

Exploring the use of text-based image generation for creative writing

Name: Gamar Ivan Azuaje Suarez

Laboratory's name: Social Computing

Supervisor's name: Professor Eiji Aramaki

Abstract

Text-to-image generation is a rapidly growing field in artificial intelligence that leverages deep learning models, particularly Generative Adversarial Networks (GANs), to transform textual descriptions into visual representations. This technology has found applications in various creative domains, enabling users to generate images from simple prompts, detailed descriptions, or even abstract concepts. This thesis explores the application of text-to-image generation in two distinct creative domains, aimed at enhancing creativity and emotional well-being.

Firstly, we introduce Visualyre, a tool designed to assist musicians in creating album art. This application employs GANs and Style Transfer to produce unique visuals that reflect the themes, moods, and identities embedded within the musicians' song lyrics and audio files. A user study involving 35 amateur and independent musicians highlights Visualyre's effectiveness in providing a valuable resource for artists, especially those with limited financial resources or design skills, to enhance their musical projects with personalized album artwork.

Secondly, we investigate the potential of real-time image generation during fictional writing exercises with StoryWriter, another GAN-based tool aimed at regulating negative emotions to enhance the writing experience. Although user studies show mixed results, they reveal the potential of StoryWriter to improve emotional outcomes. However, concerns were raised about the impact of generated images on the therapeutic benefits of writing. Overall, these projects underscore the versatility and potential of text-to-image generation in fostering artistic creativity and emotional health.