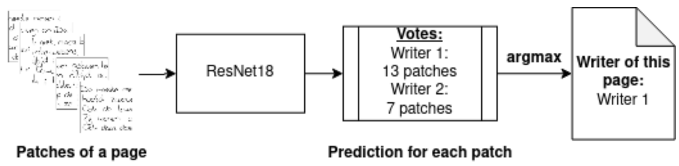

Figure 1:Visual illustration of the voting scheme used for our patch models.

Approach

Using the ResNet18 base model, we have created two model variants, one that uses the entire page as input and the other that uses patches of a page:

- **Model based on full-pages**. Receives a whole page preprocessed and generates the prediction for the writer
- **Model based on patches**. Randomly extract from a page image, n number of square patches with size $a \times a$, where n and a are intuitively chosen. Moreover, from all the patches drawn from a given page, the predictions are calculated to determine the probabilities of the potential writers. The maximum predicted writer for each patch is considered as one vote. As a result, the writer who receives the most votes is considered the predicted writer. An schematic visualization can be seen in Figure 1.

As for preprocessing, we crop the image as compact as possible to avoid unnecessary blank backgrounds and normalize so that the mean is close to 0. We also define a novel technique called ***Text padding***.

In *text padding*, a cropped image is first padded with a white background to the maximum height and width of all cropped images. Then the blank space with the original text is filled with text, making a copy from left to right and from top to bottom. In the image on the right of Figure 2 you can see the final image that is obtained. This can also be seen as a data augmentation technique.

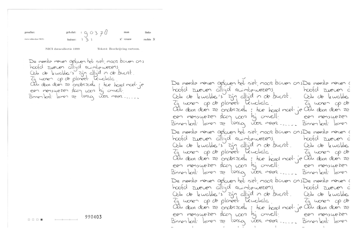

Figure 2:(Left) original page (right) cropping and applying text padding to (left).

Datasets

We experimented in the following datasets:

- **Firemaker**. A data set in which 250 writers provide four pages written in Dutch. We use the first page as the training set and the fourth page as the test set.
- **IAM**. Contains pages of handwritten text from 657 writers providing a different number of pages. In our case, for writers who contributed more than two pages, we only kept the first two, and for writers with only one page, we cut the respective pages in half.
- **ICDAR 17**. Consists of handwritten pages from the 13th to the 20th century. In this data set, 720 writers provide 5 pages each.

Results

Dataset	Full Image Accuracy (%)	Patches Accur. (%)	Patches Patch Size	Patches # Patches	Best from Table 1
IAM	91.3	96.8	600	300	**98.5**
Firemaker	98.3	**99.2**	1200	300	98.0
ICDAR17	83.0	83.6	800	64	--

Table 2:Best results with our proposed patch-based models and full-page models.

Conclusions

- We presented a comprehensive summary of writer recognition approaches, which included both traditional methods and deep learning methods in Table 1
- Proposed a new preprocessing method called *Text padding*
- Demonstrated the promising performance of using bigger patch sizes
- Provided an open-source deep learning based writer identification system that obtained competitive accuracy