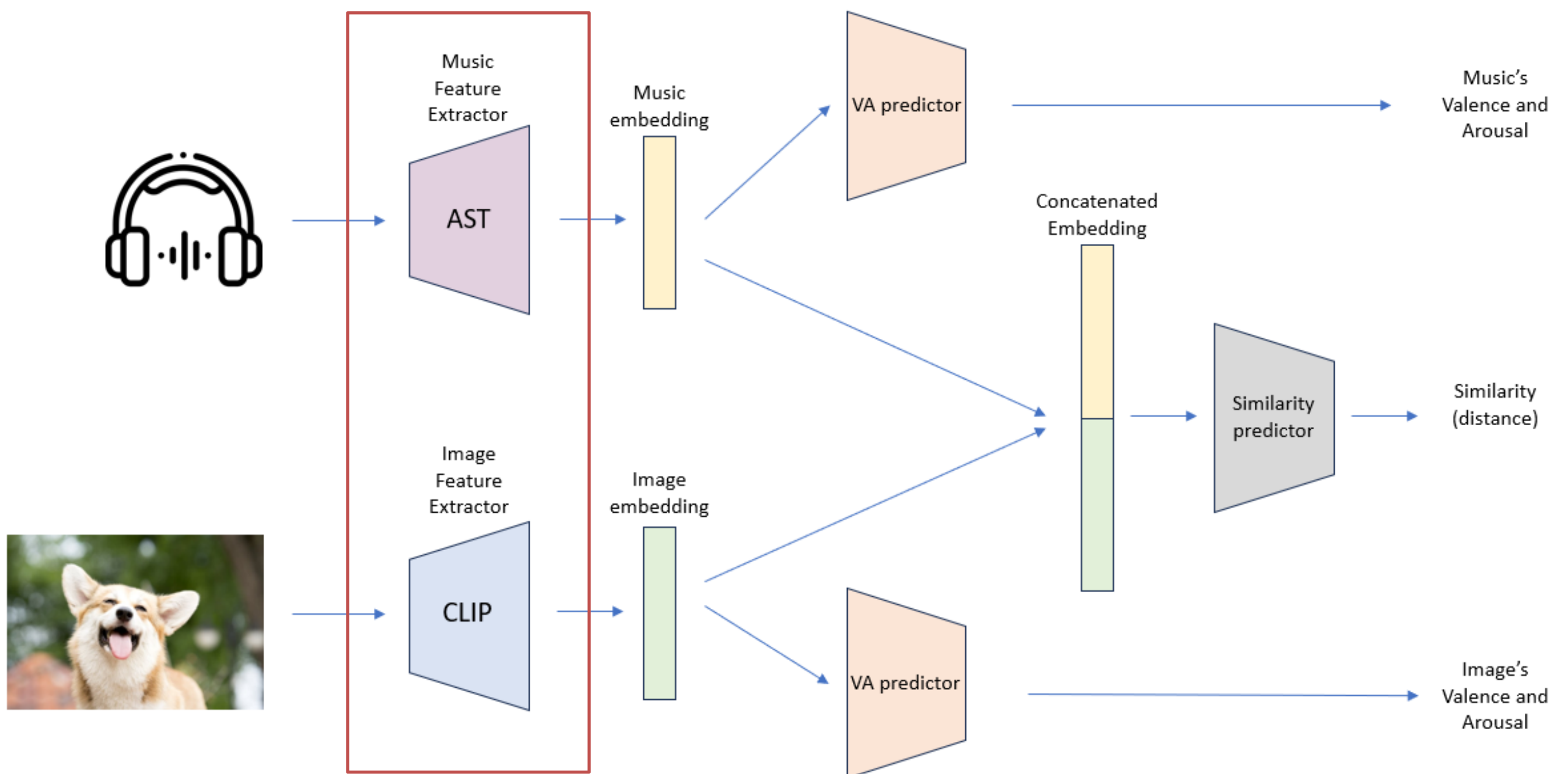


# Music Visualization

## With Deep Learning

Student: Sim Xin Ni Clodia Supervisor: Assoc Prof Alexei Sourin



### Project Objective

- This project aims to experiment using advanced transformers as feature extractors instead of Residual Neural Networks and Convolutional Neural Networks.
- The goal is to improve the performance of an emotion-based image-music matching model, and thereafter improve the visualization results of music.

### Approach & Conclusion

- The approach undertaken utilized popular transformer models, Audio Spectrogram Transformer (AST) and Contrastive Language-Image Pretraining (CLIP) as feature extractors.
- Results revealed a successful improvement in performance and was further verified by a user study.

### Sample (Beethoven, Ode to Joy)

