Text Extraction and Script Completion in Images of Arabic Script-Based Calligraphy: A Thesis Proposal

Dilara Zeynep Gürer & Ümit Atlamaz & Şaziye Betül Özateş

Boğaziçi University
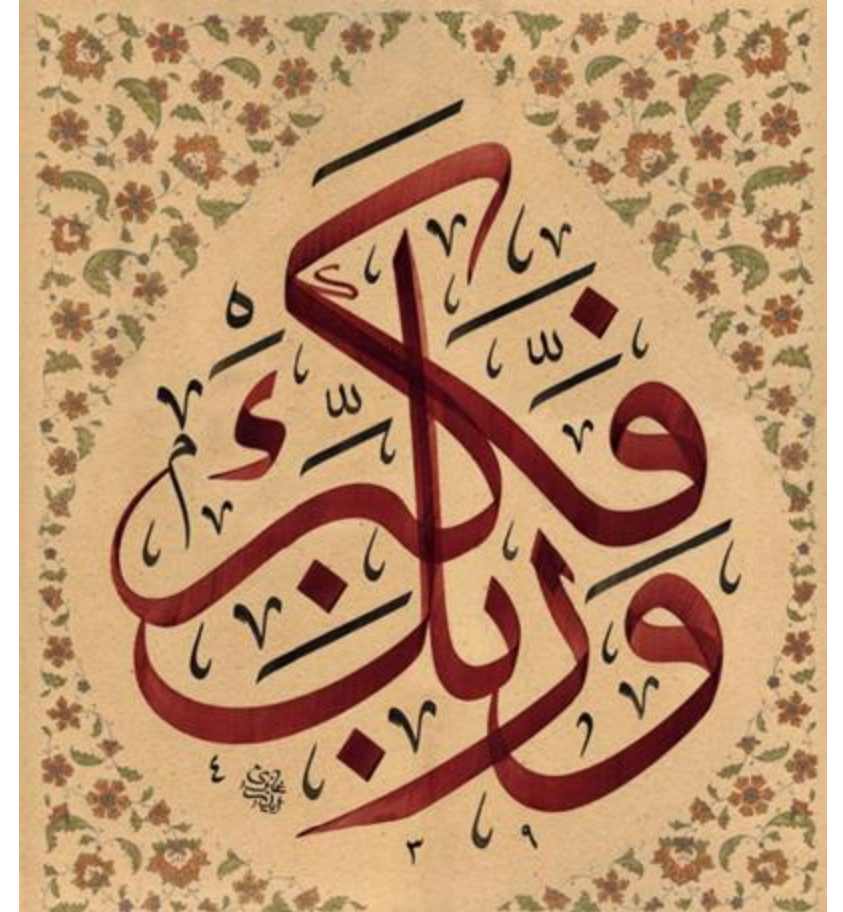İstanbul,Türkiye

Presenter
Dilara Zeynep Gürer

1

Outline

1. Introduction

2. Research Goals and Questions

3. Conclusion

1. Introduction

وَرَبَّكَ فَكَبِّرْ

- Calligraphy is the art of artistic, expressive writing

- Found on historical buildings, and Islamic manuscripts
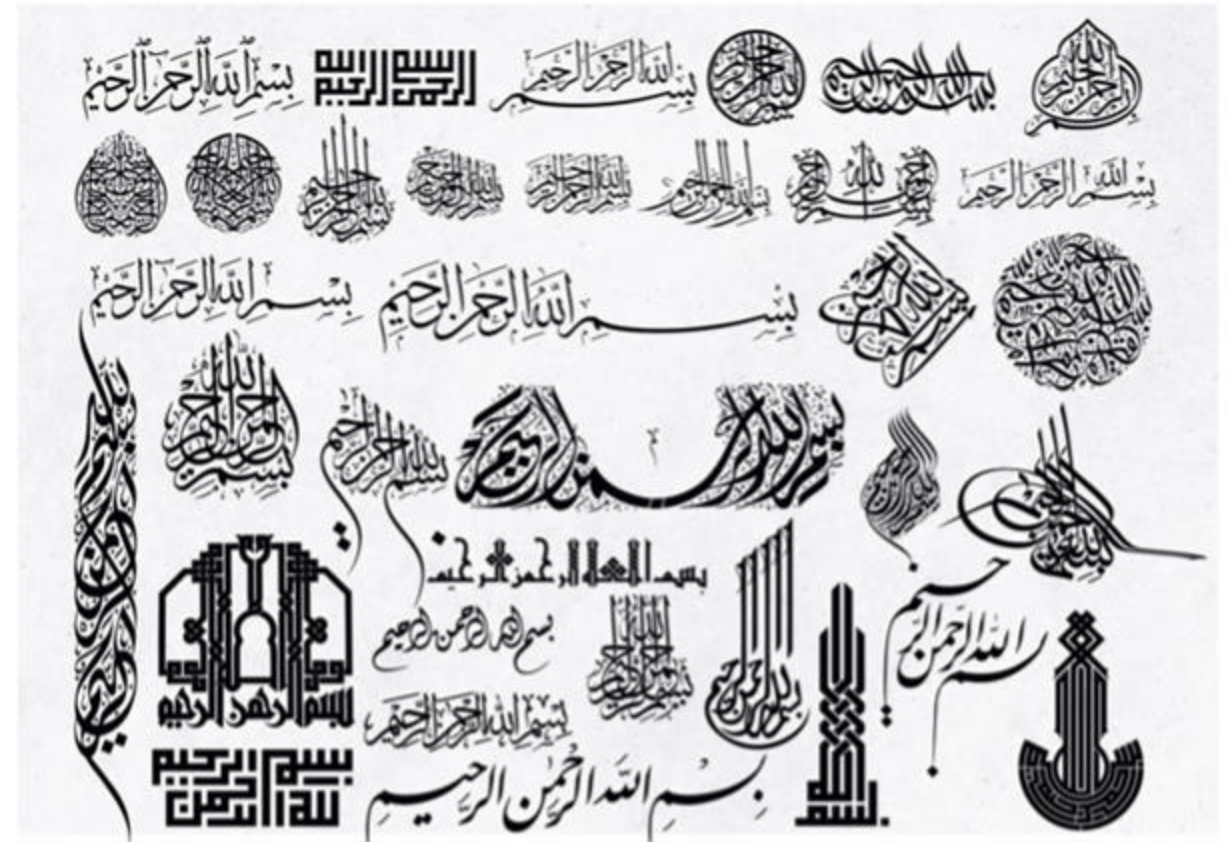
- Conveys Islamic thought and cultural history


(Aydın Kızılyar and Berna Karabulut)

2. Problem

- Traditional OCR struggles with overlapping and stylized text.

Variants of بِسْمِ ٱللَّهِ ٱلرَّحْمَٰنِ ٱلرَّحِيمِ
(Bismillahirrahmanirrahim)

4

3. Literature

- Several studies focus on letter and style classification in Arabic calligraphy
  (Torres Aguilar, 2024; Alyafeai et al., 2022; Kaoudja et al., 2021)

- Only one study tackles text extraction, with limited results due to a small dataset.
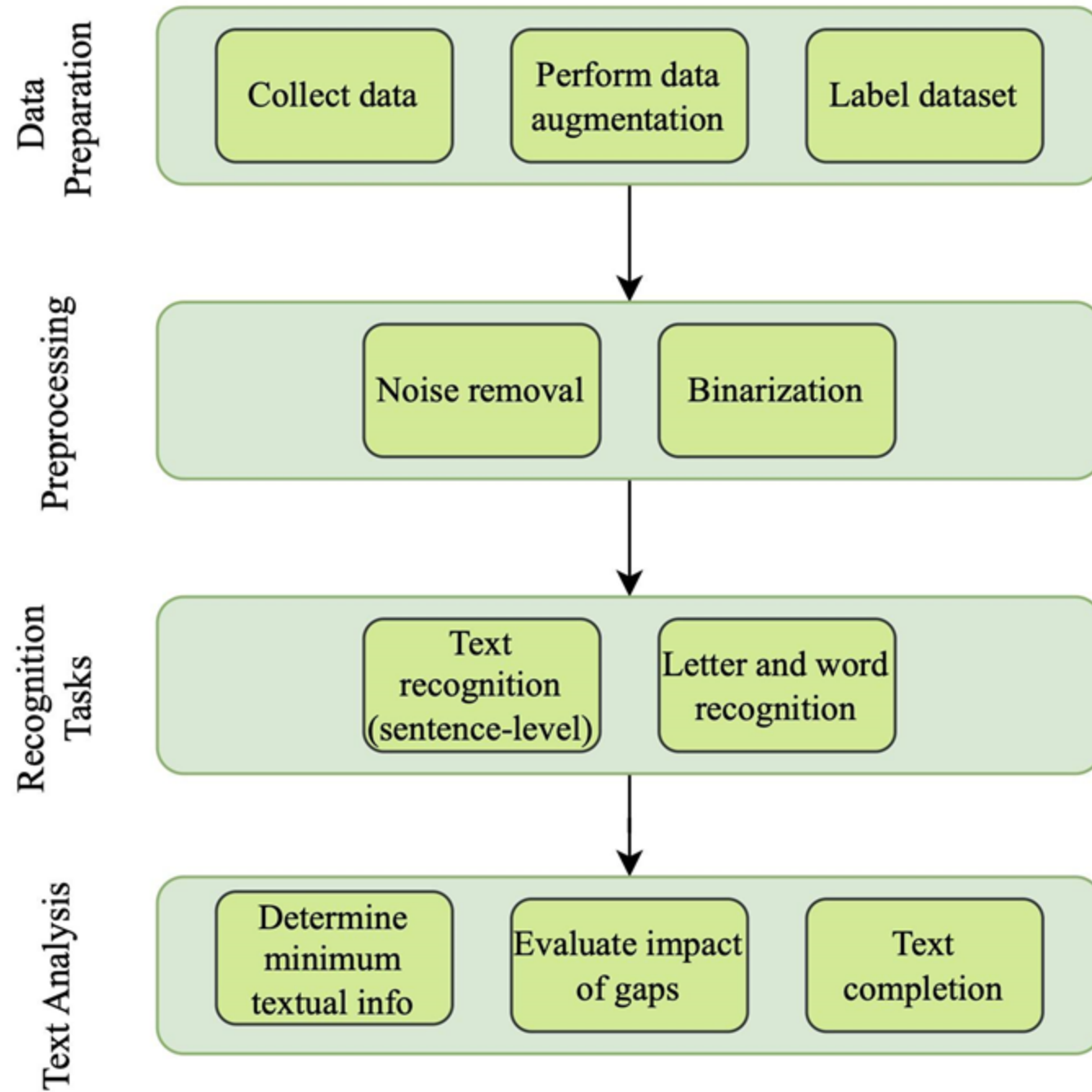 (Alsalamah, 2020)
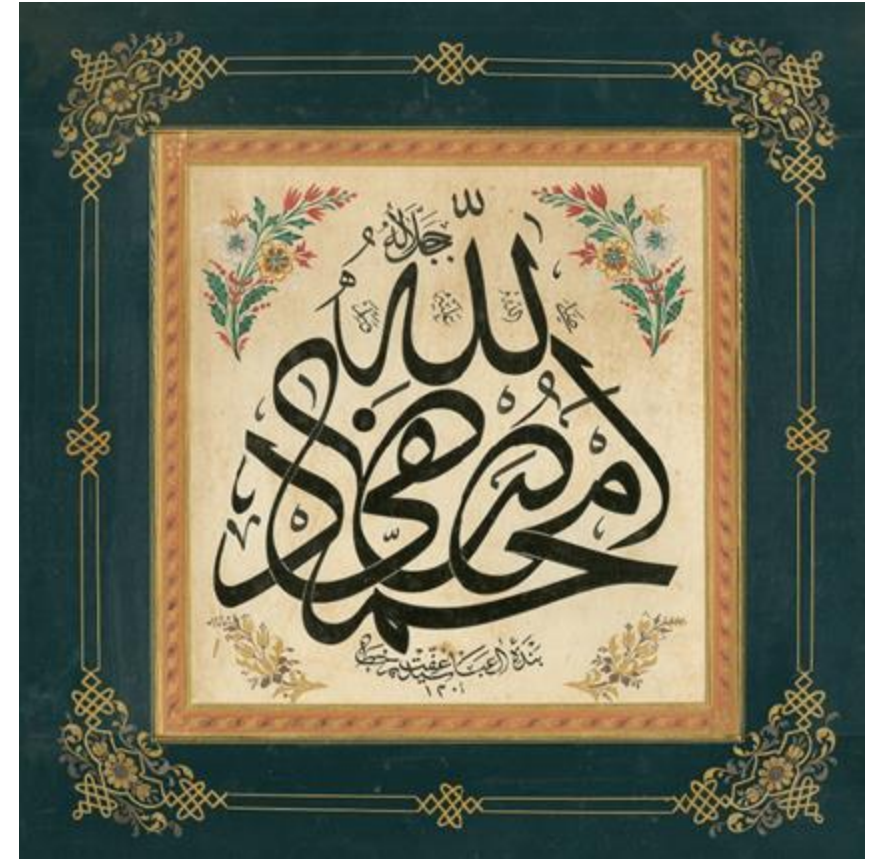
(Abdullah İffet)

4. Research Goals and Questions

Goal

• Optimal methods for accurate extraction and reconstruction

Data Preparation
- Collect data
- Perform data augmentation
- Label dataset

Preprocessing
- Noise removal
- Binarization

Recognition Tasks
- Text recognition (sentence-level)
- Letter and word recognition

Text Analysis
- Determine minimum textual info
- Evaluate impact of gaps
- Text completion

RQ1 How do we obtain authentic data?

- Web & On-site Image Collection

- Arabic & Ottoman Turkish Focus

- Persian & Urdu Expansion

- 136K-Page Archive Utilization
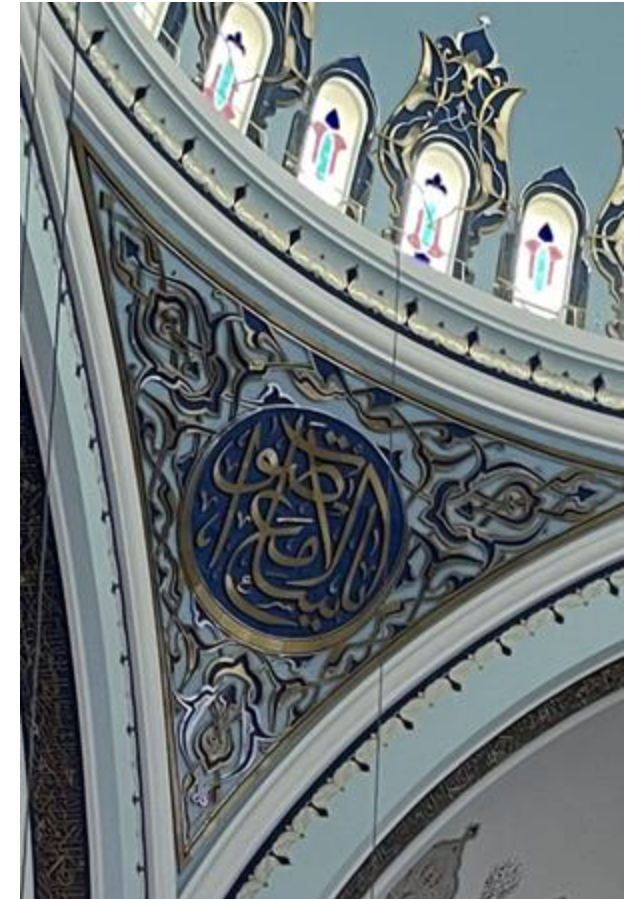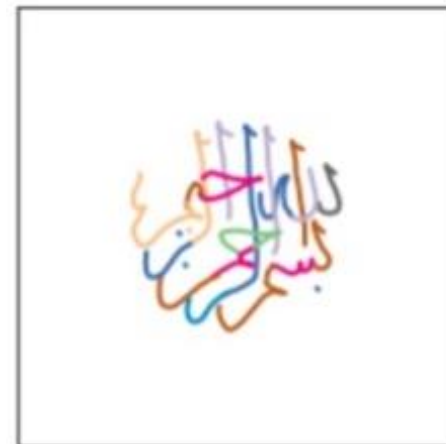
RQ2 How to label the data?

- **Automated Labeling**
  Web-sourced data

- **Manual Labeling**
  Physical sources

- **Image-Text Dataset**

- **Online Dataset Training**
  Guide offline text labeling

- **Semi-Supervised Learning:**
  Expand dataset with pseudo-labeling

RQ3 How to ensure a balanced distribution of various styles in the dataset?

- **Artistic Variation**:

  Simulate diverse calligraphy styles

- **Data Augmentation**:

  Rotation, scaling, flipping and shearing

- **Dataset Enrichment**:

  Structured variations for robustness

RQ4 How to effectively remove noise from the images?
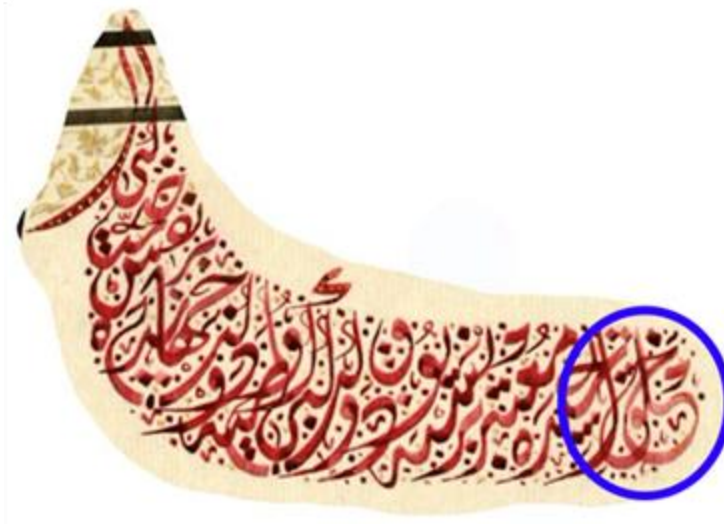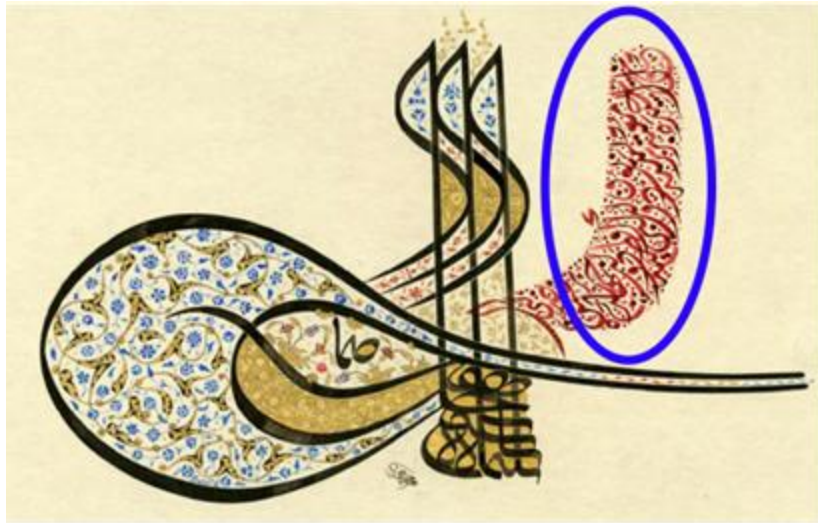
- **Template matching**

- **Morphological operations**

- **ROI detection**

RQ5 Which recognition method is most effective for analyzing the text?

- **Complexity**: Handling overlaps & curvature in calligraphy

- **Approach**: Character, word, and sentence-level recognition

- **Metrics**: CER, WER, Levenshtein Distance

- **Comparison**: Baseline OCR & Transformer-based models

RQ6 Is preprocessing necessary?

1. **Fine-tune the language component of VLMs**
   - e.g. BLIP-2 (Li et al., 2023) , LLaVA (Liu et al., 2023)
   - using our textual archive

2. **Further fine-tune**
   - using our image-text dataset

3. **Use targeted prompts with VQA models on original images**
   - e.g. "focus on the text in the given image"

RQ7 What is the minimum required information to understand the content of the images?

Conclusion

- **Goal**
  - Prepare ready-to-use materials for Arabic calligraphy text recognition

- **Method**
  - AI architectures

- **Challenge**
  - Complex structure and artistic variation in Arabic calligraphy

- **Contribution**
  - Real-world dataset
  - End-to-end extraction pipeline

Thank You

References

- Sergio Torres Aguilar. 2024. Handwritten text recognition for historical documents using visual language models and GANs. Hal-04716654v2.

- Seetah Alsalamah. 2020. Combining Image and Text Processing for the Computational Reading of Arabic Calligraphy. Ph.D. thesis, The University of Manchester.

- Seetah AlSalamah and Ross King. 2018. Towards the machine reading of Arabic calligraphy: A letters dataset and corresponding corpus of text. In 2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR), pages 19–23.

- Z. Alyafeai, M. S. Al-Shaibani, M. Ghaleb, et al. 2022. Calliar: an online handwritten dataset for Arabic calligraphy. Neural Computing and Applications, 34:20701–20713.

- Gagan Bhatia, El Moatez Billah Nagoudi, Fakhraddin Alwajih, and Muhammad Abdul-Mageed. 2024. Qalam: A multimodal LLM for Arabic optical character and handwriting recognition. In Proceedings of The Second Arabic Natural Language Processing Conference, pages 210–224, Bangkok, Thailand. Association for Computational Linguistics.

- Chaouki Boufenar, Adlen Kerboua, and Mohamed Batouche. 2018. Investigation on deep learning for offline handwritten Arabic character recognition. Cognitive Systems Research, 50:180–195.

References

- M. U. Derman. 1997. Hat. [Accessed: 28 November 2024].
- H. Gündüz. 1988. Türk hat sanatında Ṣeyh Hamdullah ve Ahmed Karahisari ekolleri. Master's thesis, Mimar Sinan Fine Arts University.
- Dilara Zeynep Gürer and İnci Zaim Gökbay. 2024. Arabic calligraphy images analysis with transfer learning. Electrica, 24(1):201–209.
- N. A. Jebril, H. R. Al-Zoubi, and Q. Abu Al-Haija. 2018. Recognition of handwritten Arabic characters using histograms of oriented gradient (hog). Pattern Recognition and Image Analysis, 28:321–345.
- Zineb Kaoudja, Mohammed Lamine Kherfi, and Belal Khaldi. 2021. A new computational method for Arabic calligraphy style representation and classification. Applied Sciences, 11(11). Wakana Kuwata, Ryota Mibayashi, Masanori Tani, and Hiroaki Ohshima. 2024. Glyph generation for Japanese calligraphy based on encoding both content and style. In 2024 IEEE International Conference on Big Data and Smart Computing (BigComp), pages 207–214.
- B. H. Nayef, S. N. H. S. Abdullah, R. Sulaiman, et al. 2022. Optimized leaky relu for handwritten Arabic character recognition using convolution neural networks. Multimedia Tools and Applications, 81:2065– 2094.

References

- Arshia Sobhan, Philippe Pasquier, and Adam Tindale. 2024. Unveiling new artistic dimensions in calligraphic Arabic script with generative adversarial networks. Proc. ACM Comput. Graph. Interact. Tech., 7(4).
- Yuanbo Wen and Juan Alberto Sigüenza. 2020. Chinese calligraphy: Character style recognition based on full-page document. In Proceedings of the 2019 8th International Conference on Computing and Pattern Recognition, ICCPR '19, page 390–394, New York, NY, USA. Association for Computing Machinery.
- Adam Wong, Joseph So, and Zhi Ting Billy Ng. 2024. Developing a web application for Chinese calligraphy learners using convolutional neural network and scale invariant feature transform. Computers and Education: Artificial Intelligence, 6:100200.
- Yuqing Zhang, Baoyi He, Yihan Chen, Hangqi Li, Han Yue, Shengyu Zhang, Huaiyong Dou, Junchi Yan, Zemin Liu, Yongquan Zhang, et al. 2024. Philogpt: A philology-oriented large language model for ancient Chinese manuscripts with dunhuang as case study. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, pages 2784–2801.