Autonomous University of San Luis Potosí
Engineering Faculty
Machine learning

SynthTalkOne

An AI system to synthesize and cloning human voices for use in chatbots, virtual assistants, and other applications.

- Name of participant: Fernando Antonio Ramírez Martínez
- Adviser: Juan Carlos Cuevas Tello

Types of generated sounds

- Speech (Text-to-Speech)
- Music
- Music notes (samples)
- Sound design
- ...

https://github.com/musikalkemist/generating-sound-with-neural-networks

Sound representations

- Raw-audio
- Spectrograms

https://github.com/musikalkemist/generating-sound-with-neural-networks

Generation from raw audio: Challenges

- Difficult to capture long-range dependencies

Pitch

Melody

Structure

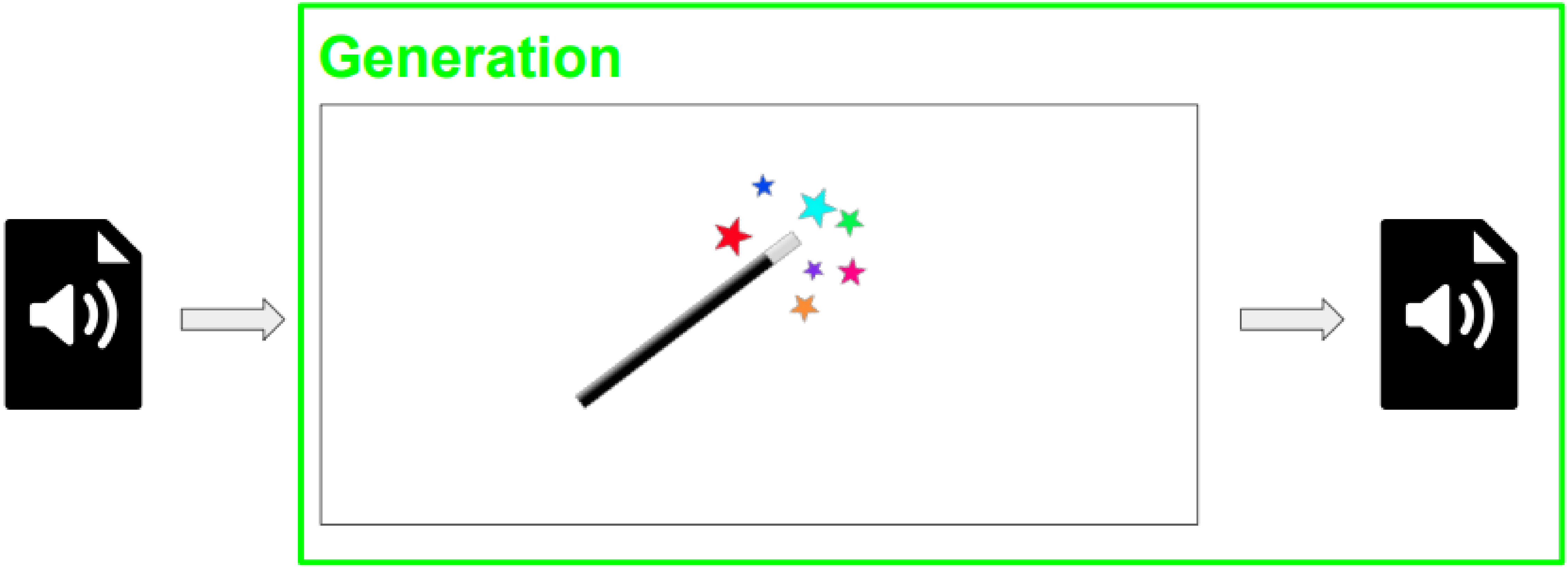
Timbre

Rhythm

Harmony

https://github.com/musikalkemist/generating-sound-with-neural-networks

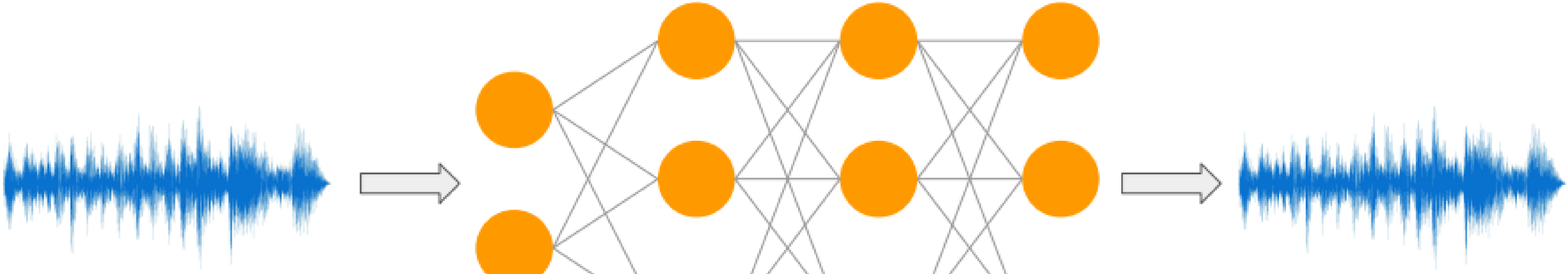
Sound generation task



Generation

https://github.com/musikalkemist/generating-sound-with-neural-networks

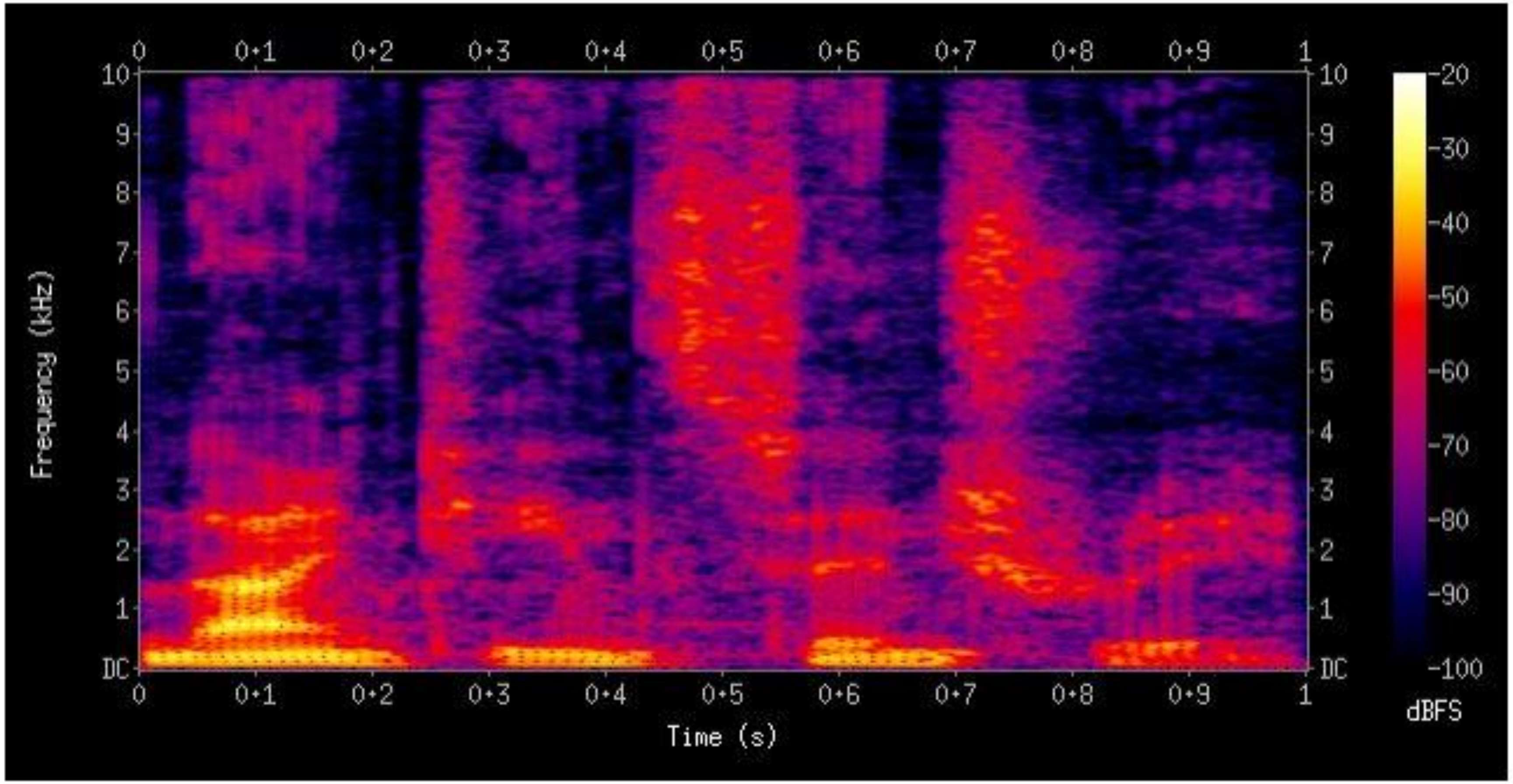
Generation from raw audio



https://github.com/musikalkemist/generating-sound-with-neural-networks

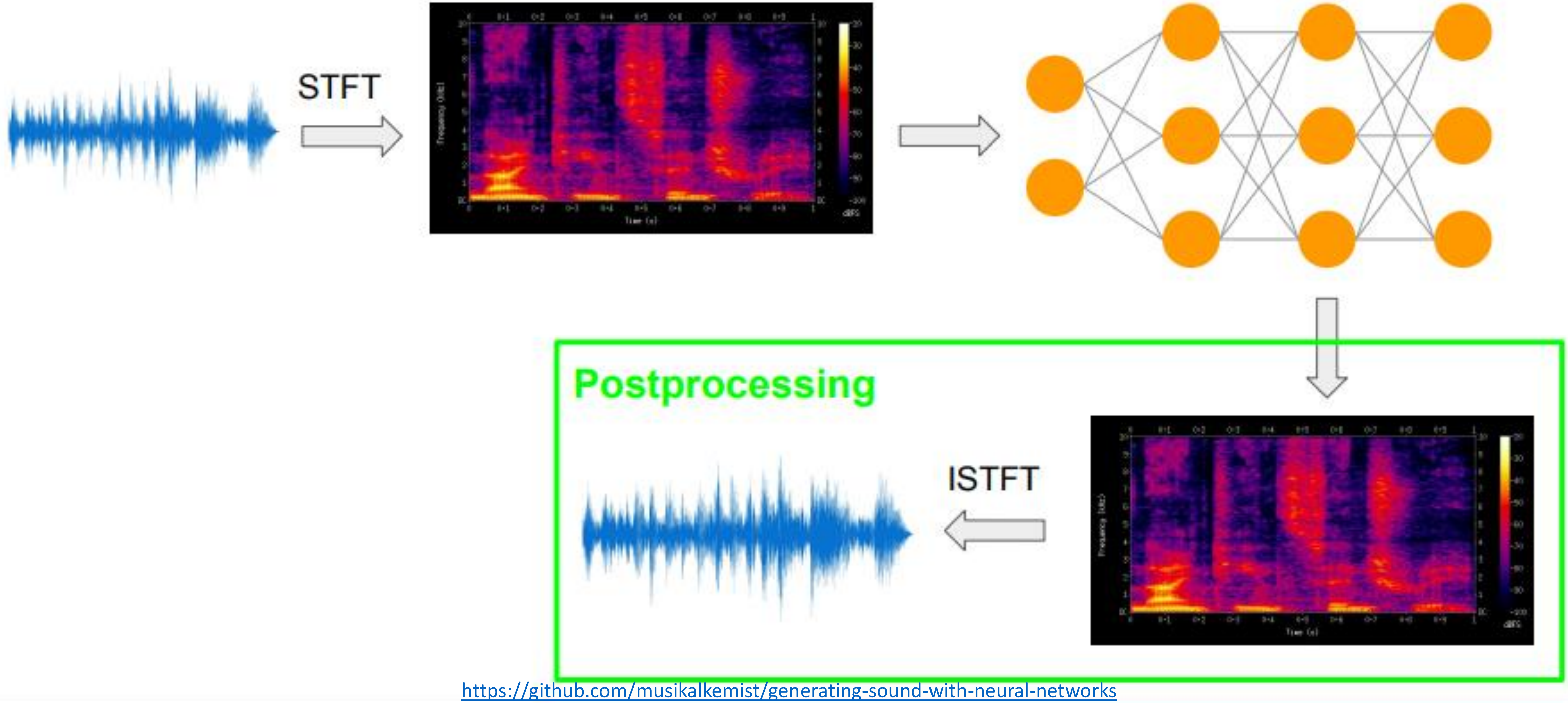
Use a more compact representation of sound

https://github.com/musikalkemist/generating-sound-with-neural-networks



https://github.com/musikalkemist/generating-sound-with-neural-networks

Generation from spectrograms



STFT

Postprocessing

ISTFT

https://github.com/musikalkemist/generating-sound-with-neural-networks

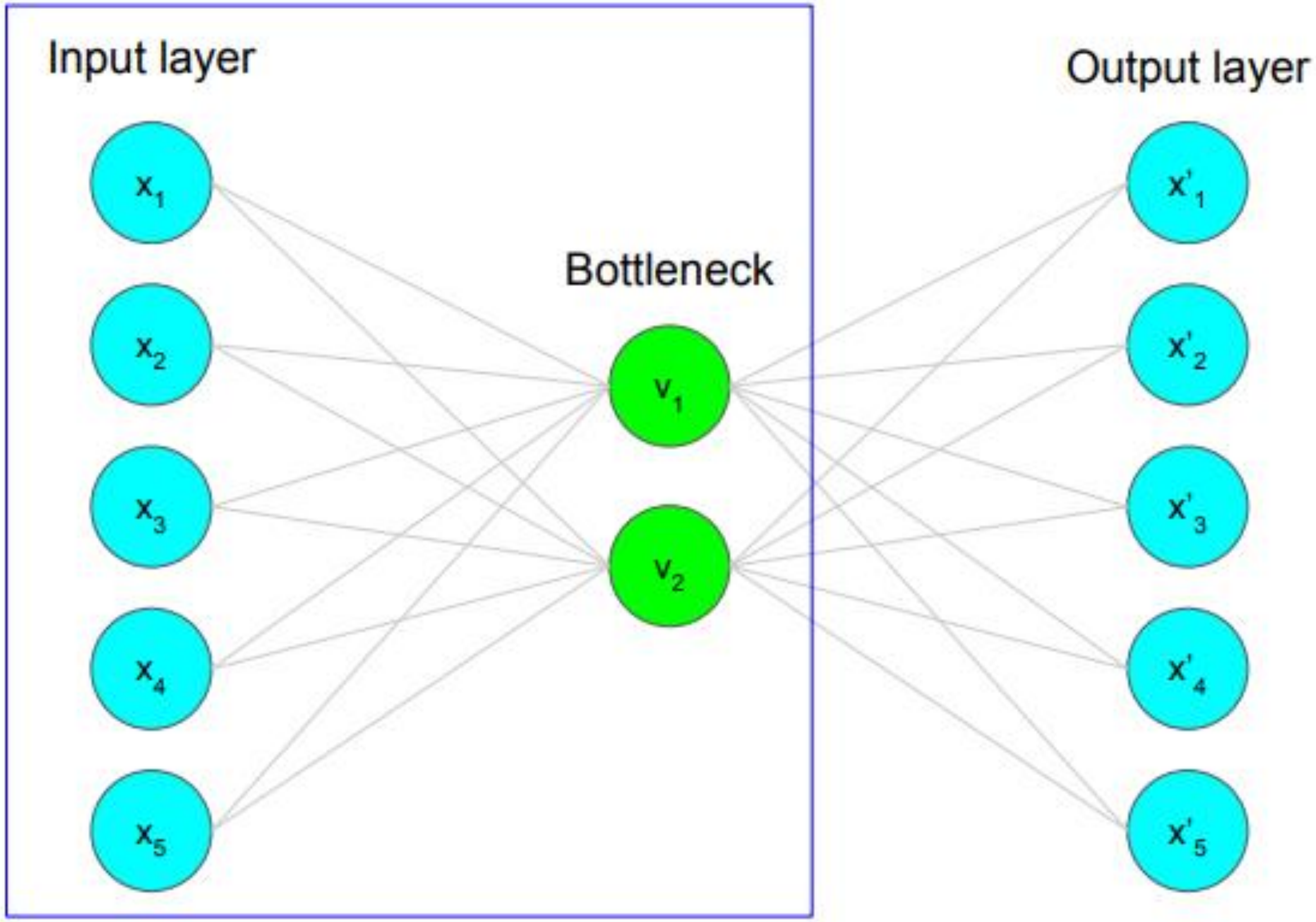
Autoencoders: The sneaky idea

Create an architecture with a bottleneck, which ensures a lower-dimensional representation of the original data.

https://github.com/musikalkemist/generating-sound-with-neural-networks

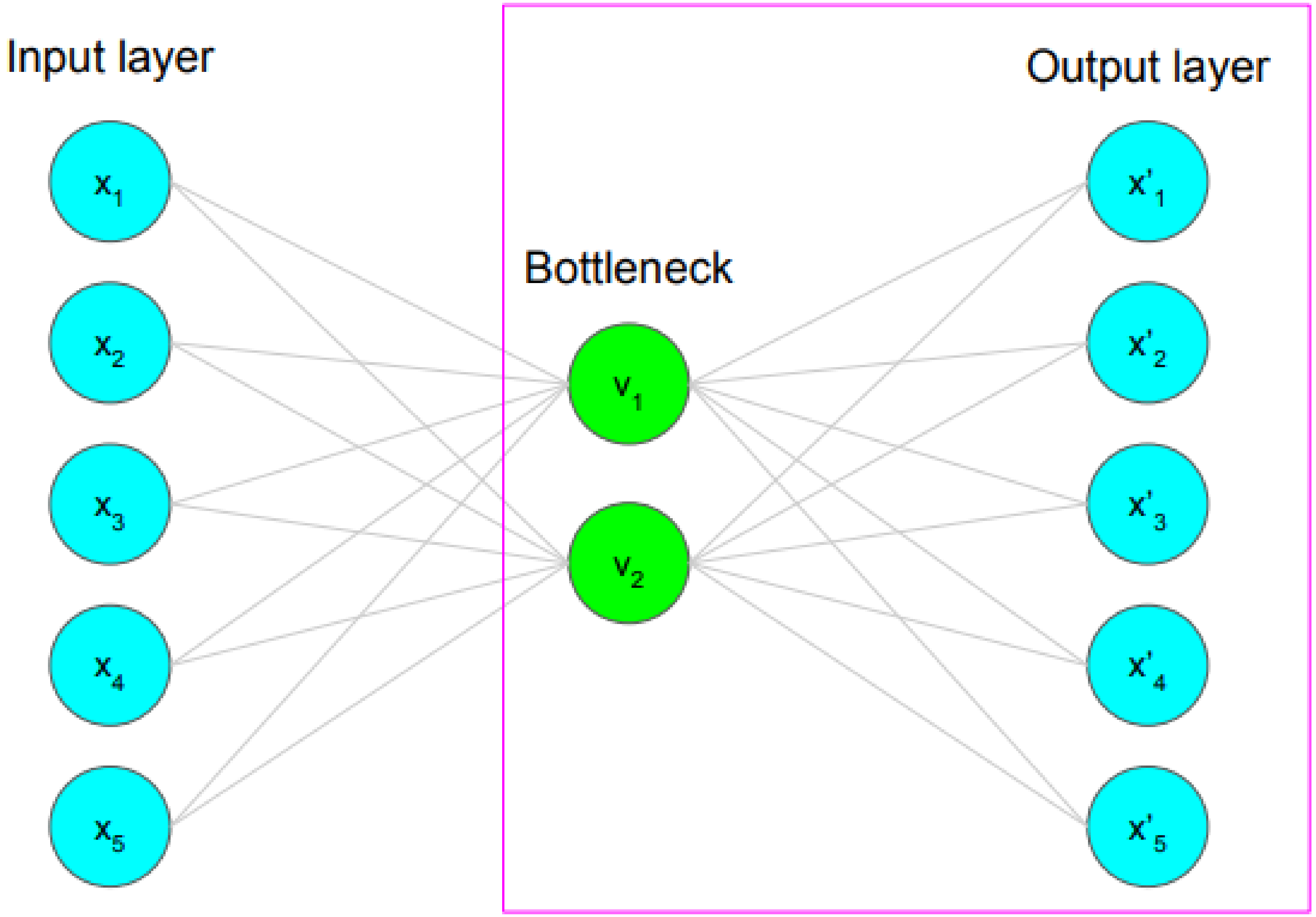
Autoencoder = Encoder + Decoder

https://github.com/musikalkemist/generating-sound-with-neural-networks



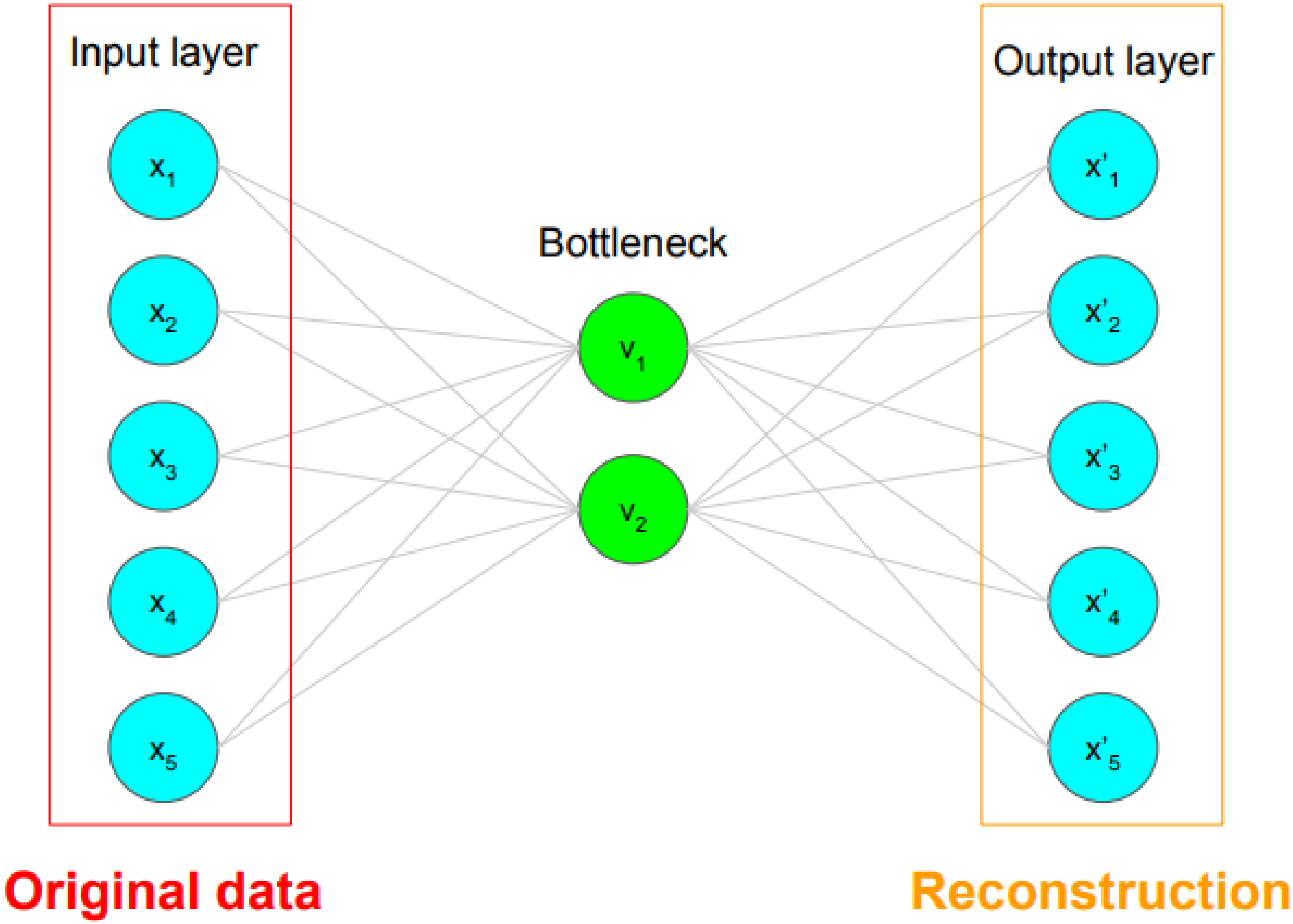
Encoder = compress data into lower-dimensional representation (*latent* space)

https://github.com/musikalkemist/generating-sound-with-neural-networks



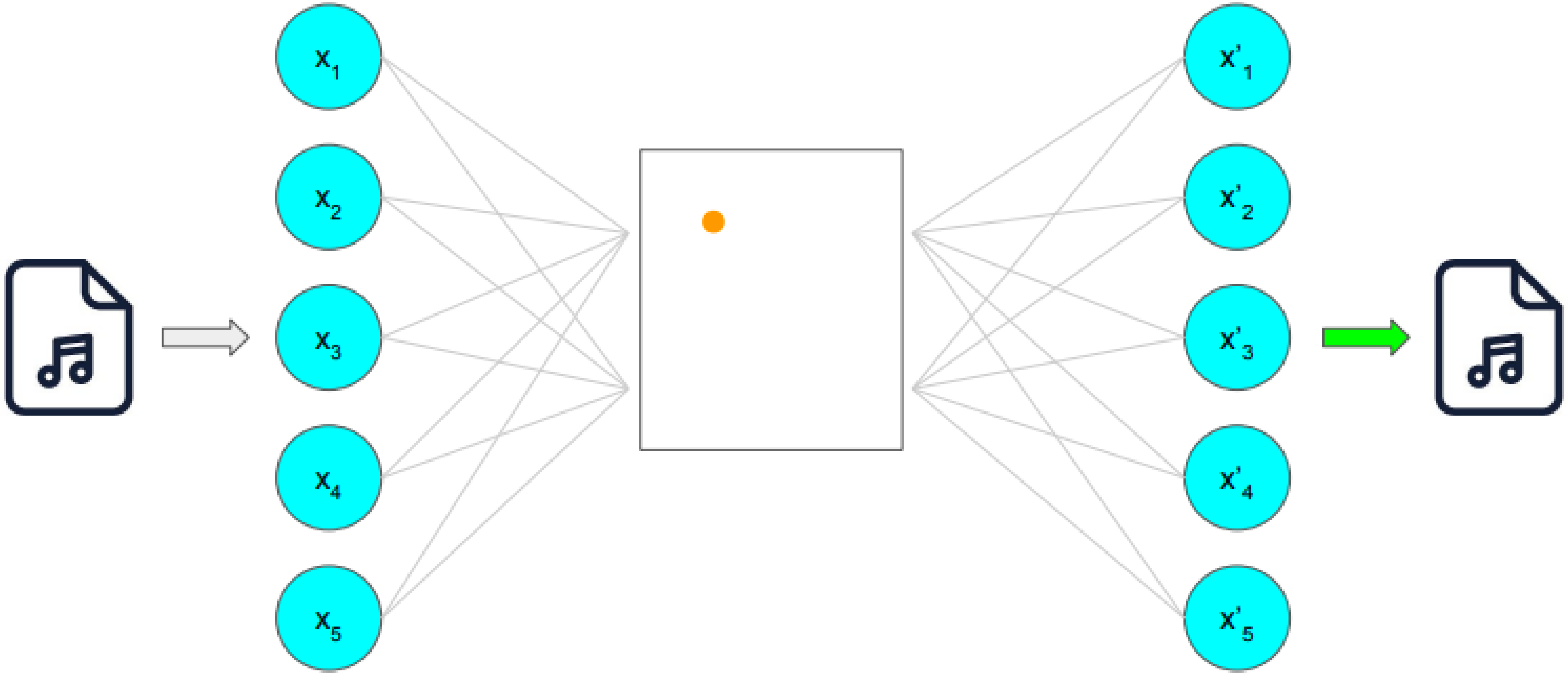
Decoder = Decompress representation back to original domain

https://github.com/musikalkemist/generating-sound-with-neural-networks



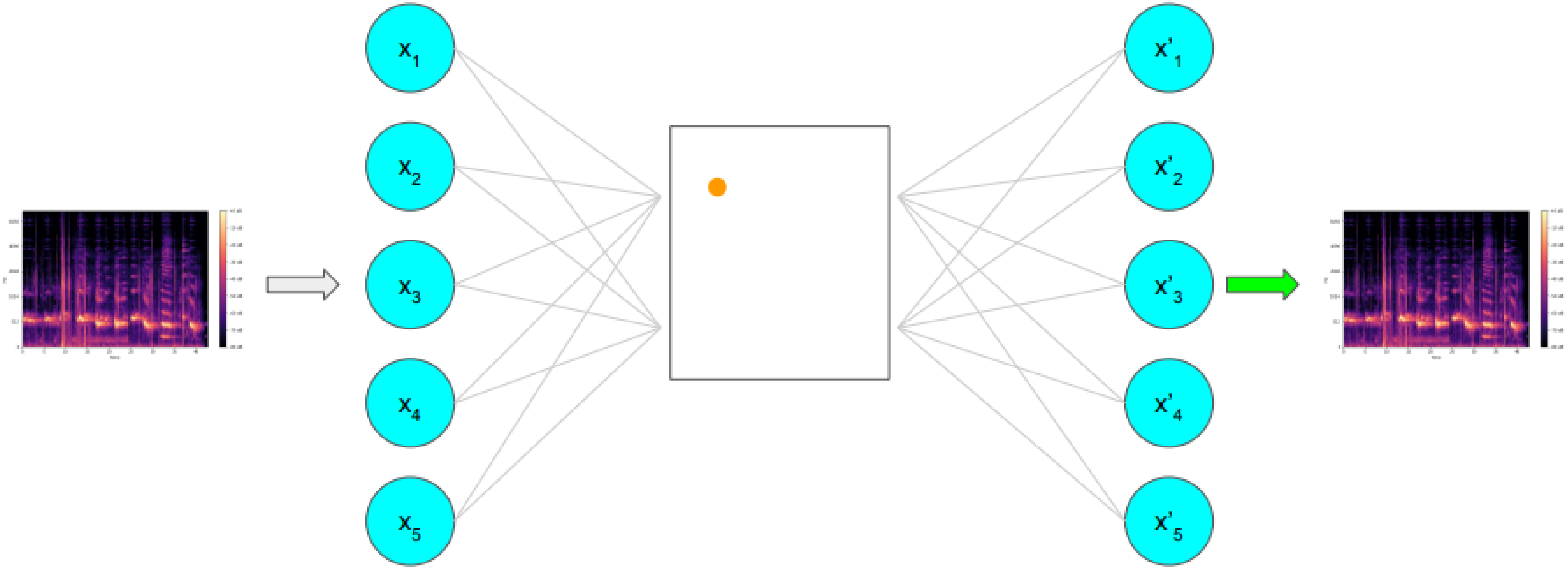
https://github.com/musikalkemist/generating-sound-with-neural-networks

Generation with AEs



https://github.com/musikalkemist/generating-sound-with-neural-networks

Generation with VAEs



https://github.com/musikalkemist/generating-sound-with-neural-networks

This content of this slides was obtained by the following repository:

https://github.com/musikalkemist/generating-sound-with-neural-networks