



Meeting Minutes - @13/06/2022

Chair: Alan, Pham

Secretary: Kelin “Colin” Zhu

Members:

- a1829445, Kelvin, Ting
- a1808891, Pham Gia “Alan” Nghi
- a1805827, Kelin “Colin” Zhu

Time and Place

Meeting on Zoom with Thomas and between team members on the 13th of June, 2022

Purpose

- Show Thomas the poster and receive feedback
- Ask questions on the Final Report

Outcome

- Decided to rewrite the poster with more appropriate content.
- Make the poster to have less text and more visual graphs
- Focus on 3 main key points: Data Generation, Machine Learning and Image Generation

Full notes:

Feedback poster

Project Intro

What does the system does as a whole

What's invloved

Current is not quite clear

Textual description → to a visual depiction

Use machine learning and web based tool to achieve the result

Maybe a 2nd heading

Background - Context

Police receives text description of a criminal via the computer in the police car. This project aims to reduce the cognitive load of the officer by converting text description into an image.

This project uses Machine Learning and Web-based Application to achieve such results

→ **This could be the INTRO**

At least 1 example

Put graphic that show a sentence and the image generated from that (give 2-5 examples if you can)

Proposed Solution Methodology

A bit too much text and arrows, just the core idea

4 components

1. Text input
2. Machine Learning
3. Structure visual/data format (JSON output)
4. 3D rendering

Use arrows to guide

We can combine a bit academic and creative/marketing

May need to talk more about Data Generation

Where do get the training data?

Synthetic Data Generation

Use templates, pre-defined sentences and structured output to generate the data

We embed visual text description from law documents → highlight the output

We could delete Experimental Results, What could be improved?

→ Talk more about the Machine Learning Model

→ Show how the transformer works

List of used software is a bit too big

→ Save space for other parts

Talk more about how we do the 3D generation (model? Color? Render?)

3 main parts

1. Machine Learning Model
2. Data Generation
3. Image Generation

Conclusion

Introduce police in the intro since we conclude with the police.

Visual ID is an innovation

Be bold

This also be used in the fashion industry or any equivalent industry as a whole

The boxes are a big too big

Short and concise text/paragraph

Change the center text (repetition)

Textual description → to a visual depiction

New intro

In the police force, a textual description of a perpetrator is sent to the officers via the computers in their police cars. Often, the description is hidden in amongst the other information about the crime. It is difficult for the officers to quickly glance at the screen and get a mental picture of the perpetrator.

Visual ID is a web-based application that can help the officers to virtualize the perpetrator by converting text to an image. In this work, multiple datasets are created from a manual script that contains 20,000 texts, and 13 appearance categories and each category include 35 different classes. These datasets are used for training an optimized machine learning model (RoBERTa) for natural language tasks that can classify the text into appearance categories and its corresponding classes. Finally, an image generator renders these appearance categories into 3D images.

→ Could be a bit too long.

→ Focus on what it does

NEW NEW INTRO

In the police force, a textual description of a perpetrator is sent to the officers. It is difficult for the officers to quickly glance at the screen and get a mental picture of the perpetrator. Visual ID is a web-based application that can help the officers to virtualise the perpetrator by converting text to an image