

DHQ: Digital Humanities Quarterly

2021

Volume 15 Number 4

Nouf Alrasheed <nalrasheed_at_mail_dot_umkc_dot_edu>, Department of Computer Science. & Electrical Engineering. University of Missouri-Kansas City

Praveen Rao <praveen_dot_rao_at_missouri_dot_edu>, Department of Health Management & Informatics, Department of Electrical Engineering & Computer Science. University of Missouri-Columbia

Viviana Grieco <griecov_at_umkc_dot_edu>, Department of History. University of Missouri-Kansas City

Abstract

Handwritten character recognition is a challenging pattern recognition problem due to the inconsistency of the handwritten scripts and the lack of accurate labeled data. Historical documents written in cursive are even more challenging as characters have unique and varying shapes. Frequently, words are linked by lines and ornamental doodles. When historical documents are digitized, the images contain various types of noise and degradation, which further complicates the recognition of characters. In this paper, we present an empirical study of how well state-of-the-art convolutional neural networks (CNNs) for image classification perform for the task of recognizing handwritten characters in seventeenth-century Spanish American notarial scripts. Professional historians, paleography experts and trained labelers were involved in preparing the labeled dataset of Spanish characters for training the CNNs. The labeled dataset used in this experiment was created from the manuscripts written by one of the multiple scribes that contributed to the collection of approximately 220,000 digitized images of notary records housed at the *Archivo General de la Nación Argentina* (National Archives). We removed the noise in these images by applying standard image processing techniques. After training different CNNs, we computed the classification accuracy for all the characters. We observed that ResNet-50 achieved a promising accuracy of 97.08% compared to InceptionResnet-V2, Inception-V3, and VGG-16, which achieved 96.66%, 96.33% and 70.91%, respectively.

I. INTRODUCTION ^[1]

Notary records contain a wealth of information for understanding different aspects of the human experience. For that reason, historians specialized in different regions and time periods employ them in writing social, economic, political, and cultural histories. The seventeenth-century Spanish American notarial scripts housed in the National Archives of Argentina are among the most challenging collections, as they were written by multiple hands, for an audience of experts, and at a time the written Spanish language underwent significant transformations [Silva Prada 2001] [Wasserman 2018] .^[2] Consequently, it takes years of training and practice in seventeenth-century Spanish paleography to become proficient in reading and analyzing these notarial scripts (Fig.1). On an average, it takes expert Spanish speaking readers about one hour to read a four to five pages long notarized deed. The task is even more daunting for non-native Spanish speakers.

1

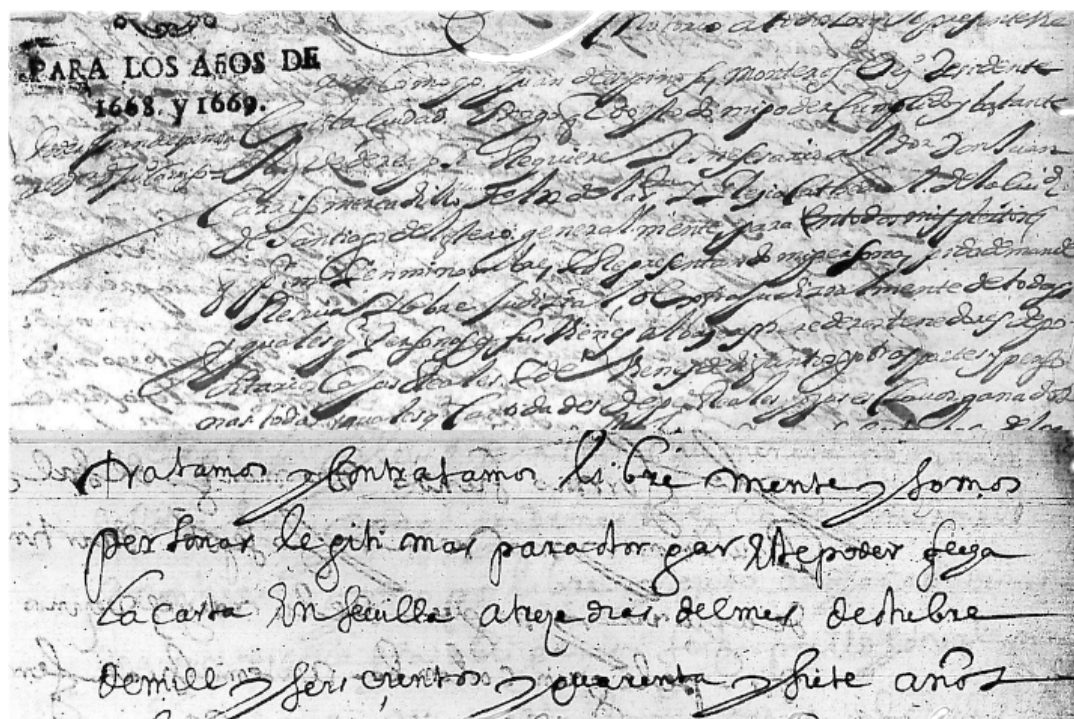


Figure 1. Samples of different handwriting styles present in the collection of seventeenth-century notarial scripts used for this study

Digitization significantly aided in the preservation of these records and made them relatively more accessible primarily by enabling their duplication without damaging the originals. However, despite the quantity and variety of documents this collection compiles, these records are still waiting to be fully utilized in scholarly endeavors. To this day, researchers and students rely on traditional, time consuming, and expensive methods of archival research to access these documents and archival discovery primarily depends on the skill, patience and luck of the scholar. The development of a system capable of storing, reading, querying, and analyzing this historical collection is crucial as it will make these manuscripts accessible to a broader community of researchers without requiring extensive and expensive paleography training. Character recognition is the first step in the development of such a system.

Today, using optical character recognition (OCR), we can automatically convert printed or handwritten text into machine-readable, editable, and searchable text. For that reason, OCR is regarded as the heart of many document analysis systems. Unlike the recognition of printed text, historical handwritten text presents unique challenges. Written in cursive, historical scripts usually employ irregular characters and capitalization, abbreviations, archaic spelling, and linked words (Fig. 2).^[3] Additionally, the scans contain different types of noise including discoloration, stains, as well as ink bleeds and smudges. In order to enable OCR tasks, researchers applied different methods including Support Vector Machine (SVM) [Vellingiraj et al. 2016], K-NN [Chammas et al. 2018], and deep learning [Granell et al. 2018].

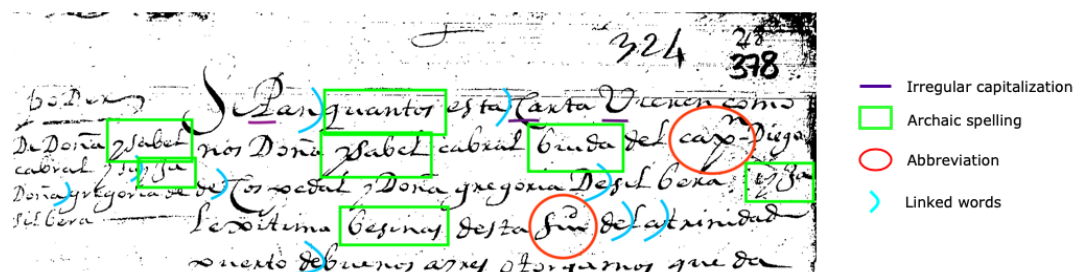


Figure 2. Example of some of the challenges present in the manuscripts utilized for this study. On average, one page has 24 lines. The variations showed in this short paragraph are standard throughout the deeds.

The challenges of converting Spanish American historical texts into machine-readable text have been pointed out by Hannah Alpert-Abrams. She used Ocular, the OCR tool developed by Taylor Berg-Kirkpatrick, to machine read the

Primeros Libros, a sixteenth-century, printed, bilingual (Spanish and Nahuatl) text [Alpert-Abrams 2016]. Her experiments delivered not only dirty OCR but also linguistic errors that could lead to inaccuracies in cultural translation. We experimented with Transkribus, another popular tool, which also yielded an inaccurate output for our collection of seventeenth-century, handwritten, Spanish American notarial records. However, in recent years, deep learning has achieved remarkable success for image understanding and classification, image segmentation, speech recognition, and natural language processing. In this work, we explore if deep learning, and more specifically CNNs, can enable accurate recognition of characters in Spanish American notarial scripts.

Deep learning techniques perform better and achieve higher accuracy when large labeled datasets are available (e.g., ImageNet, MS COCO) [Deng et al. 2009] [Lin et al. 2014]. However, there is no readily available labeled dataset for seventeenth-century historical texts. Although data labeling is an expensive, error-prone, and time-consuming task, with the help of professional historians and paleography experts, we manually prepared a labeled dataset. Additionally, historical texts have various kinds of noises and degradation and the uneven scanning quality of the images pose additional challenges for image preprocessing and cleaning as it has to be performed without losing any of the essential features that define each of the characters.

In this paper, we present an empirical study of how well state-of-the-art CNNs perform for the task of recognizing handwritten characters in seventeenth-century Spanish American notarial scripts. To the best of our knowledge, this is the first effort towards automatically recognizing characters in seventeenth-century Spanish American notarial scripts.

The key contributions of our work are the following:

- With the assistance of professional historians as well as labelers proficient in Spanish and trained in paleography, we prepared the training dataset in two steps. Firstly, we collected from our labelers 250 unique samples of each of the characters present on the manuscripts. Secondly, we augmented the dataset and generated additional characters by applying random distortions and rotations to the original ones. For quality control, before training CNNs on the generated dataset, our professional historians verified that the expanded characters resembled the original ones. There are certain characters that are rare in both, the Spanish language and the notarial scripts. For those characters we had fewer labels.
- We selected four state-of-the-art CNNs, namely, Inception-v3 [Szegedy et al. 2015] [Szegedy et al. 2016], ResNet-50 [He et al. 2016], VGG-16 [Simonyan and Zisserman 2014] and InceptionResNet-v2 [Szegedy et al. 2017] to recognize high frequency characters as well as rare characters (i.e., x and z). We trained these networks by configuring the hyperparameters to achieve the best classification accuracy. Our experiments showed that ResNet-50 achieved the best classification accuracy of 97.08% whereas other networks achieved lower accuracy with VGG-16 being the poorest.
- For broader use by the academic community and to foster new research in transcribing historical texts, our labeled dataset and software are publicly available on GitHub via <https://github.com/UMKC-BigDataLab/DeepLearningSpanishAmerican>.

This paper is organized as follows. Section II discusses the relevant related works and the motivation for this study. Section III presents the methodology we employed and discusses our evaluation results. Finally, we conclude in Section IV and outline our future work.

II. RELATED WORK & MOTIVATION

Deep learning approaches have been widely used for handwritten text recognition of many modern languages. CNNs [Krizhevsky et al. 2012] are among the most popular deep learning methods and have a proven record of outstanding performance when applied to image recognition tasks. Additionally, CNNs have shown an outstanding success when applied to the MNIST dataset [LeCun et al. 1998]. Ashiquzzaman et al. proposed a CNN-based model using ReLU activation function and dropout as a regularization layer that has achieved 97.4% accuracy [Ashiquzzaman and Tushar 2017]. Tsai investigated various convolutional neural network architectures for handwritten Japanese character recognition and created a model with a 96.1% recognition rate for character classification [Tsai 2016]. Another CNN-based model has been proposed by Rabby et al. to classify Bangla handwriting characters. A 95.71% validation accuracy was achieved for the BanglaLekha-Isolated dataset [Rabby et al. 2018].

However, applying these methods to historical documents present unique challenges due to the quality of the scanned images, writing style variations, and the lack of labeled data. Consequently, only a few studies have taken this path. For instance, Kölsch et al. used a Fully Convolutional Neural Network (FCNN)-based approach for historic German documents, which achieved 95.6% accuracy [Kölsch et al. 2018]. Clanuwat et al. proposed a KuroNet model that jointly recognizes an entire page of text by using a residual U-Net architecture and predicts the location and identity of all characters on a given page. Additionally, their proposed system was able to successfully recognize a significant fraction of pre-modern Japanese documents [Clanuwat].

Researchers tend to combine CNNs with recurrent neural networks (RNNs) to further improve accuracy. That was the case for Granell et al. who proposed a handwritten text recognition system to transcribe a corpus of Spanish medieval scrips based on a CNN and RNN [Granell et al. 2018]. The authors showed that deep learning approaches outperform the traditional machine learning models such as Hidden Markov Model-based systems. Dona Valy et al. evaluated different deep learning approaches for character recognition that have been constructed from Khmer palm leaf manuscripts [Valy et al. 2018]. The authors showed that the combination of CNN and RNN-based architectures achieves better results with a 5.01% error rate. Finally, Chammas et al. presented a CRNN system for text- line recognition of historical documents [Chammas et al. 2018]. They showed how to train the system with only 10% manually labeled text-line data from the READ 2017 dataset.

Next, we describe the salient features of four recent CNNs that we used in our study.

1. VGG

In 2014, Karen Simonyan and Andrew Zisserman (2014) proposed the VGGNet for the Large Scale Visual Recognition Challenge (ILSVRC2014). The key contribution from this model was to increase the depth of the architecture by using a 3x3 convolutional filters to achieve higher performance. The VGG model achieved 92.7% top-5 accuracy on the ImageNet dataset and won the ILSVRC2014 challenge. For our experimental study, we chose VGG-16 as a representative of the VGGNet due to its smaller number of parameters compared to VGG19.

2. Inception

Inception architecture was first proposed in 2014 by Szegedy et al. The authors claimed that deeper networks are more prone to overfitting and consume computational resources. They solved that challenge by moving from fully connected to sparsely connected architectures. They introduced the inception layer, which is a combination of three different convolutional layers (1x1 convolutional layer, 3x3 convolutional layer, and 5x5 convolutional layer) with a max pooling layer that operates at the same level. Their outputs are concatenated to be the input of the next layer. This architecture has been updated to increase the accuracy further and proved that any convolution with kernel size more substantial than 3x3 could be represented efficiently with a series of smaller convolutions. In our experimental study, we used Inception- V3 [Szegedy et al. 2015] [Szegedy et al. 2017].

3. ResNet

He et al. introduced the deep residual neural network (ResNet) architecture and won the first place in the ILSVRC 2015 classification competition. ResNet introduces the idea of identity connections that skip one or more layers to train deeper neural networks. This resolved the vanishing gradient problem by allowing the gradients to flow directly through the skipped connections backward from later layers to the initial filter. For our experimental study, we used ResNet-50 as a representative [He et al. 2016].

4. InceptionResNet

Inception-ResNet is a convolutional neural network proposed by Szegedy et al. in 2016. It was trained on more than one million images from the ImageNet database and achieved a 3.08% top-5 error on the test set of the ImageNet classification (CLS) challenge [Szegedy et al. 2017]. The success of residual connections in training very deep architectures and the performance of the Inception-V3 inspired the authors to replace the Inception filter concatenation step with residual connections. This combination allows Inception to obtain all the advantages of the residual approach but with the preservation of its computational efficiency. We used InceptionResNet-v2 as a representative for our experimental study.

Despite these advances, to this day, there is a lack of end-to-end systems capable of managing and analyzing historical documents in general and those in Spanish in particular. This gap, coupled with the professional training needs of 21st century humanities scholars, draw our attention and drives our experimentation efforts to make these manuscripts accessible to a broader community of researchers without requiring extensive and expensive paleography training. Our effort is the first step in this direction and will open up a wide range of research opportunities for others in the academic community.

III. METHODOLOGY

In this section, we present the methodology for conducting this empirical study. The overall steps are illustrated in Figure 3. There are four main stages: (a) pre-processing, (b) dataset preparation, (c) training and validation (to tune the hyperparameters), and (d) testing the accuracy of character recognition.

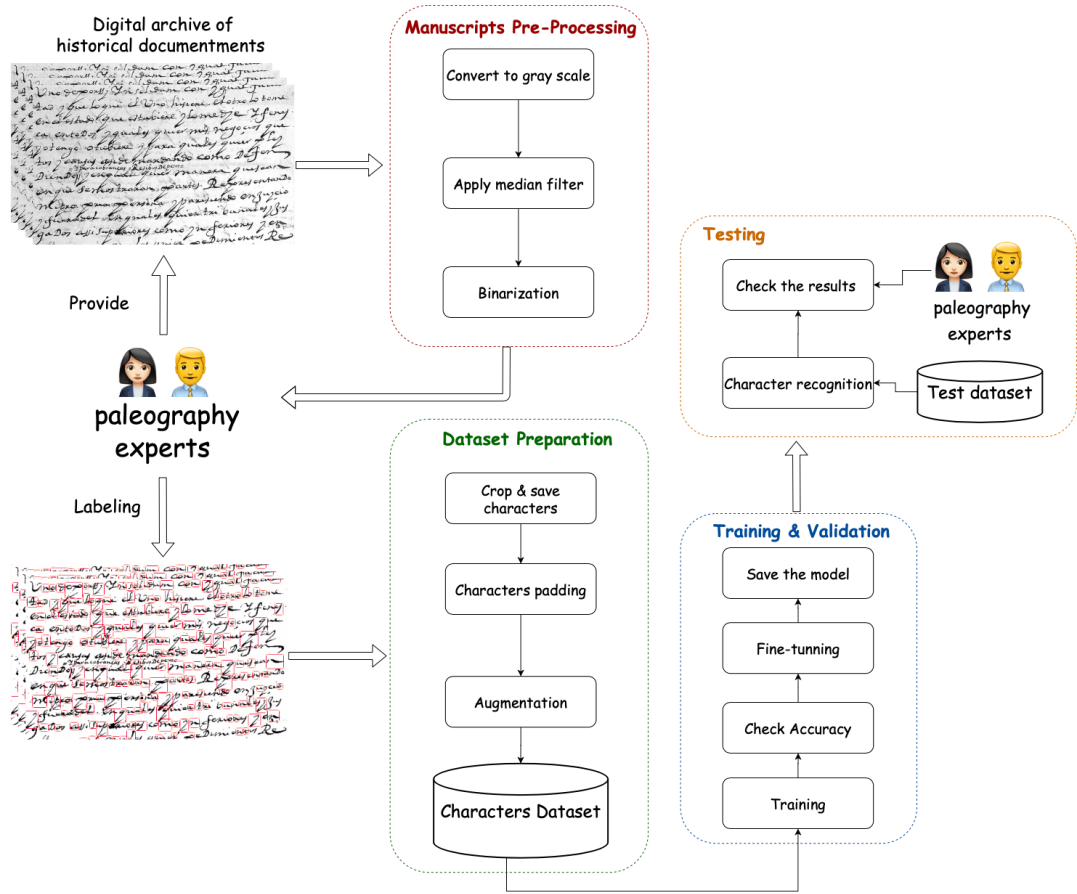


Figure 3. Overall steps involved in character recognition of seventeenth-century Spanish American notarial scripts

a) Pre-processing:

The manuscripts scans contained noise including spurious ink markings, ink smudges and bleeds. Such noise affects the feature extraction as well as the classification. Before constructing the character dataset, we reduced the noise on the manuscripts images (Figure 4). The following preprocessing techniques allowed us to clean the images without affecting the quality of the written content. Firstly, we converted the original manuscript images into grayscale. Secondly, we applied a median filter to soften the images backgrounds and remove background noise. Finally, we applied image binarization to convert the grayscale images into black and white ones as this technique significantly reduces the information contained in an image and increases the training speed [Chamchong and Fung 2009].

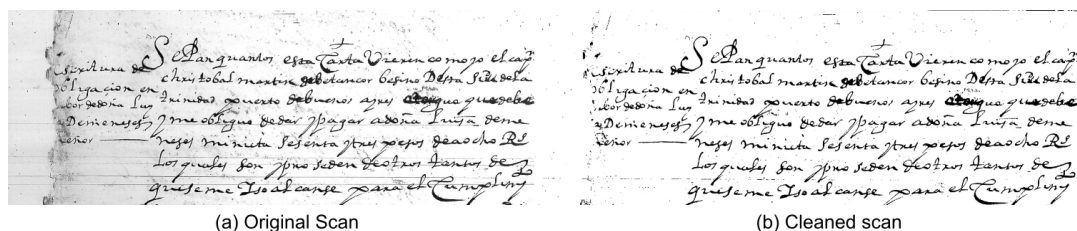


Figure 4. An example of an original scan and its cleaned version.

b) Dataset preparation:

We constructed the character dataset from clean images. Colabeler tool was used to annotate and label the characters. The annotations were exported in JSON format along with each character label and its corresponding coordinates. We ran a Python script to crop and save every character in .png format. As it was difficult to keep each character within square dimensions while annotating them, each one of them was padded with white pixels and resized to fixed dimensions for training purposes. We considered 24 characters that comprise most of the Spanish alphabet: a, b, c, c, d, e, f, g, h, i, j, l, m, n, o, p, q, r, s, t, u/v, x, y, and z. Note that the "u" and the "v" were interchangeable in seventeenth-century Spanish, and thus, the alphabets provided by most paleography manuals list them as a single character. We followed this standard and treated them as single letter (u/v). Additionally, characters such as "k" and "w" are infrequently used in modern Spanish and were so uncommon in seventeenth-century notarial scripts that paleography textbooks do not even list them on their sample alphabets.^[4] To build our sample dataset, we selected the hand of Nicolas de Valdibia y Brizuela who, by 1650, acted as an interim notary in Buenos Aires, Argentina. As opposed to those who held permanent positions, interim notaries did not receive extensive training nor were they skilled scribes. Thus, the experiments we report in this paper were based on very irregular and, therefore, hard to read scripts, as shown in Figure 2 & Figure 4.

24

Our labeling team labeled 250 unique samples for every character. This resulted in a total of 6,000 original images. As deep learning models perform well on large labeled datasets, the dataset was augmented by applying random distortion and rotation of +5 and -5 degrees without affecting the shape and/or the direction of each character. A few examples are shown in Figure 5.

25

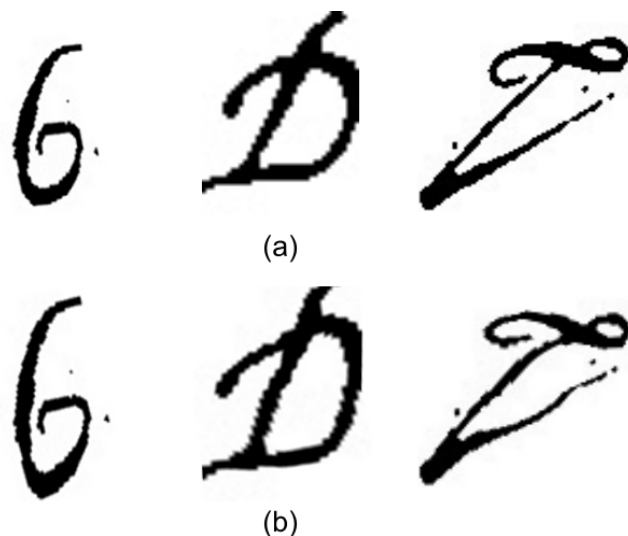


Figure 5. Example of the original characters "b", "d" and "p" b) Example of augmented characters after applying random distortion and rotation

For quality control, before training CNNs on the generated dataset, the professional historians and paleography experts in our team verified that the expanded characters resembled the original ones. Our dataset contained 1,000 samples for each character out of which 250 samples were manually labeled, and 750 samples were generated. This resulted in a total of 24,000 images.

26

c) Training and validation:

We conducted all our experiments on a GeForce RTX 2080 Ti GPU with 12GB GPU memory. Our software was implemented using Keras [Chollet et al. 2015] with TensorFlow backend [Abadi et al. 2015], TensorBoard,^[5] and OpenCV [Bradski 2000].

We trained the state-of-art CNNs using 24,000 images (1,000 images per character) in our labeled dataset prepared from our seventeenth-century Spanish American notary scripts. For each character, we split the training set into an 80-20 split. The samples not used for training were part of the testing set. The testing set contained 50 samples per character and were preprocessed in the same way as the training images. Data augmentation was applied to the training set to avoid overfitting. As we mentioned earlier, most of the paleography manuals list the “u” and the “v” as a single character as they were interchangeable in seventeenth-century Spanish. We followed the same standard and treated them as a single letter (u/v).

e) Test of the accuracy of character recognition: in this research, the character recognition accuracy was used as the primary metric to evaluate the performance of different CNNs. An empirical tuning approach has been followed to tune the hyperparameters to obtain higher character recognition accuracy. To ensure a fair comparison, we set a total of 150 epochs to train the models, and used the Adam optimizer [Kingma and Ba 2015]. Adam is a gradient descent optimization algorithm that is popularly used in training deep learning models. (Using gradient descent, it is possible to find local minima of functions during optimization.) The performance of the models with different hyperparameters values is shown in Table I and Table II.

Table I shows the recognition accuracy when we set the dropout value to 0.6. ResNet-50 achieved the highest accuracy of 97.02%, followed by InceptionResnet-v2, Inception-v3, and VGG-16 with a recognition accuracy of 96.33%, 93.83%, and 96.33%, respectively. The performance of most of the networks improved when the batch size was increased to 64 except for Inception-v3, which achieved a better recognition accuracy when the batch size was 32.

The recognition accuracy results obtained from using 0.5 dropout rate are presented in Table II. The recognition accuracy decreased for most of the networks when we changed the dropout rate from 0.6 to 0.5. VGG-16 was the only model where it performed better on a 0.5 dropout rate.

Models	Batch Size 32	Batch Size 64
VGG-16	62.50%	69.33%
Inception-V3	94.91%	93.83%
ResNet-50	95.50%	97.08%
InceptionResnet-V2	96.00%	96.33%

Table 1. Recognition accuracy obtained with 0.6 dropout rate; best accuracy is shown in blue and the worst in red

Models	Batch Size 32	Batch Size 64
VGG-16	70.33%	70.91%
Inception-V3	93.83%	96.33%
ResNet-50	96.92%	87.58%
InceptionResnet-V2	96.66%	94.08%

Table 2. Recognition accuracy obtained with 0.5 dropout rate; best accuracy is shown in blue and the worst in red

Table III shows the accuracy breakdown of each character obtained from two different experiments. We labeled the best results in bold and worse results in italics. As shown in the table, VGG-16 fails in recognizing non-confusing characters such as “o” and “u/v”. However, it performs well on few characters such as “m” and “y”.

	Batch Size = 32 & Dropout rate = 0.5				Batch Size = 64 & Dropout rate = 0.5			
	VGG	Inception	ResNet	Inception	VGG	Inception	ResNet	Inception
	16	v3	50	Resnet v2	16	v3	50	Resnet v2
A	82%	94%	94%	96%	92%	94%	90%	92%
B	82%	100%	100%	96%	86%	98%	98%	92%
C	62%	92%	94%	94%	42%	94%	86%	92%
c _s	76%	100%	98%	100%	74%	100%	98%	98%
D	74%	92%	96%	96%	74%	94%	76%	96%
E	36%	84%	90%	84%	46%	90%	76%	86%
F	38%	86%	88%	92%	60%	92%	92%	94%
G	22%	94%	96%	98%	32%	88%	68%	68%
H	74%	98%	100%	98%	68%	100%	78%	100%
I	64%	100%	96%	100%	64%	96%	44%	100%
J	88%	100%	98%	96%	86%	100%	96%	100%
L	76%	100%	100%	100%	70%	100%	80%	100%
M	98%	100%	100%	96%	96%	100%	98%	100%
N	90%	94%	96%	96%	76%	92%	56%	78%
O	82%	100%	100%	100%	82%	100%	100%	100%
P	88%	98%	98%	100%	94%	98%	98%	98%
Q	74%	100%	100%	100%	78%	100%	100%	98%
R	58%	92%	96%	98%	44%	94%	96%	90%
S	74%	94%	94%	92%	58%	96%	94%	90%
T	64%	96%	100%	100%	82%	96%	94%	96%
u/v	68%	100%	100%	100%	64%	100%	100%	100%
X	52%	62%	92%	90%	66%	90%	84%	90%
Y	94%	100%	100%	100%	88%	100%	100%	100%
Z	72%	100%	100%	98%	80%	100%	100%	100%

Table 3. Recognition accuracy per character. Best results are highlighted in bold and the worst results are highlighted in italics

Overall, VGG-16 performed the worst for most of our character datasets. Figure 6 gives the graphs of accuracy and loss values for the training set with respect to the number of epochs. It shows that VGG-16 accuracy could have been improved if it trained with more epochs.

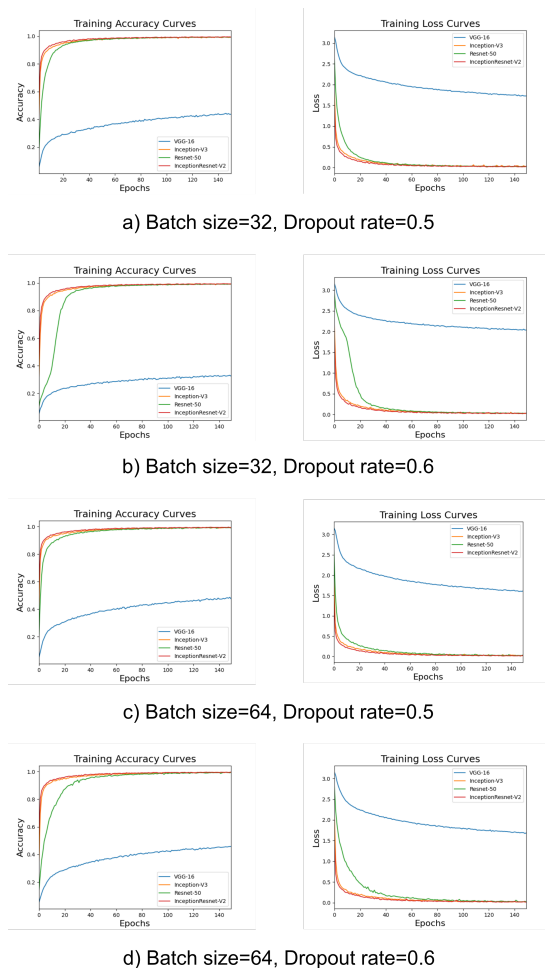
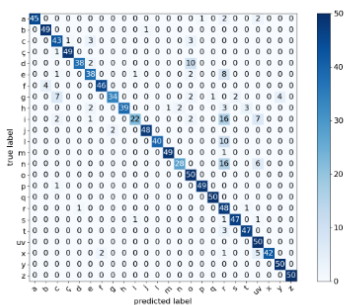


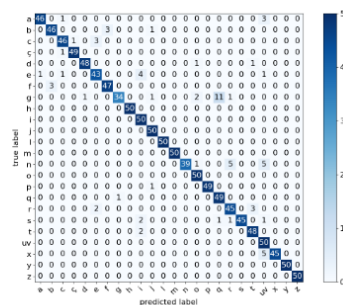
Figure 6. The accuracy (left) and loss (right) curves on the training set of the CNN models

To further understand why some models achieved low accuracy, we generated confusion matrices for all the models. Confusion matrices help us study the miss-classified characters. As seen in Figure 7, confusion matrices confirm that most of the confusions occur between the characters that are written similarly. For instance, the character “n” is confused with “r” as shown in Figure 7(a), and “g” is confused with “q” as shown in Figure 7(b). The results are not surprising as these characters generally confuse non-expert human readers and, occasionally, trained paleographers. Figure 8 shows samples of these characters. However, as shown on Table III, the recognition accuracy remains overall strong.

34



(a) ResNet-50 Confusion Matrix
(batch size=64, dropout=0.5)



(b) InceptionReset-v2 Confusion Matrix
(batch size=64, dropout=0.5)

Figure 7. (a) ResNet-50 Confusion Matrix (b) InceptionReset-v2 Confusion Matrix Confusion matrices of selected models to show the miss-classified characters

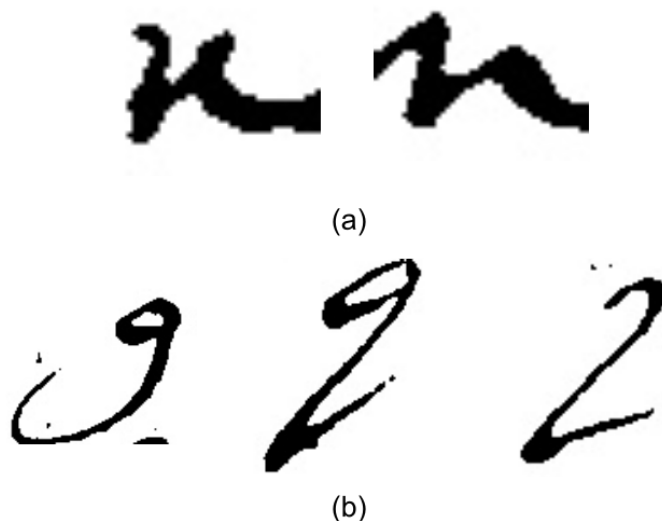


Figure 8. a) Example of the shape similarities between the characters "r" and "n". b) Example of the shape similarities between the characters "g," "q," and "y"

IV. CONCLUSION & FUTURE WORK

Historical handwritten character recognition is a challenging pattern recognition problem due to the inconsistency of the handwritten scripts and the lack of accurate labeled data. In this paper, we presented an empirical study on how state-of-the-art CNNs (developed for image classification) perform for the task of recognizing handwritten characters in seventeenth-century Spanish American notarial scripts. The labeled dataset employed in this study was carefully curated with the help of paleography experts and professional historians. Data augmentation was employed to increase the number of training samples. We observed that ResNet-50 achieved the promising accuracy of 97.08% compared to InceptionResnet-V2, Inception-V3, and VGG-16, which achieved 96.66%, 96.33%, and 70.91%, respectively.

Our study demonstrates that recent CNNs are promising to detect characters in seventeenth-century Spanish notarial scripts. Our future work will test the performance of deep learning-based OCR models such as Keras-OCR, YOLO-OCR, Tesseract and Kraken for the detection and recognition of handwritten words on these manuscripts. Accurate word recognition will be a necessary step in the development of a tool for reading, querying, and analyzing this historical collection. We plan model the content of these manuscripts in a form that would make information retrieval faster and better.

The labeled dataset and software used in this study are publicly available on GitHub via <https://github.com/UMKC-BigDataLab/DeepLearningSpanishAmerican>.

ACKNOWLEDGEMENTS

This research was supported by the NEH Digital Humanities Advancement Grant (HAA-271747-20), UMKC's Missouri Institute for Energy and Defense (MIDE), UMKC's Funding for Excellence Program, and a Collaborative Data Science Grant from UMKC's Institute for Data Education, Analytics and Science (IDEAS). The authors would like to thank Ryan Rowland, Maha Alrasheed, and Vania Todorova for labeling data as well as the *Archivo General de la República Argentina* for granting their permission to use in this study their digitized collection of notary records. Dr. Martin L. E. Wasserman contributed to this project with his expertise in Spanish paleography. The first author (N. A.) would like to thank UMKC's Women's Council Graduate Assistance Fund (GAF) as well as the University of Tabuk in Saudi Arabia for sponsoring her scholarship.

Works Cited

Abadi et al. 2015 Abadi M., Agarwal A., Barham P., Brevdo E., Chen Z., Citro C., Corrado G. S., Davis A., Dean J., Devin M., Ghemawat S., Goodfellow I., Harp A., Irving G., Isard M., Jia Y., Jozefowicz R., Kaiser L., Kudlur M., Levenberg J., Mané D., Monga R., Moore S., Murray D., Olah C., Schuster M., Shlens J., Steiner B., Sutskever I., Talwar K., Tucker P., Vanhoucke V., Vasudevan V., Viégas F., Vinyals O., Warden P., Wattenberg M., Wicke M., Yu Y., and Zheng, X.

"TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available <https://www.tensorflow.org/>

Alpert-Abrams 2016 Alpert-Abrams, H., "Machine Reading the *Primeros Libros*," *DHQ* 10, no. 4, 2016.

Ashiquzzaman and Tushar 2017 Ashiquzzaman A. and Tushar A. K., "Handwritten Arabic numeral recognition using deep learning neural networks," in 2017 IEEE International Conference on Imaging, Vision & Pattern Recognition (icIVPR). IEEE, 2017, pp. 1–4.

Bradski 2000 Bradski G., "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

Chamchong and Fung 2009 Chamchong R. and Fung C. C., "Comparing background elimination approaches for processing of ancient Thai manuscripts on palm leaves," in 2009 International Conference on Machine Learning and Cybernetics, vol. 6. IEEE, 2009, pp. 3436–3441.

Chammas et al. 2018 Chammas E., Mokbel C., and Likforman-Sulem L., "Handwriting recognition of historical documents with few labeled data," in 2018 13th IAPR International Workshop on Document Analysis Systems (DAS). IEEE, 2018, pp. 43–48.

Chollet et al. 2015 Chollet F. et al., "Keras," <https://github.com/fchollet/keras>, 2015.

Clanuwat 2019 Clanuwat T., Lamb A., and. Kitamoto A., "Kuronet: Pre-modern Japanese Kuzushiji character recognition with deep learning," arXiv preprint arXiv:1910.09433, 2019.

Deng et al. 2009 Deng J., Dong W., Socher R., Kai Li L. Li, and Li Fei-Fei, "Imagenet: A large-scale hierarchical image database," in 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248–255.

Granell et al. 2018 Granell E., Chammas E., Likforman-Sulem L., Martínez-Hinarejos C.-D., Mokbel C., and Cîrstea B.-I., "Transcription of Spanish historical handwritten documents with deep neural networks," *Journal of Imaging*, vol. 4, no. 1, p. 15, 2018.

He et al. 2016 He K., Zhang X., Ren S., and Sun J., "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

Kingma and Ba 2015 Kingma, Diederik P. and Ba, Jimmy, "Adam: A Method for Stochastic Optimization," arxiv:1412.6980, published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015.

Kolsch et al. 2018 Kölsch A., Mishra A., Varshneya S., Afzal M.Z., and Liwicki M., "Recognizing challenging handwritten annotations with fully convolutional networks," in 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR). IEEE, 2018, pp. 25–31.

Krizhevsky et al. 2012 Krizhevsky A., Sutskever I., and Hinton G. E., "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

LeCun et al. 1998 LeCun Y., Bottou L., Bengio Y., and Haffner P., "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

Lin et al. 2014 Lin, T.-Y., Maire M., S. Belongie, J. Hays, Perona P., Ramanan D., Dollár P., and. Zitnick C. L., "Microsoft coco: Common objects in context," in Computer Vision – ECCV 2014, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham: Springer International Publishing, 2014, pp. 740–755.

Rabby et al. 2018 Rabby A. S. A., Haque S., Islam S., Abujar S., and Hossain S. A., "Bornonet: Bangla handwritten characters recognition using convolutional neural network," *Procedia computer science*, vol. 143, pp. 528– 535, 2018.

Silva Prada 2001 Silva Prada, N. "Paleografías americanas," <https://www.openedition.org/21549>, 2001.r

Simonyan and Zisserman 2014 Simonyan K. and Zisserman A., "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

Szegedy et al. 2015 Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., and Rabinovich A., "Going deeper with convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.

Szegedy et al. 2016 Szegedy C., Vanhoucke V., Ioffe S., Shlens J., and. Wojna Z., "Rethinking the inception architecture for computer vision," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2818–2826.

Szegedy et al. 2017 Szegedy C., Ioffe S., Vanhoucke V., and Alemi A. A., "Inception-v4, inception-resnet and the impact of residual connections on learning," in Thirty-first AAAI conference on artificial intelligence, 2017.

Tsai 2016 Tsai C., "Recognizing handwritten Japanese characters using deep convolutional neural networks," 2016.

Valy et al. 2018 Valy D., Verleysen M., Chhun S., and Burie J.-C., "Character and text recognition of Khmer historical palm leaf manuscripts," in 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR). IEEE, 2018, pp. 13–

18.

Vellingiriraj et al. 2016 Vellingiriraj E., Balamurugan M., and Balasubramanie P., "Information extraction and text mining of ancient vattezhuthu characters in historical documents using image zoning," in 2016 International Conference on Asian Language Processing (IALP). IEEE, 2016, pp. 37–40.

Wasserman 2018 Wasserman, M.L.E., "La escritura paleográfica Iberoamericana: letras procesales y encadenadas," in *Introducción a la Paleografía. Herramientas para la Lectura y Análisis de Documentos Antiguos*, ed. Rosana Vassallo (La Plata: Facultad de Humanidades y Ciencias de la Educación, 2018)



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.