Music Composition with Deep Learning: A Review

Carlos Hernandez-Olivan*, Jose R. Beltran

April 2022

*carloshero@unizar.es

Universidad Zaragoza

Carlos Hernandez-Olivan

José Ramón Beltrán

Carlos Hernández Oliván is a PhD student in Electronics at the Universidad de Zaragoza under the supervision of Dr. José R. Beltrán. He received the B.E. and M.Sc. degrees in Industrial Engineering in 2017 and 2019, respectively. He studied viola at the Professional Conservatory of Zaragoza where he received his professional certificate in 2013. His research interests are focused on Music Information Retrieval, in particular, on music analysis and generation systems with Artificial Intelligence.

José R. Beltrán is an Associate Professor with the Department of Electronic Engineering and Communications, University of Zaragoza. He has been involved in different research and development projects on Audio Analysis and Processing. His research interests are focused on the study of Automatic Learning Systems for the analysis, processing and synthesis of the musical signal. In 2008, he was a promoter of an academic spin-off: ARSTIC Audiovisual Solutions S.L. devoted to the use of technologies for the artistic and audiovisual fields.

Prof. Beltrán is a member of the Aragon Institute for Engineering Research (I3A), Reseach Group in Advanced Interfaces (AffeciveLab).

Universidad
Zaragoza
1542

Outline

➤ What is Automatic Music Composition?

➤ From Algorithmic Composition to AI-based Music Composition

➤ Music Composition with Deep Learning

➤ AI-based Music Composition Evaluation

➤ New Trends: Human-Computer Interaction, style stransfer...

Universidad Zaragoza
1542

Carlos Hernandez-Olivan, Jose R. Beltran

Some Open Questions

▶ Are the current DL models capable of generating music with a certain level of creativity?

▶ What is the best NN architecture to perform music composition with DL?

▶ Could end-to-end methods generate entire structured music pieces?

▶ Are the composed pieces with DL just an imitation of the inputs or can NNs generate new music in styles that are not present in the training data?

Universidad
Zaragoza
1542

Carlos Hernandez-Olivan, Jose R. Beltran

A Few more...

➤ Should NNs compose music by following the same logic and process as humans do?

➤ How much data do DL models for music generation need?

➤ Are current evaluation methods good enough to compare and measure the creativity of the composed music?

➤ Are we (researchers, developers...) designing DL models according to the music basic principles, or are we just taking models from other fields to do the music composition task?

Outline: Music Classification (1880s – 2020s)

"UTILITY" MUSIC

Religious
Soundtracks
(TV, Cinema...)
...

MUSIC OF THE WORLD

Habanera
Salsa
...

Musicmap | The Genealogy and History of Popular Music Genres

CLASSICAL MUSIC

Genres/Periods

Medieval
Renaissance
Baroque
Classical
Romantic
Modern
Contemporary

Forms

Symphony
Chamber
Opera
Vocal

POPULAR MUSIC / FOLK

Rock 'N' Roll
Pop
Punk
...
Country

Blues
Jazz
Gospel

Rap

Techno

Universidad Zaragoza

Carlos Hernandez-Olivan, Jose R. Beltran

Outline: Why Music Composition is difficult

➤ Music is subjective

➤ Each music Genre has its own rules ——→ Tonal/atonal...

➤ Music means repetition ——→ Has local and high-level structures

➤ 2 axes ——→ Harmony and time are dependent to each other

➤ Intricacies
- Passing notes
- Appogiated chords
- Ornamentations
- ...

Universidad Zaragoza
1542

Carlos Hernandez-Olivan, Jose R. Beltran

The "Power" of Motifs: Examples

Composition:

> ▶ "mystery" 🔊   Only piano, high notes, minor chords with altered notes...

Composition + Orchestration + Performance:

> ▶ "power" 🔊

Universidad Zaragoza 1542

Carlos Hernandez-Olivan, Jose R. Beltran

Carlos Hernandez-Olivan, Jose R. Beltran

Universidad
Zaragoza
1542

Carlos Hernandez-Olivan, Jose R. Beltran

Universidad Zaragoza
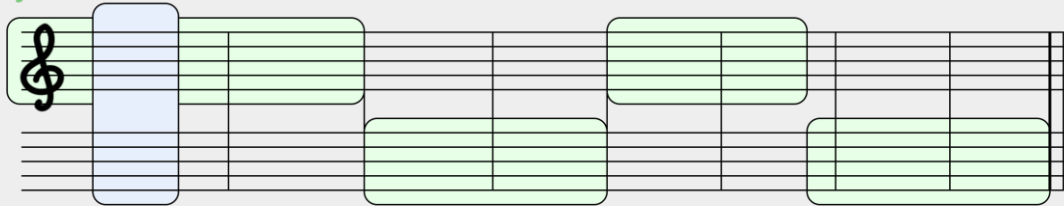
Music Composition with DL: Motivation

Music Generation papers per publication year

Source: "music generation" search title and abstract (dimensions.ai)

Carlos Hernandez-Olivan, Jose R. Beltran

Universidad Zaragoza
1542

Music Composition with Deep Learning:
A Review

Carlos Hernandez-Olivan, Jose R. Beltran

RNN?

modify

RL?

Human

idea

Composition process
(write some chords, melody...)

f

AI

TRAINING

inputs

Model/Algorithm

outputs

inference

"NEW" MUSIC

The beginnings: Algorithmic Music

[Nierhaus, 2009] Algorithmic Composition: paradigms of automated music generation.

Algorithmic Music Models

- Markov chains
- Generative grammars
- Cellular Automata
- Chaos Theory
- Genetic Algorithms
- Transition Networks

Carlos Hernandez-Olivan, Jose R. Beltran

Universidad
Zaragoza
1542

The beginnings: Algorithmic Music

Compositions and Projects

➤ ILLIAC Suite
  [Hiller & Isaacson, 1957]

➤ EMI or Emmy [David Cope, 1980s]

➤ Analogiques A and B by Iannis Xenakis

➤ Project1 (PR1) [Koening, 1964]

Universidad
Zaragoza
1542

Carlos Hernandez-Olivan, Jose R. Beltran

First Music composition DL-based models

[Lewis, 1988] Creation by refinement: a creativity paradigm for gradient descent learning networks

> ▶ Training + Creation phases

[Todd, 1988] A Sequential Network Design for Musical Applications.

[Todd, 1989] A connectionist approach to algorithmic composition.

> ▶ Sequential network ⟶ Use memory (notes already produced) by feedback connections

[Mozer, 1994] Neural network composition by prediction: Exploring the benefits of psychophysical constraints and multiscale processing.

> ▶ CONCERT Network [Elman, 1990] ⟶ Continues a sequence of notes
> ⟶ Output is the probability dist. over the candidates

Universidad Zaragoza 1542

Carlos Hernandez-Olivan, Jose R. Beltran

Music composition with DL: Dimensions

[Briot et al., 2019] Deep Learning Techniques for Music Generation – A Survey

REPRESENTATION
Inputs
Nature

NN ARCHITECTURE
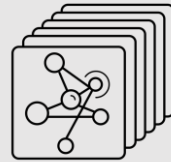
OBJECTIVE
Output
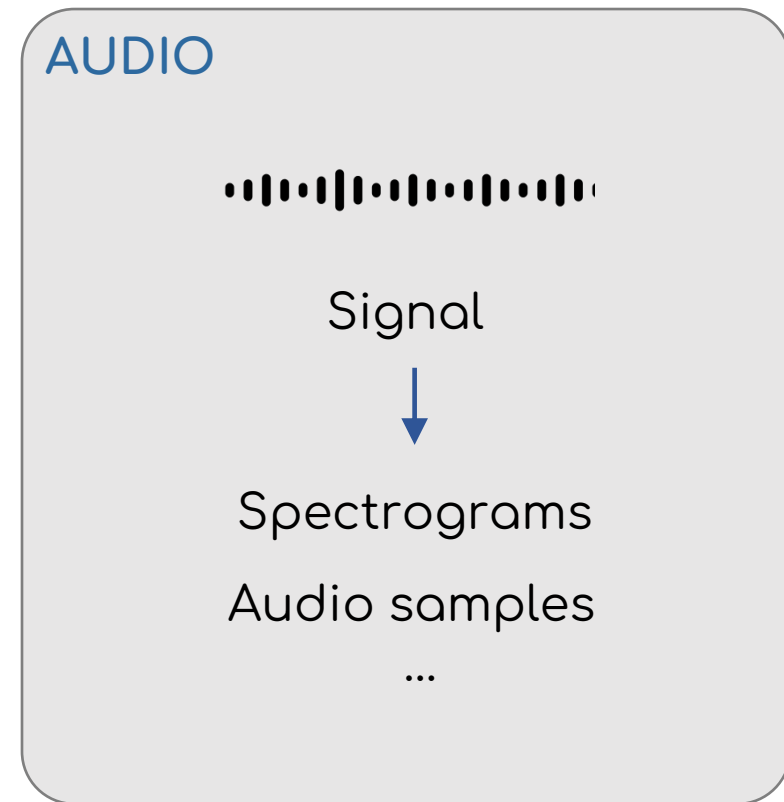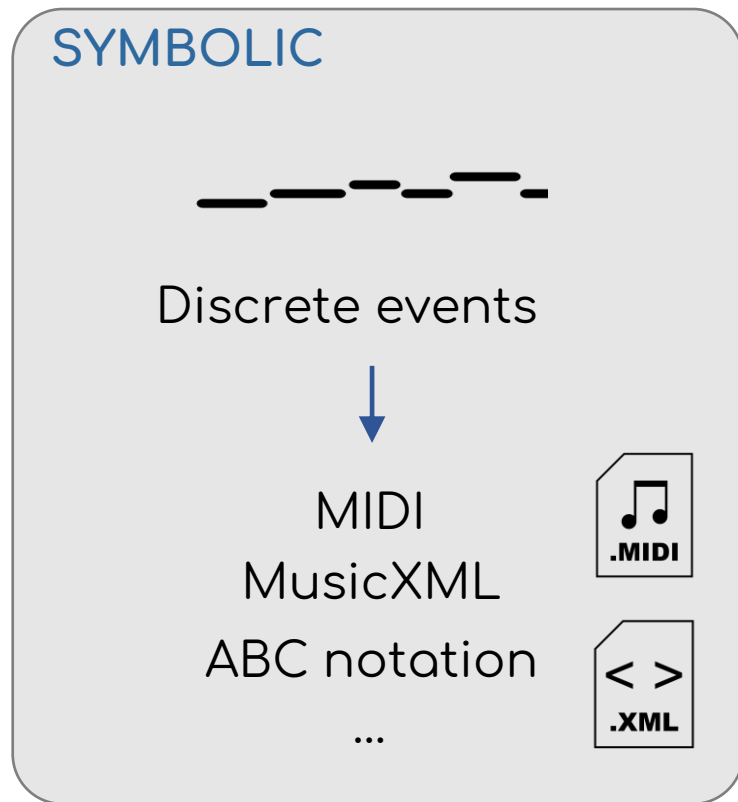Texture and
Form

CHALLENGES
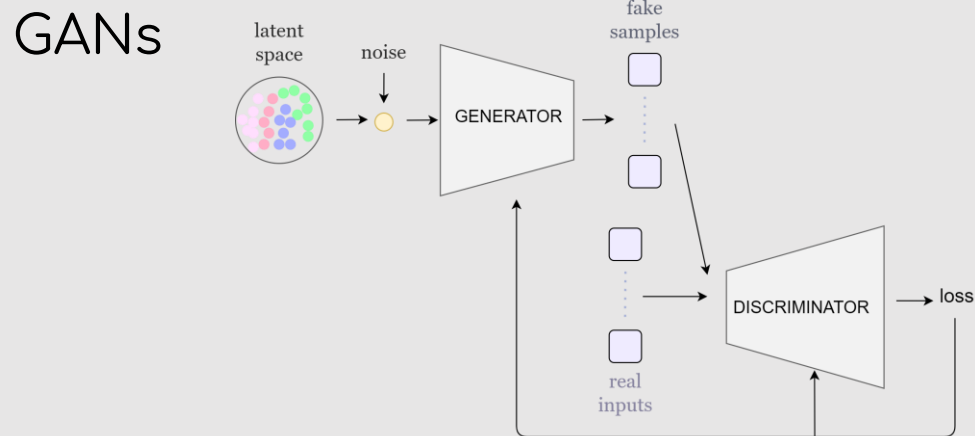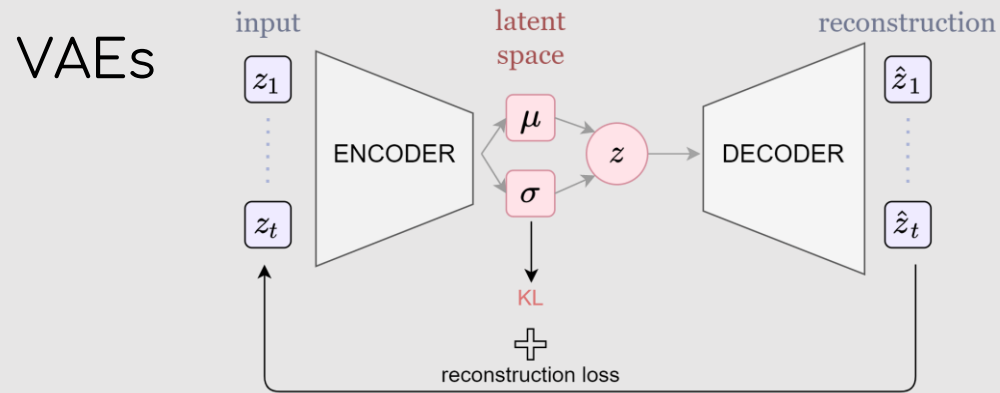Future
Directions

STRATEGY
How Outputs
are
generated

Universidad
Zaragoza
1542

Carlos Hernandez-Olivan, Jose R. Beltran

Input Representation

Input Nature:
Representation

SYMBOLIC

Discrete events

↓

MIDI
MusicXML
ABC notation
...

♫ .MIDI

< > .XML

AUDIO

Signal

↓

Spectrograms

Audio samples

...

Carlos Hernandez-Olivan, Jose R. Beltran

Universidad
Zaragoza
1542

NN Architectures

Generative Models

VAEs

GANs

Sequence Models

Recurrent Neural Networks (RNNs)

Transformers

Carlos Hernandez-Olivan, Jose R. Beltran

Universidad
Zaragoza
1542
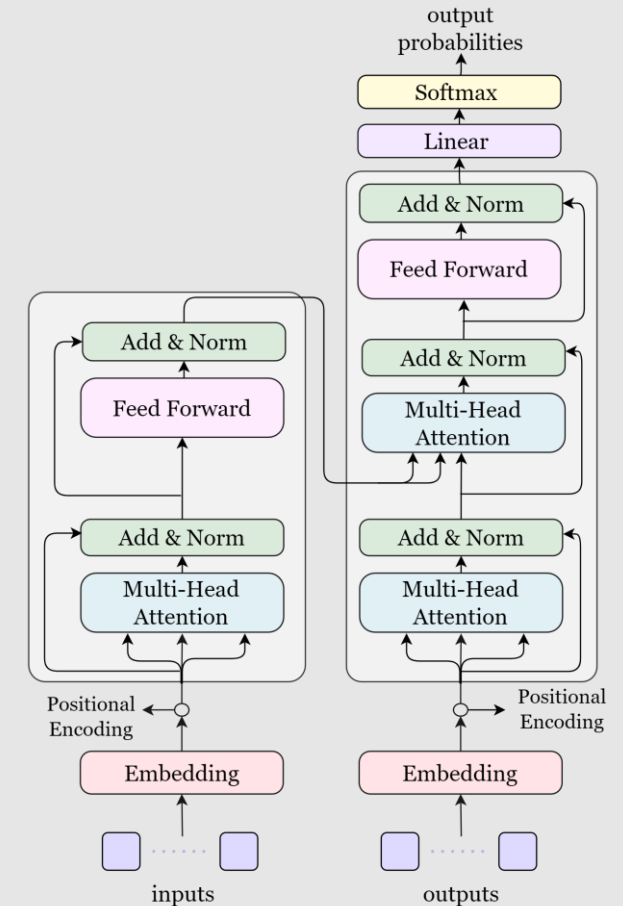
Music composition DL

Objective: Texture and Form

MELODY

Monophonic

Polyphonic

Texture

Melody +
Accompaniment

...

HARMONIZATION

Melody

Harmony

MULTI-INSTRUMENT

Ensemble

Orchestra

...

Carlos Hernandez-Olivan, Jose R. Beltran

Universidad
Zaragoza
1542

Music composition DL: Melody Generation

Monophonic vs Polyphonic melody

Texture: melody with accompaniment (other possibilities: monophonic or polyphonic melody...)

Melody

Harmony   chord   chord progression

Structure   motif   phrase or theme
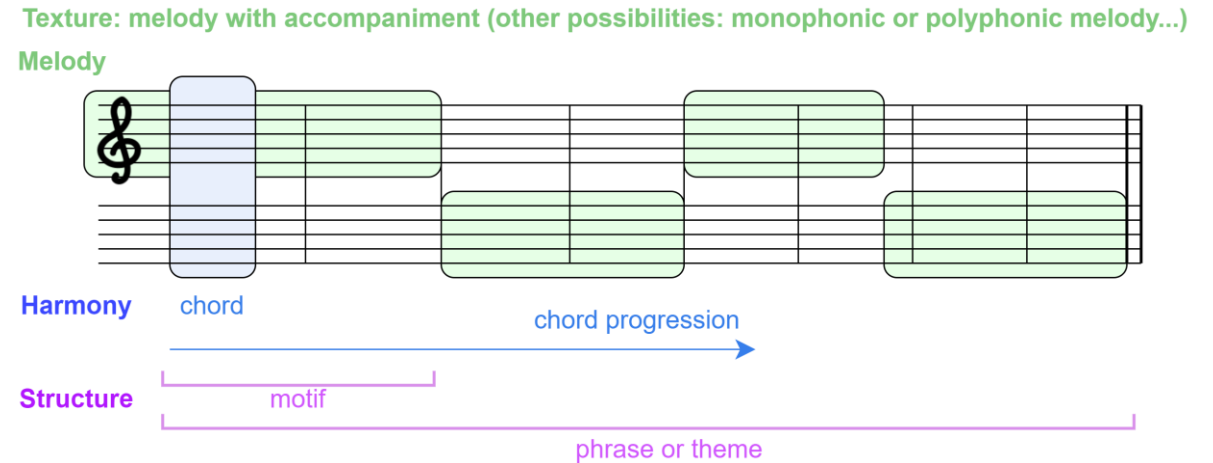
[Eck and Schmidhuber, 2012] A First Look at Music Composition using **LSTM** Recurrent Neural Networks.

▶ Learning Chords

▶ 12-bar blues genre

▶ Learning Melody and Chords

Universidad Zaragoza
1542

Carlos Hernandez-Olivan, Jose R. Beltran

Music composition DL: Melody Generation

[Hadjeres et al., 2017] DeepBach: a Steerable Model for Bach Chorales Generation

> Bach Chorales generation (4 voices) with RNNs + Gibbs sampling (MCMC)

[Huang et al., 2018] Music Transformer: Generating Music with Long-Term Structure

> Piano pieces trained with MAESTRO dataset

Structure Awareness

[Lattner et al., 2016] Imposing higher-level Structure in Polyphonic Music Generation using Convolutional Restricted Boltzmann Machines and Constraints

Music composition DL: Multi-Instrument

[Dong et al., 2017] Muse**GAN**: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment

[Roberts et al., 2018] A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music – "Music**VAE**"

[Ens and Pasquier, 2020] Exploring Conditional Multi-Track Music Generation with the **Transformer**

More: LakhNES...

Universidad Zaragoza

Carlos Hernandez-Olivan, Jose R. Beltran
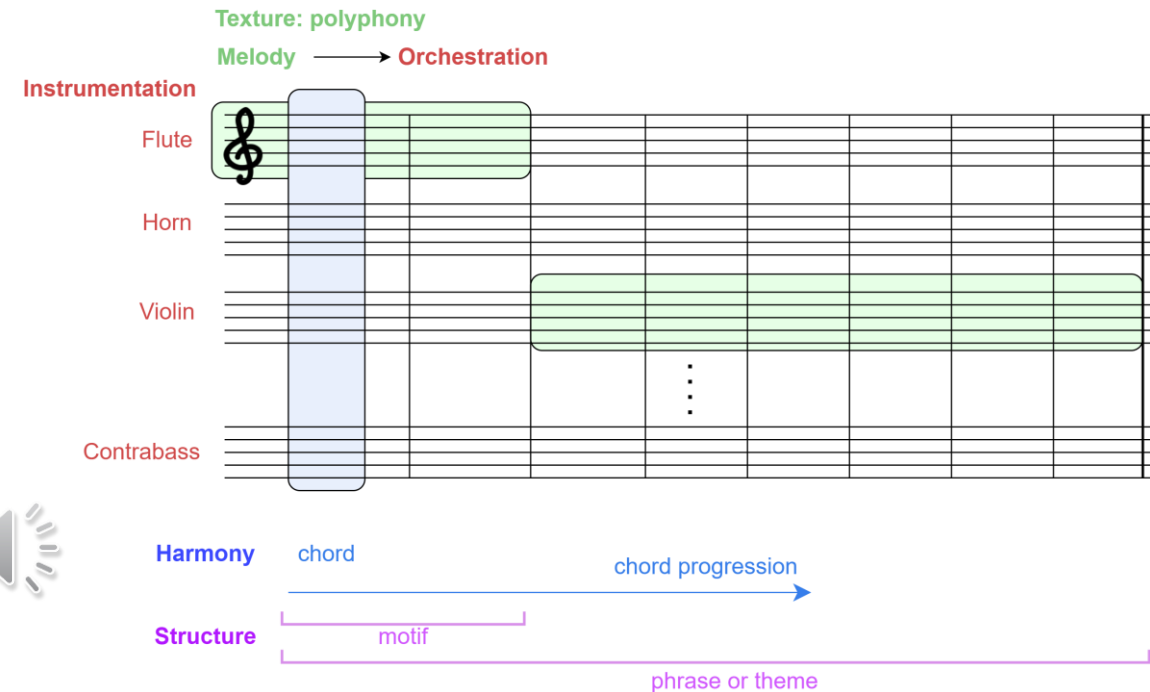
Evaluation

SUBJECTIVE
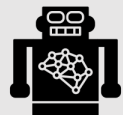
Surveys: User's opinions

Turing test

AI   vs   Human

OBJECTIVE

Feature Extraction

Feature Comparison

Ideally: A combination of both

Evaluation: Subjective

[Hernandez-Olivan et al., 2022] Subjective Evaluation of Deep Learning Models for Symbolic Music Composition.

- Questions for different user levels based on their music knowledge

    - beginner
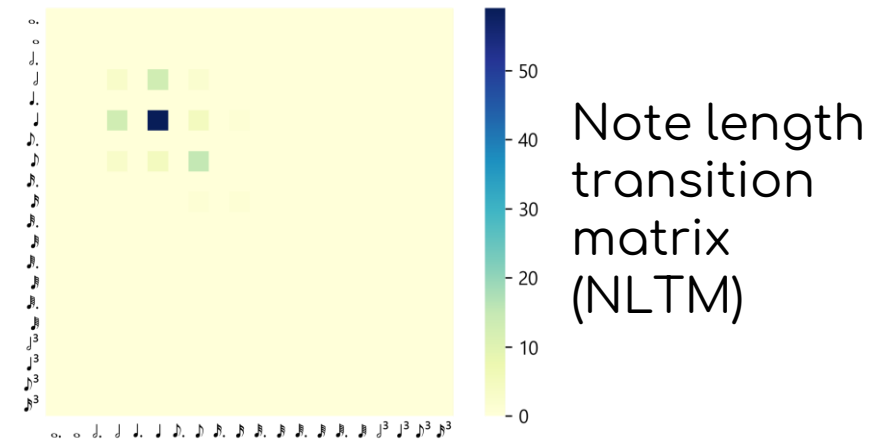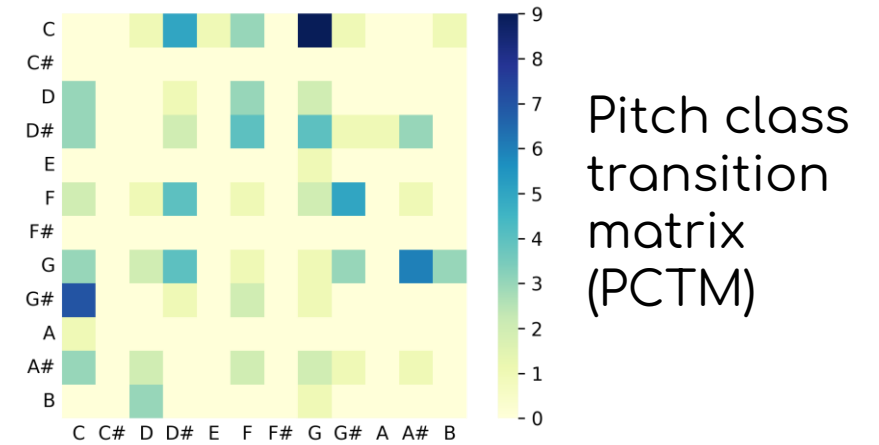
    - intermediate

    - pro

Cons:

- Difficult to reproduce

- Lack of automation

Universidad Zaragoza
1542

Carlos Hernandez-Olivan, Jose R. Beltran

Evaluation: Objective

[Li-Chia Yang and Alexander Lerch, 2019] On the evaluation of generative models in music.
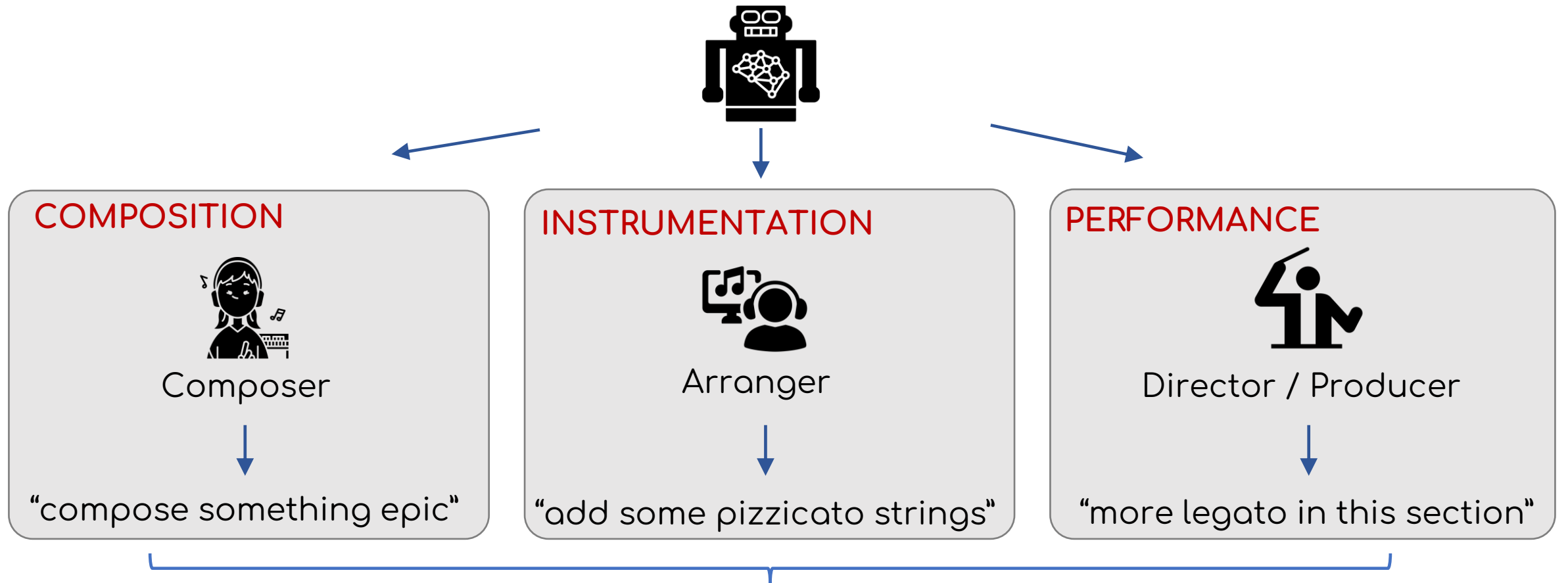
- ▶ Objective measure from music features at different levels (piece, instrument, bar)

- ▶ Inter and intra (Euclidean) distances between generated and input datasets

- ▶ PDF from the distance histograms per feature

- ▶ KLD and Overlapping Area between the PDFs

Pitch class transition matrix (PCTM)

Note length transition matrix (NLTM)

Carlos Hernandez-Olivan, Jose R. Beltran

Universidad
Zaragoza
1542

Future Directions

- Specialize models to learn music and genres

- Objective evaluation

- Structure modelling ⟶ Deal with longer sequences

- Human-Computer interaction ⟶ Inpainting...

- Generation by conditioning

Carlos Hernandez-Olivan, Jose R. Beltran

Universidad
Zaragoza
1542

Future Directions

COMPOSITION

Composer

"compose something epic"

INSTRUMENTATION

Arranger

"add some pizzicato strings"

PERFORMANCE

Director / Producer

"more legato in this section"

Symbolic Music Generation with Performance attributes + Synthesis / Audio Music Generation
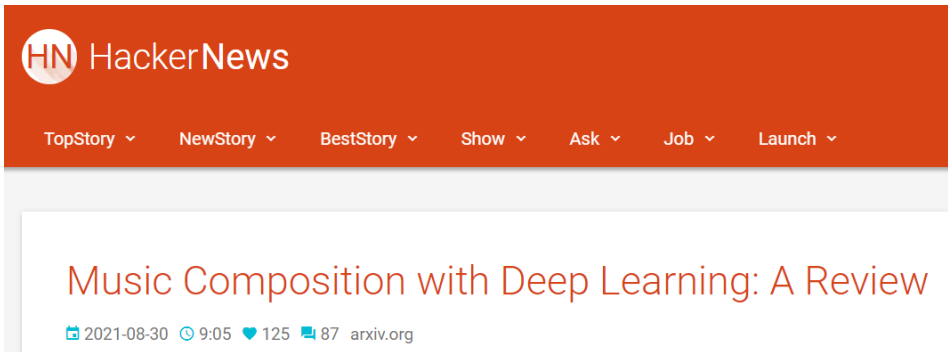
OpenAI DALLE for music? ...

Recap: Open Questions

▶ Are the current DL models capable of generating music with a certain level of creativity?

▶ What is the best NN architecture to perform music composition with DL?

▶ Could end-to-end methods generate entire structured music pieces?

▶ Are the composed pieces with DL just an imitation of the inputs or can NNs generate new music in styles that are not present in the training data?

Universidad
Zaragoza
1542

Carlos Hernandez-Olivan, Jose R. Beltran

Recap: Open Questions

➤ Should NNs compose music by following the same logic and process as humans do?

➤ How much data do DL models for music generation need?

➤ Are current evaluation methods good enough to compare and measure the creativity of the composed music?

➤ Are we (researchers, developers...) designing DL models according to the music basic principles, or are we just taking models from other fields to do the music composition task?

Universidad
Zaragoza
1542

Carlos Hernandez-Olivan, Jose R. Beltran

Survey Paper

Accepted as a book chapter in "Advances in Speech and Music Technology: Computational Aspects and Applications" of Springer book series "Signal and Communication Technology".
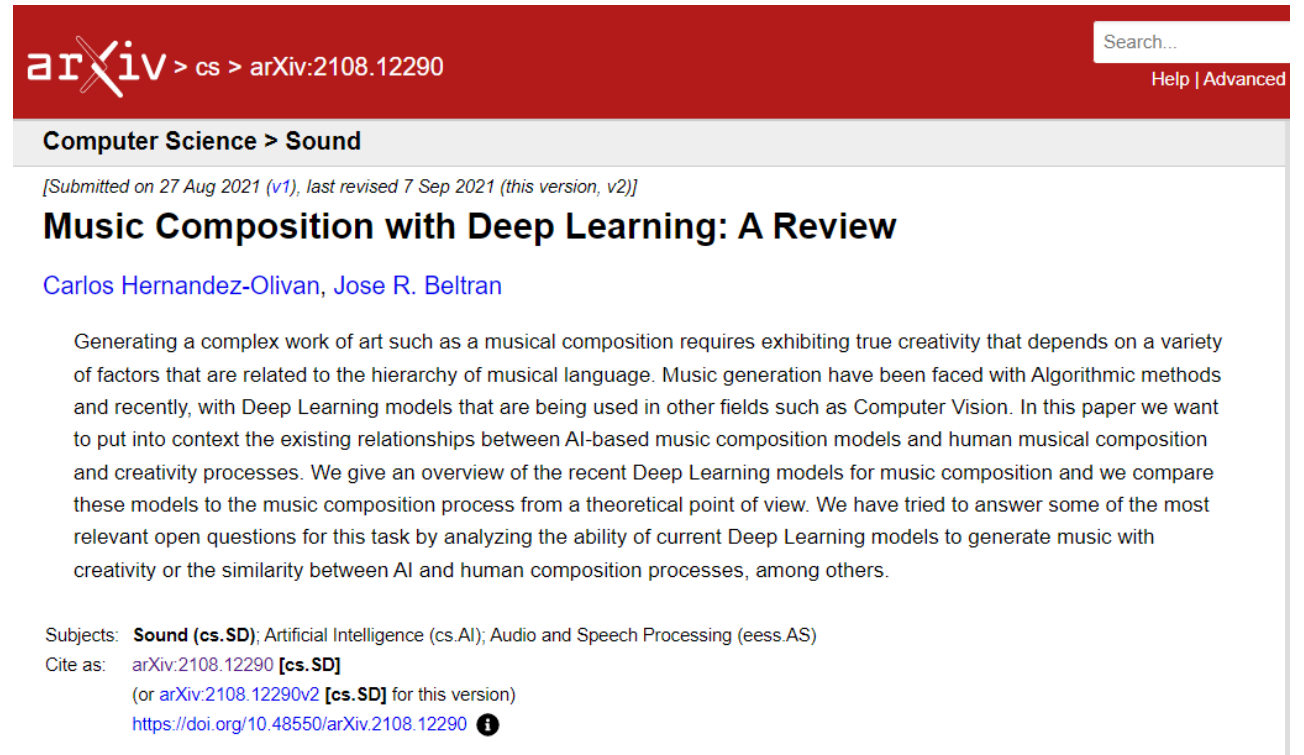
HackerNews

TopStory NewStory BestStory Show Ask Job Launch

Music Composition with Deep Learning: A Review

2021-08-30 9:05 125 87 arxiv.org

https://hacker-news.news/post/28353358

https://github.com/carlosholivan/DeepLearningMusicGeneration

arXiv > cs > arXiv:2108.12290

Search...
Help | Advanced

Computer Science > Sound

[Submitted on 27 Aug 2021 (v1), last revised 7 Sep 2021 (this version, v2)]

Music Composition with Deep Learning: A Review

Carlos Hernandez-Olivan, Jose R. Beltran

Generating a complex work of art such as a musical composition requires exhibiting true creativity that depends on a variety of factors that are related to the hierarchy of musical language. Music generation have been faced with Algorithmic methods and recently, with Deep Learning models that are being used in other fields such as Computer Vision. In this paper we want to put into context the existing relationships between AI-based music composition models and human musical composition and creativity processes. We give an overview of the recent Deep Learning models for music composition and we compare these models to the music composition process from a theoretical point of view. We have tried to answer some of the most relevant open questions for this task by analyzing the ability of current Deep Learning models to generate music with creativity or the similarity between AI and human composition processes, among others.

https://arxiv.org/abs/2108.12290

Acknowledgements

> Spanish Science, Innovation and University Ministry Aragonese Government - AffectiveLab

> Thanks to Jürgen Schmidhuber for his recommendations

Some images of this presentation have been taken from the Noun Project

Universidad
Zaragoza
1542

Carlos Hernandez-Olivan, Jose R. Beltran

Music Composition with Deep Learning:
A Review

Carlos Hernandez-Olivan*, Jose R. Beltran

April 2022

*carloshero@unizar.es