Travail Pratique Transmédia

Uandha Fernandes Barbosa,
BMus and BScEng

MUS3329X
Projet en Informatique Musicale
February 2026

Université Concordia University

Travail Pratique

Transmédia

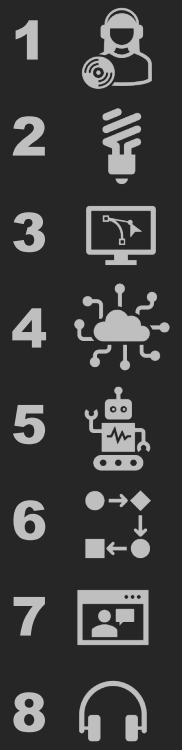
Uandha Fernandes Barbosa,

BMus and BScEng

MUS3329X

Projet en Informatique Musicale

February 2026



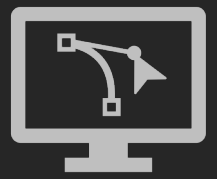
1 ABOUT ME

- ❑ **Bachelor of Engineering – Brazil**
- ❑ **Bachelor of Music in Composition – Canada**
- ❑ **Master’s Candidate at Concordia University:**
  - **My research combines creative practice and technology focusing on machine learning and human-computer interaction in arts creation.**
  - **Currently investigating “embodied knowledge” and LLM co-creation in Uzulangs and live coding environments.**

2 PROJECT IDEA

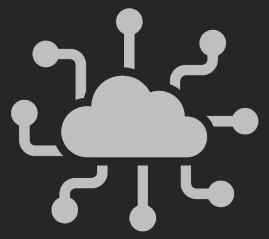
- Live coding performance often requires the performer to remain highly focused on the computer interface.
- This mode of interaction may not fulfil performer's sensation of physical embodiment when compared to instrumental performance (e.g. playing a MIDI keyboard).
- To address embodied interaction, I propose for this project the development of a small neural network that amplifies gestural presence.

3 PYTHON SCRIPT



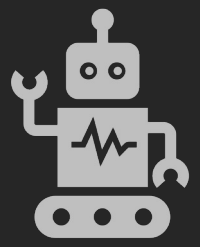
- In live coding performances, performers are typically either standing or seated in front of their computers.
- Since the performer is already positioned in front of the computer, I chose to use the computer's built-in camera as an input device to capture existing body movements.
- For this investigation, I focus on facial expressions that are initially captured through a Python script using MediaPipe as 10 parameters.
- After preliminary testing, these were reduced to three distinct expressive configurations that provide clear positional differentiation: rounded mouth, kiss position, smile (measured bilaterally).

4 MODEL'S DATASET



- The 3 facial expression measurements (normalized between 0–1) serve as the input dataset for the neural network.
- The output consists of 7 values (also normalized), which control parameters across two different 2 VST plugins - OXID and AAS (Applied Acoustics Systems).
- These VSTs receive MIDI input from the Strudel REPL, which the performer uses to control the musical structure in real time.
- Using the defined input/output relationships, a dataset was constructed for training the model. Each data point was adjusted point-by-point to refine mappings and then saved (total of 205 training points).

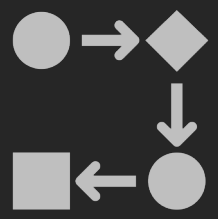
		PARAMETERS							
		Hidden layers	Activation	Output activation	Batch size	Max iter	Learn rate	Validation	Loss
MODEL	1	20 20 20	3	1	1	1000	0,01	0,05	0,005
	2	20 20 20	1	1	10	1000	0,01	0,1	0,025
	3	20 20 20	1	1	10	1000	0,01	0,03	0,029
	4	20 20 20	3	1	10	1000	0,01	0,05	0,019
	5	20 20 20	1	1	5	1000	0,01	0,01	0,021
	6	20 20 20	3	1	3	1000	0,01	0,05	0,099
	7	20 20 20	1	1	3	1000	0,01	0,1	0,031



5 MODEL'S TRAINING

- **Using the same dataset (205 training points), smaller batch sizes produced lower training loss.**
- **Activation 3 (nonlinear) demonstrated better results compared to activation 1 configuration.**

6 MODEL'S REFLECTIONS



- **Larger Dataset: expanding the dataset with more recorded samples would improve model stability and training process.**
- **Reducing Positional Bias: the current model sometimes interprets different head positions as different facial expressions.**
- **Multimodal Comparison: investigate diverse body capture configuration (e.g. upper-body tracking) would allow comparative evaluation of different input modalities and improve mapping accuracy.**

Email:
uandha.fernandesbarbosa@mail.concordia.ca

SCAN ME

7

INFO

Email:
uandha.fernandesbarbosa@mail.concordia.ca

SCAN ME

8

DEMO