International Collaboration on Handwriting Analysis

Meeting Minutes – Client Meeting Week 4

Location:	Virtual Discord Meeting
Date and Time:	6.30 PM – 7.15 PM, Thursday, 16th of March 2023
Participants:	
Client:	Dr Sophie Constable
Our Team:	Mohammad Saddam Mashuri, Rui Lu, Sage Labios, Xiangda Li, Yang Ding, Neo Han (Yuyang Han), Zhishang Bian
Minute Taker:	Mohammad Saddam Mashuri
Absences:	-

Agenda

1. Overall progress report to the Client

2. Image pre-processing progress and result report

3. Convolutional Neural Network (CNN) progress and result report

4. K-means clustering method progress and result report

Discussion Points and Decisions

1. Xiangda Li reported to the client that the group had tested to pre-process some of the leaflets (grayscaling, skew correction, noise reduction, etc.). A few leaflets gave trouble with the image processing (like J104.jfif) and some imperfections with the pre-processing. Dr Sophie Constable wanted the software to have images resized but characters still the same size and suggested a fix to have some of the leaflets with troublesome sections cropped.

2. Dr Sophie Constable suggested the group meet with Dr Muhammed Hussein for advice about our project, and she will try to arrange a meeting with him. There is no exact date and time at the moment. There is 14 hours difference between Canberra and Germany (e.g., 7 PM in Canberra = 9 AM in Germany).

3. Rui Lu and Neo Han showed the CNN benchmark, and the accuracy is quite terrible on the test dataset. Dr Sophie Constable suggested two solutions for the lack of ground truth issue: split each leaflet into two paragraphs and have the software figure out which paragraphs pair with the original leaflet. Neo Han suggested splitting the dataset to increase the accuracy.

4. Zhishang Bian showed an issue with the K-means method for feature extraction on the image of the leaflets. This issue is that too many features are being measured, which can not be statistically shown in 3D space (4 or more would need a four or more-dimensional space, which is not feasible to view in the real-world – too confusing). He thinks that dimensional reduction would be a good fix for this as there may not be a need to extract too many features to classify the images/leaflets.

5. Dr Sophie Constable raised one more risk in this work. The risk was that this work might be sensitive for the family of the POWs.

Action items	Owner(s)	Deadline	Status
Prepare agenda for Week 4 tutorial	All team members	Thursday Afternoon, 16th of March 2023.	Complete