

昆曲·錯山記

黃令東, 姜政, 孫璿承, 柏彤

Kunqu: Legend of Wrong Mountain

Lingdong Huang, Zheng Jiang, Giada Sun, Tong Bai

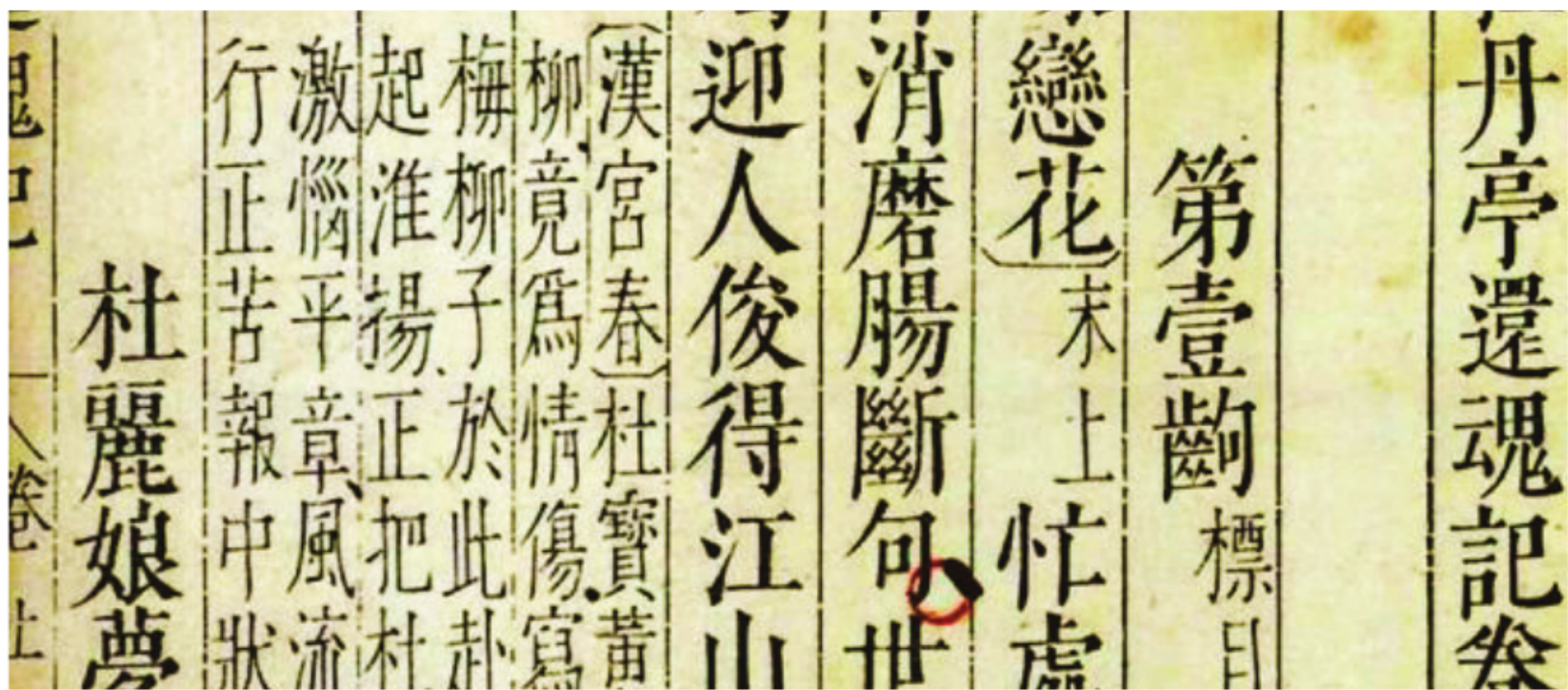
Kunqu by Machine

We try to generate a traditional Chinese Opera — Kunqu by Machine Learning method. Our technique is combined with the generation methods of text, video, and sound.



Motivation

Kunqu opera, one of the greatest and most classic forms of Chinese Art, embodies China's rich cultural background and aesthetics in its melodic singing and expressive performance. Many scripts and music were written for Kunqu over the past hundreds of years. Combining this reservoir of traditional masterpieces with cutting-edge Machine Learning methods, our group produced a number of aesthetically and technically interesting results. Among them we present The Legend of Wrong Mountain, a Kunqu Opera entirely hallucinated by a computer, with its script, music and singing all generated using machine learning techniques. Our work is an attempt to explore the relationship between Machine Learning and traditional art forms, and through this mixture the emergence of novel and exciting contemporary elements in art.

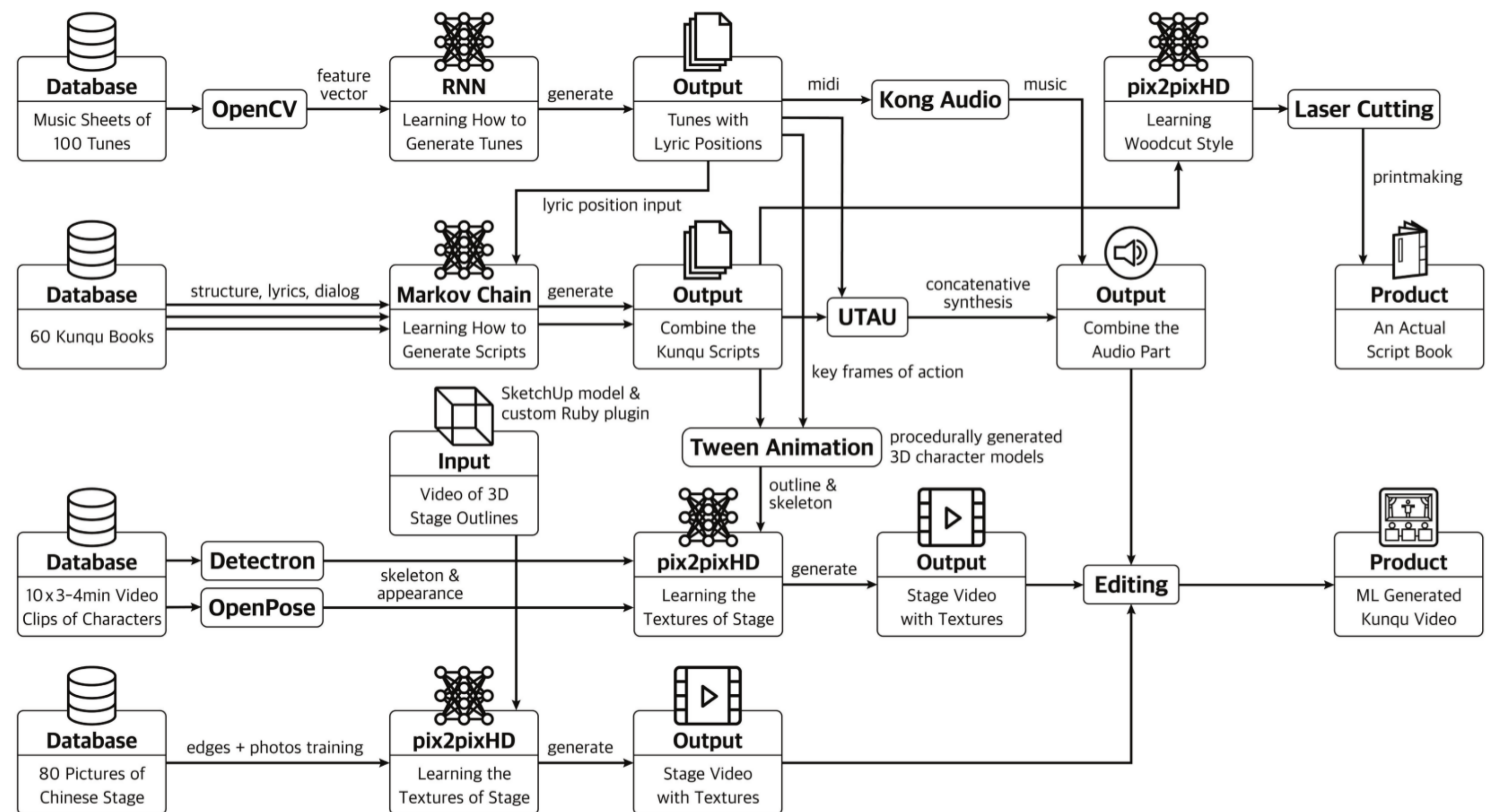


References

- [1] He, Kaiming, et al. "Mask r-cnn." Computer Vision (ICCV), 2017 IEEE International Conference on. IEEE, 2017.
- [2] Cao, Zhe, et al. "Realtime multi-person 2d pose estimation using part affinity fields." CVPR. Vol. 1. No. 2. 2017.
- [3] Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." arXiv preprint (2017).
- [4] Wang, Ting-Chun, et al. "High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs." arXiv preprint arXiv:1711.11585 (2017).

Process

Our group mainly use three generation methods to generate the whole project. We used Markov Chain to generate the story script. For the vision expression part, we collected our multi-category paired dataset using Detectron[1], OpenPose[2], and we mainly used Pix2pix[3] and Pix2pixHD[4] to generate the appearance of performers. Also, we trained a calligrapher transfer model to display our generated script with traditional Chinese calligrapher. We trained different model for each character and our environment and used those models to transfer the information map rendered by 3D sketches. For sound and music part, we collected music sheets of tunes to train a Recurrent Neural Network (RNN) and used it to generate tunes. Then, we converted tunes to midi format and imported them to Kontakt and Pro Tools to create multichannel Chinese style music. In addition, we used UTAU to generate voice on the pitch of tunes.



Results

We created an audio-visual experience featuring a performance of Kunqu Opera hallucinated by the computer and an actual woodblock-printed script book. Through the understanding and re-interpretation of existing literature and sound materials, our algorithm is able to generate convincing music and lyrics conforming to Qupai, melodic rulesets of Chinese opera, as well as artistically expressive facial and bodily motions from the traditional art form. There's still room for improvement of generating natural performer's voice and interpreting the relation between motion and script. Obviously, drama is not only comprehensive but also very complicated art form which is not easy for machine to learn and mimic.

