

ISDN3150 Week 2 Assignment – Part 4

From Image to Text — Understanding Visual Content with Language Models

Man Chun Hei William

https://github.com/hacker-is-undefeatable/ISDN3150_Week2_Assignment/tree/main

February 2026

1 Objective

The purpose of this assignment was to convert personal portrait photographs into plain-text ASCII art and then test whether a text-only large language model could correctly interpret the result as depicting a human face or upper body.

2 Methodology

2.1 Images Used

Three portrait images were processed:

- `image_1.jpg` (https://github.com/hacker-is-undefeatable/ISDN3150_Week2_Assignment/blob/main/imgs/image_1.jpg?raw=true)
- `image_2.jpg` (https://github.com/hacker-is-undefeatable/ISDN3150_Week2_Assignment/blob/main/imgs/image_2.jpg?raw=true)
- `image_3.jpg` (https://github.com/hacker-is-undefeatable/ISDN3150_Week2_Assignment/blob/main/imgs/image_3.jpg?raw=true)

All images were downloaded from GitHub using the `requests` library and opened with `PIL.Image`.

2.2 ASCII Art Generation Pipeline

The function `mono()` performs the following steps:

1. Convert image to grayscale (if necessary)
2. Resize while preserving aspect ratio
3. Apply **histogram equalization** with `cv2.equalizeHist(resized_img)`
4. Normalize pixel values to the range [0, 255]
5. Map brightness to characters from the ramp.

2.2.1 Role of histogram equalization

After heavy downscaling, facial details (eyes, mouth, hair boundary, nose) are often compressed into a narrow intensity range. `cv2.equalizeHist()` spreads the intensity values more evenly across the full 0–255 range. This increases local contrast and makes important mid-tone features (cheeks, eye sockets, hair strands) more distinguishable in the resulting low-resolution ASCII art.

2.3 Resolution Experiments

I first generated ASCII art at

```
target_h = 50, target_w = 100
```

Later I tried higher resolutions such as

```
target_h = 100, target_w = 200
```

(and several values in between) hoping the increased character density would allow the model to recognize facial structure (symmetry of eyes, nose bridge, mouth, hair outline, shoulders). Unfortunately this did not lead to correct recognition.

3 LLM Analysis Results (gpt-4o-mini)

All three ASCII arts (at 50×100 resolution) were given to gpt-4o-mini with the same prompt asking for:

1. description of what is observed
2. inferred object/subject
3. visible features or characteristics

3.1 Summary of model interpretations

- Image 1: abstract pattern / landscape / water-like or foliage-like structure
- Image 2: stylized landscape / skyline / mountains / natural scene
- Image 3: possible stylized figure / character / animal / mask / humanoid creature (closest, but still uncertain)

None of the responses used words such as “face”, “person”, “portrait”, “human”, “head” or “shoulders”. Increasing the resolution did not change this outcome.

4 Discussion & Conclusion

The main observations are:

1. Histogram equalization noticeably improves the visibility of facial contours and mid-tone details when the ASCII art is viewed by a human in a monospace font.
2. This preprocessing step is not sufficient for the text-only LLM to classify the pattern as a human face/portrait.
3. Going from 50×100 to 100×200 characters adds more structural information but still does not reach the recognition threshold for gpt-4o-mini.
4. Current text-only models appear to lack strong enough priors to decode dense symbol arrangements as faces without additional training or multimodal input.

In summary, even with contrast enhancement and resolution tuning, plain-text ASCII art remains very difficult for text-only language models to interpret as human portraits. This highlights a limitation in transferring fine spatial understanding from the vision modality into the pure text domain using only the classic ASCII character set.

5 Appendix

Here are the output for the program:

5.1 ASCII Art Outputs (target_h=50, target_w=100)

5.1.1 Image 1

5.1.2 Image 2

5.1.3 Image 3

```
1 ==^~^~~~~~ ( CWEWE00$$$$@$$$$##$$EE=/_.‘‘.~-----  
2 =^~^~~~~~ X$EEEEOEEOOOW$O$$$00WCO//[_~~~-----  
3 =^~^~~~~~ _~/([W$$0$WEEEXE$00XEXX#[CXO(?)/.-----
```

5.2 LLM Analysis Outputs (gpt-4o-mini)

5.2.1 Image 1

1. ****Description of the Text-Based Image:****
The ASCII art representation consists mainly of letters, symbols, and various characters arranged in a structured pattern. This arrangement creates a visual texture rather than a clear image. The background contains a mixture of characters that include 'W', 'C', 'O', 'X', and various punctuation marks. There are regions that

appear clearer and denser, while others seem more chaotic in arrangement.

2. ****Type of Object or Subject:****

While the ASCII art is quite abstract and does not depict a recognizable object or subject clearly, the patterns and forms generated by the text suggest a structure, possibly a landscape or abstract representation. The repetition of certain characters like 'W', 'C', and 'O' could imply an environment, perhaps something that represents a complex pattern like water or foliage, although it is not definitive.

3. ****Visible Features or Characteristics:****

- The art showcases various patterns created by repeated letters and symbols, which suggest a layered or textured appearance.
- There are no clear features (like eyes, a mouth, or distinct identifiable shapes) that suggest a face or figure.
- Some areas consist of more dense clusters of characters, which might imply depth or shadows.
- Overall, the chaotic yet systematic nature conveys a feeling of complexity and possibly fluidity or movement, akin to natural elements or abstract geometric forms.

In summary, this ASCII art is quite abstract and difficult to classify as a recognizable form, leaning towards a more textured and complex representation of an environment rather than depicting a specific object or subject.

5.2.2 Image 2

The text-based image representation you provided appears to depict a complex arrangement of ASCII characters that, when viewed collectively, forms an abstract or stylized representation of a scene or subject. Here's a breakdown of the analysis:

1. ****Description of Observations:****

- The image is predominantly made up of various symbols, letters, and punctuation marks arranged in a grid-like pattern. It incorporates specific characters such as '#', '@', 'E', 'C', 'W', 'X', and '\$', indicating they might carry visual weight or significance in the overall design.
- There are certain areas where the arrangement seems denser, likely indicating shadowed or darker regions, while other parts are more sparse, suggesting highlights or lighter areas.

2. ****Inference about the Object/Subject:****

- Given the complexity and the use of specific characters, the ASCII art seems to represent a stylized landscape or possibly an abstract representation of a natural scene. The inclusion of symbols might suggest a fantastical or surreal element within the landscape. However, it's not easily discernible as a specific object like a face or a recognizable animal.

3. ****Identifiable Features/Characteristics:****

- The use of varying characters suggests different textures; for example, the repetition of '#' and '@' might signify solid or structured areas, while 'E' and other letters might provide more fluid lines or softer textures.

- 12 - The overall shape and forms are somewhat reminiscent of skyline
silhouettes or patterns that could represent mountains, trees, or
abstract architectural structures.
- 13 - There are linear paths or curves within the image that may represent
roads, rivers, or flow of light, leading the viewer's eye through
the artwork.

14

15 Overall, while there are identifiable features, the abstract nature and
the complexity of the arrangement leave much to interpretation,
creating a unique and individualized understanding for each viewer.

5.2.3 Image 3

- 1 1. **Description of the ASCII Art**: The text-based image appears to be a
complex arrangement of symbols and characters that create a visual
pattern. The overall structure shows a range of dense symbols
interspersed with lighter characters and spaces. The vertical and
horizontal lines created by the use of symbols give the artwork a
somewhat pixelated appearance. However, due to the nature of ASCII art,
the image likely conveys some recognizable shape or form.
- 2
- 3 2. **Inferred Subject**: Given the arrangement of the characters and the
regularity in the flow of the symbols, I can infer that this ASCII art
likely represents an abstract figure, possibly a character or an animal
. There appear to be features reminiscent of a face or a mask,
suggesting it could be depicting a humanoid character or a creature.
- 4
- 5 3. **Visible Features and Characteristics**:
- The use of symbols such as '@', '\$', '#', 'E', and 'O' suggests depth
and texture, which could represent different materials or colors in
a real-world context.
- 6
- 7 - The overall outline formed by the symbols indicates a symmetry that
might be characteristic of a face or a stylized creature with
prominent features like eyes, a mouth, or other facial aspects.
- 8 - There are sections that seem to represent varying contrast,
potentially highlighting eyes or other features of the figure/object
represented.
- 9 - The use of escape characters and surrounding structures gives it a
framed or contained look, often mimicking the way art retains a
boundary.

10

11 In conclusion, this ASCII art likely depicts a stylized figure, possibly a
face or another characterized creature, built from a mosaic of symbols
that evoke a sense of depth and detail despite being created in a
simple text format.