Integration of Deep Learning in Automatic Music Generation, Aiming at Preserving and Developing Don Ca Tai Tu

Nguyen, Duc (School: Le Hong Phong High School for the Gifted)
Ha, Bao (School: Le Hong Phong High School for the Gifted)

The application of modern technologies is essential for preserving and revitalizing Don ca tai tu, a UNESCO-recognized Intangible Cultural Heritage of Humanity, increasingly overshadowed by the rise of contemporary music genres. However, current efforts primarily reach a limited audience, while younger generations remain largely uninterested in this traditional art form. To address this challenge, we developed a deep learning framework for automatic music generation, allowing users to compose musical pieces infused with the stylistic essence of Don ca tai tu using natural language prompts. We curated a specialized dataset by collecting recordings from multiple online sources, experienced musicians, and cultural experts. The data was meticulously filtered and processed to ensure a clean, high-quality training set. Moreover, we introduce DcttGen, a unified framework to generate high-fidelity, long-form compositions based on text inputs. It achieves this through (1) Music Tokenizer to extract essential musical features from the raw audio waveform efficiently, (2) Autoregressive Transformer to maintain long-term structural coherence in the generated music, and (3) Rectified Flow Transformer to synthesize high-fidelity audio outputs from low sampling rate representations. Furthermore, we redesigned the Chain-of-Thought prompting technique tailored for the Don ca tai tu music, enhancing the resulting compositions' coherence and creativity. With this combination, DcttGen can produce high-quality, well-structured musical pieces compared to other baselines. Experimental results show that DcttGen matches or even surpasses recent methods in various objective metrics. Through this work, we introduce a novel AI-driven approach to the preservation and revitalization of Don ca tai tu.

Awards Won: