Honors Project Proposal Form (rev FA24)

Name: Merrick Shorter

Major: Computer Science

Project advisor: Dr. Nishatul Majid

Research question: Can we noninvasively process large amounts of historical documents efficiently to aid in the restoration of lost knowledge of our past?

Goal/purpose: Using artificial intelligence to transcribe words has been a hot topic for the past two decades. I want to uphold an example for others to help the human race learn more about their history through the transcription of historical documents including birth, marriage, and death documents. Using one of the latest offline handwriting recognition artificial intelligence model, YOLOv10 we hope to retrieve the genealogy data from historical documents in many different languages to form a dataset and enable quick searching through services like familysearch.org and ancestry.com. This can connect families and expand our knowledge of genealogy and human history in a manner where IRB approval is not needed. This is a useful and harmless use of artificial intelligence that expands upon existing literature that relieves humans of heavy and sometimes impossible labor while not affecting the human experience and jeopardizing privacy.

Rationale: A large amount of written human history has been lost to us collecting dust either because it requires too much manpower and time to transcribe or because it is too faded or tattered for a human to read. Artificial intelligence's task diversity and efficiency have rapidly improved to the point where we can transcribe historical documents to help the field of genealogy (Martinez et al., 2024). This task can be done by first creating a You Only Look Once (YOLO) neural network (Tervan et al., 2023, Martinez et al.,2024) which takes greyscale pixel values as inputs, passes these values through a series of hidden layers with adaptive weights which return a value corresponding to a character. Throughout the training epochs (number of complete passes of the dataset) the weights in the hidden layer get adjusted until the network can consistently get the correct answer. Then we conduct some image preprocessing using OpenCV (Xie & Lu, 2013, Deng et al., 2020). This trains the neural network to expect data in certain spots using bounding boxes created in Roboflow (Pellegrino et al., 2024, Martinez et al. 2024), a service we will use to label data in our dataset for accuracy testing latter. Preprocessing is also where you can address blemished data (Li et al., 2023, Islam, 2024). After preprocessing, you train the neural network using the Identity and Access Management (IAM) dataset which contains 13k+ greyscale images of offline handwriting to test and train neural networks (Islam, 2024, Deng et al., 2020). Once the neural network is trained and tested over many epochs, it is time to put it to use using object detection to transcribe offline handwriting (Deng et al., 2020, Graves & Schmidhuber, 2008) to be added to databases of genealogy services such as familysearch.org or ancestry.com. Using our knowledge of offline handwriting recognition and the speed of artificial intelligence models we can fill gaps in family trees and regain the knowledge of our future to learn about ourselves and our lineage.

Project timeline:

Spring 25 – Create the AI using YOLOv10, Label the dataset with Roboflow, conduct image preprocessing using open cv, and train the AI on a small number of epochs.
Summer 25 – Continue to read articles, especially concerning the number of epochs needed to train the AI, and begin drafting the paper**.**
Fall 25 –Have the final AI trained, tested, deployed, and revise the final paper.
Spring 26 – Prepare final lit review

Annotated bibliography:

Article: J. Deng, X. Xuan, W. Wang, Z. Li, H. Yao, and Z. Wang, "A review of research on object detection based on deep learning," Journal of Physics: Conference Series, vol. 1684, no. 1, p. 012028, nov 2020. [Online].

My abstract: Computer vision and object detection have become a hot topic in 20 years. The authors highlight the six steps of object detection: pre-processing, window sliding, feature extraction, feature selection, feature classification, and post-processing. The model training method allows for two types, single-stage and double-stage detection.

Article: J. Terven, D.-M. Cordova-Esparza, and J.-A. Romero-Gonz ´ alez, "A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS," Machine Learning and Knowledge Extraction, vol. 5, no. 4, pp. 1680–1716, 2023. [Online].

My abstract: You Only Look Once (YOLO) has become a central real-time object detection system for robotics, driverless cars, and video monitoring applications. This article is a comprehensive analysis of YOLO's evolution from v1 to v8 and beyond. The focus of this article is on changes in network architecture and training tricks for each model.

Article: G. Xie and W. Lu, "Image Edge Detection Based On Opencv," International Journal of Electronics and Electrical Engineering, vol. 1, pp. 104–106, 2013. [Online].

My abstract: The authors explain the process of determining if there is just enough copper core in a very small copper wire. This task is done by taking a high-resolution picture of the internal structure of the wire, then using OpenCV, they implement image preprocessing, and after morphological opening and closing operations, the exact number of copper cores can be distinguished using contour tracking

Article: A. Graves and J. Schmidhuber, "Offline handwriting recognition with multidimensional recurrent neural networks," in Advances in Neural Information Processing Systems, D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, Eds., vol. 21. Curran Associates, Inc., 2008. [Online].

My abstract: Offline handwriting recognition combines computer vision with sequence learning. These two elements usually handled separately can now be combined in a hidden Markov model neural network. This article introduces a globally trained model that takes raw pixel data as input therefore not relying on alphabet-specific processing. This model performed over two percent better than the competition in the Arabic recognition competition, despite the authors understanding no Arabic.

Article: R. Li, S. Chen, F. Zhao, and X. Qiu, "Text detection model for historical documents using CNN and MSER — IGI Global," Journal of Database Management (JDM), 2023. [Online].

My abstract: The article introduces a solution for documents with illegible characters. This solution is an ATD model including a CNN-based text-box generation network and an NMS-based MSER text-box generation model. The results show much higher accuracy than previous models and this is a strong step forward for the detection of historical documents.

Article: J. Martinez, L. Wortham, N. Majid, and E. Smith, “Extracting From Handwritten Historical Data from Registry Forms Using YOLOv8” Publisher, year. [Online]

My abstract: The authors created an algorithm that if successful, can quickly and accurately recognize and transcribe historical documents using an object detection network called YOLOv8. In the study, the authors tested their approach using birth, marriage, and death documents from Rio Grande do Sul, Barracao, Brazil dated between 1748-1998. The author’s model got within 2% of the highest detection accuracy reported which was 99.42%.

Article: Pellegrino, Rosemarie V., Jethro Hoyt T. Lacuesta, and Carl Ferione L. Dela Cuesta. "Nail Abnormality Identification Using Roboflow and Yolov8." *2024 14th International Conference on Biomedical Engineering and Technology*. 2024.

My Abstract: The article uses a YOLOv8 neural network trained on roboflow data to identify abnormalities in human fingernails including, Terry’s Nail, Spoon Nail, and Splinter Hemorrhage that may be missed by medical professionals. The authors used roboflow to easily add bounding boxes, adjust image-level, and color-level. The detection was done using a Raspberry Pi 4 Model B which is not a very powerful machine. Despite the limited processing power, the model was 95% accurate in classifying a healthy nail and 75-94% accurate in correctly classifying the three unhealthy nails.

Article: Islam, Nazrul. "Handwritten recognition of English crossed-out words using YOLO." (2024).

My Abstract: The author looks at offline handwriting recognition from a different angle. Many of the popularized models can very accurately decipher unobstructed text but fail when a word is crossed out or the paper has blemishes. The author wants to find a solution for this problem so they train a model using YOLOv8 and training data from the IAM dataset to accurately decipher text with six different obstruction styles.