Music Generation with Deep Learning Techniques
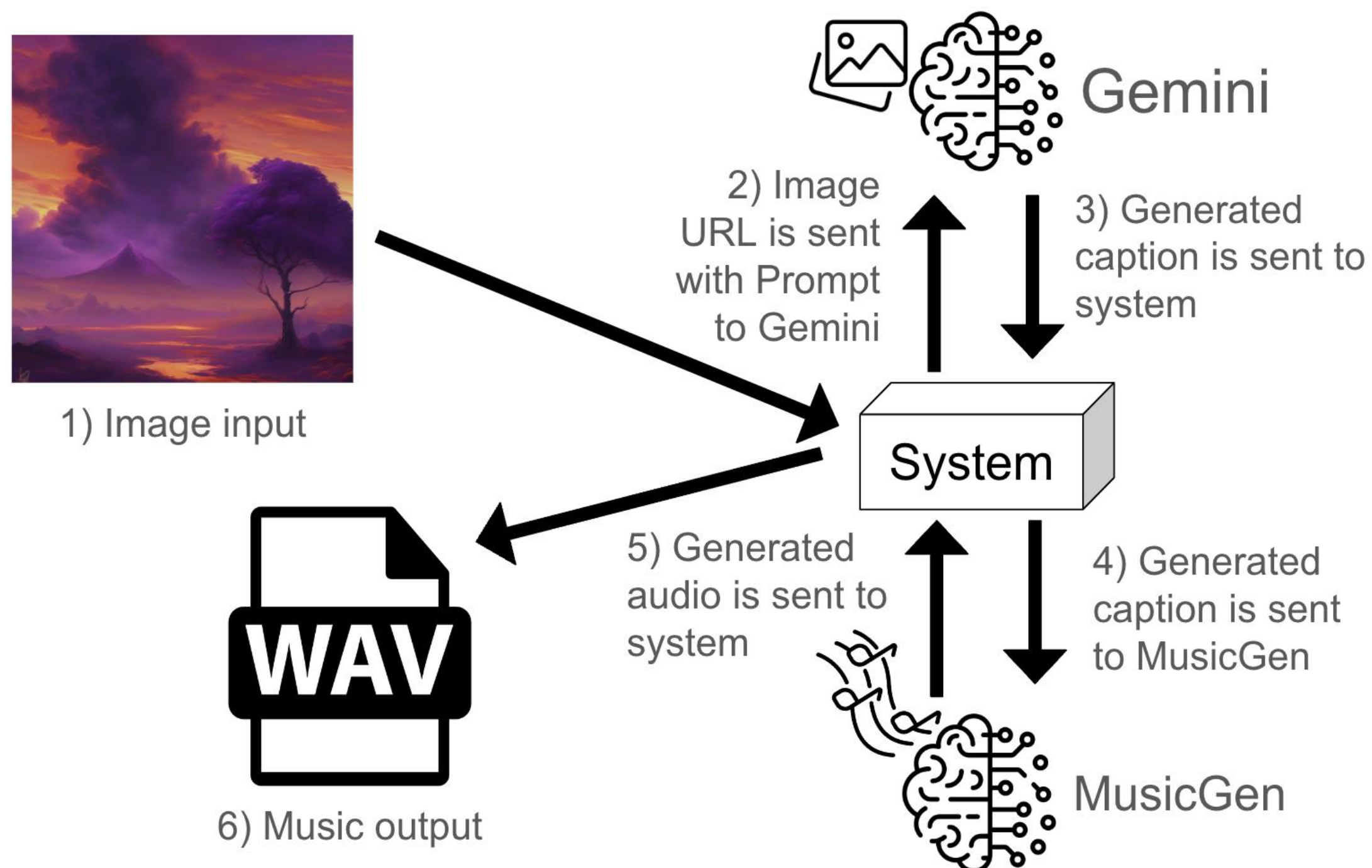
Generating music from images using captions as an intermediary

Student: Chong Yi An   Supervisor: Alexei Sourin

Research on music generation is rapidly evolving and improving but are mainly focused on generating music from text prompts. This project attempts a novel approach of generating music from images by first generating a caption then generating feeding the caption into the music generator. It also paves the way for future research into applications of deep learning music generation.

System Design

Existing deep learning models were integrated – Gemini for generating text captions and MusicGen for generating music from text prompts

Results

✓ Generated captions generally matched the content of the images and were suitable for music generation
✓ Generated music articulates mood conveyed by images

Listen to Excerpts by scanning this QR code!