

D7047E - Group 19 Lab 2

Willmer Wohlén

April 2025

1 Introduction

For this task two generative LLM's were used to generate images of humans in real life scenarios. The images were then analyzed by the group to find inconsistencies.

2 Results

Using ChatGPT and Gemini 5 images using different prompts were generated. The results shown in the figures below. The results show that the models have a difficult time generating realistic human body composition and that it gets worse when there are multiple individuals and small details. From a discussion with the group the conclusion was drawn that this probably has to do with the models only learning to generate images from a prompt and has no understanding of actual human anatomy.



Figure 1: Image generated using GPT-4 using the prompt "Cowboys in a saloon partaking in a thumb war competition"



Figure 2: Inconsistency found: Multiple left arms on the same person



Figure 3: Image generated with Gemini using the prompt "A photograph of a man driving his car"

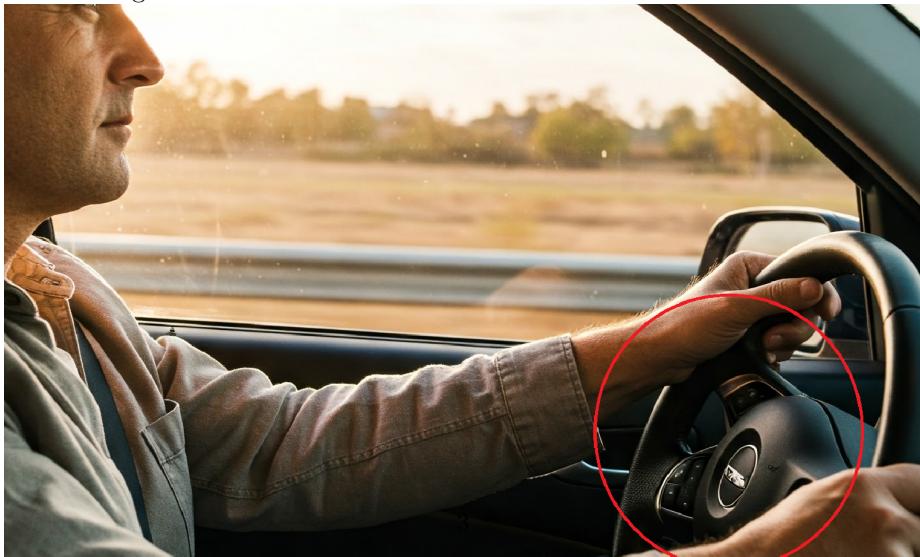


Figure 4: Inconsistency found: Odd driving wheel structure. Placement of the wheel does not match and is not symmetrical



Figure 5: Image generated using Gemini with the prompt "A swedish family doing a typical swedish family activity"



Figure 6: Inconsistency found: Unnatural hands



Figure 7: Image generated using Gemini with the prompt "Typical floridian family"





Figure 9: Image generated using Gemini of a woman in a car with a pug. Multiple inconsistencies in the structure of the car, the seats and handles