



## АНОТАЦІЯ

Нині комп'ютер є основним засобом зберігання та обробки інформації про музику. Це можуть бути ноти, композиції цілком, метадані про звук. Як наслідок необхідності отриманні інформації про цифровий або аналоговий звук, з'являється велика кількість задач, що пов'язані з обробкою звуку – ідентифікація композиції, пошук схожих композицій та інші.

Метою роботи є розробка методу розпізнавання послідовності акордів в звуковому сигналі, над яким зроблено аналогово-цифрову трансформацію. Цей метод дозволить з високою точністю вказати акорд, що звучить в певний момент часу цього звукового сигналу.

В результаті аналізу та порівняння існуючих методів вирішення задачі, запропонований та обґрунтований метод розпізнавання на основі згортальної нейронної мережі. Реалізована система, що складається з модулів генерації та розмітки даних, навчання нейронної мережі, розмітки WAV-файлів. Розпізнаючий компонент системи – згортальна нейронна мережа. WAV-файли трансформуються в вектори за допомогою віконного семплювання. З семплів та файла розмітки формується набір даних для навчання нейронної мережі. За допомогою цих даних навчається нейронна мережа. На базі рішення нейронної мережі про те, який акорд звучить в даному семплі та алгоритмі згладжування результату розмічаються WAV файли, що відповідають умовам задачі.

Якість результату визначається метрикою асигасу нейронної мережі та фактичною точністю розпізнавання послідовності акордів у WAV-файлі. До системи можна додати більше акордів та відрегулювати параметри векторизації WAV-файла з метою подальшої розробки та покращення результату.

## АННОТАЦИЯ

В настоящее время компьютер является основным средством хранения и обработки информации о музыке. Это могут быть ноты, композиции целиком, метаданные о звуке. Как следствие необходимости получения информации о цифровом или аналоговом звуке появляется значительное количество задач, которые связаны с обработкой звука – идентификация композиции, поиск похожих композиций и другие.

Целью работы является разработка метода распознавания последовательности аккордов в звуковом сигнале, над которым сделано аналогово-цифровую трансформацию. Данный метод позволит с высокой вероятностью указать аккорд, звучащий в определенный момент времени этого звукового сигнала.

В результате исследования и сравнения существующих методов решения задачи, предложен и обоснован метод распознавания на основе сверточной нейронной сети. Реализована система, состоящая из модулей генерирования и разметки данных, обучения нейронной сети, разметки WAV файлов. Распознающий компонент системы – сверточная нейронная сеть. WAV-файлы трансформируются в векторы с помощью оконного семплирования. Затем, из семплов и файла разметки формируется набор данных для обучения нейронной сети. С помощью набора данных тренируется нейронная сеть. На основании решения нейронной сети о том, какой аккорд звучит в этом семпле и алгоритме сглаживания результата размечаются WAV файлы, соответствующие ограничениям задачи.

Оценка качества результата определяется метрикой ассигасы нейронной сети и фактической точностью распознавания последовательности аккордов в WAV файле. В систему можно добавить больше аккордов, отрегулировать параметры векторизации WAV файла с целью дальнейшего исследования и улучшения результата.

## ABSTRACT

Nowadays, computer is a main instrument to store and process sound information. It includes notes, songs, sound metadata. As a consequence of necessity to get information of digital or analog sound, lots of sound-processing tasks appeared: identification, search of similar compositions, etc.

The aim of this thesis is development of the algorithm, which is capable of chord sequence recognition from the sound signal, over which digital-analog transformation has been made. This approach should determine the chord which sounds at certain time of musical composition.

As the result of research and analysis of existing approaches to this task, convolutional neural network-based method is proposed. Implemented system consists of data generation and markup module, neural net training module, WAV files markup module. Solving component of system is convolutional neural network. WAV files are transformed to input vectors with sliding window sampling. Test and train data are created based on samples and markup file. Neural net is trained with that dataset. WAV files that confirm to task restrictions are translated into chords basing on neural net chord recognition and chord wiping algorithm.

Result quality evaluation is defined by accuracy metrics of neural network and actual chord recognition precision in WAV file. System is capable of known chord list expansion, WAV vectorization parameters change to continue research and improve existing result.

## ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ ТА ТЕРМІНІВ .....	6
ВСТУП .....	7
1 ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ .....	9
1.1 Огляд методів розпізнавання акордів .....	9
1.2 Постановка задачі .....	11
2 ПРОЕКТУВАННЯ СИСТЕМИ ДЛЯ РОЗПІЗНАВАННЯ ПОСЛІДОВНОСТІ АКОРДІВ В ЗВУКОВОМУ СИГНАЛІ .....	12
2.1 Формати звукових файлів .....	12
2.2 Перетворення звуку у вхідний вектор нейронної мережи .....	14
2.3 Згортальні нейронні мережі .....	16
2.4 Обрання засобів реалізації .....	19
3 РЕАЛІЗАЦІЯ СИСТЕМИ ДЛЯ РОЗПІЗНАВАННЯ ПОСЛІДОВНОСТІ АКОРДІВ В ЗВУКОВОМУ СИГНАЛІ .....	22
3.1 Опис програмної реалізації .....	22
3.2 Опис роботи програми .....	28
4 ТЕСТУВАННЯ СИСТЕМИ ДЛЯ РОЗПІЗНАВАННЯ ПОСЛІДОВНОСТІ АКОРДІВ В ЗВУКОВОМУ СИГНАЛІ .....	31
4.1 Метод оцінки точності результату .....	31
4.2 Оцінка гіперпараметрів векторизації WAV-файла .....	32
4.3 Результати роботи на різних музикальних колекціях .....	36
4.4 Перспективи участі у змаганнях з цифрового пошуку у музиці .....	41
ВИСНОВКИ .....	43
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ .....	44
ДОДАТОК А. СТРУКТУРА ЗГОРТАЛЬНОЇ НЕЙРОННОЇ МЕРЕЖІ .....	46

## ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ ТА ТЕРМІНІВ

Скорочення:

LSTM – спеціальна архітектура-підтип рекурентної нейронної мережі;

ЗНМ – згортальна нейронна мережа.

НМ – нейронна мережа;

ПЗ – програмне забезпечення;

ШІ – штучний інтелект.

Терміни

MP3 – формат файла, який зберігає зберігати звук із використанням стиснення в частотно-часовому форматі.

WAV – формат файла, який зберігає звук без використання стиснення в частотно-часовому форматі.

MIDI – формат файла, який зберігає цифрову конфігурацію музикальної композиції, зокрема ноти та інформацію про них.

## ВСТУП

Нині комп'ютер є основним засобом зберігання та обробки музики і будь-якої інформації про музику. Це можуть бути ноти, композиції цілком, метаінформація про звук. З необхідністю отримання інформації про цифровий або аналоговий звук впливає велика кількість задач, які пов'язані з обробкою звуку – ідентифікація композиції, пошук схожих композицій, та інші.

Звук [1] можна собі уявити як функцію  $x(t)$ , що показує тиск середовища в даній точці в залежності від часу. Музикальна композиція є сумою окремих звуків, які відповідають нотам і які можна почути у певні моменти часу. Кожен звук має свою частоту, початок та кінець звучання. Живі музикальні інструменти, в свою чергу, мають тембр – набір окремих гармонійних коливань (частотних компонент), які утворюють звук, що відповідає ноті.

Музичний інтервал [2] – відношення висот двох тонів, що визначається відношенням частот їх коливань.

Акорд [3] ритмічно одночасне сполучення кількох (не менше трьох) різних за висотою. Тобто, декілька нот, вибраних за певним правилом, що звучать одночасно. Залежно від кількості різнойменних звуків, що можуть бути розташовані по терціях, акорди поділяються на тризвуки (3 ноти), септакорди (4 ноти), нонакорди (5 нот) та інші.

Задачу розпізнавання послідовності акордів можна поставити наступним чином – нехай маємо звуковий сигнал в певний момент часу і групу акордів, яку необхідно вміти розпізнавати. Необхідно для кожного моменту часу вказати акорд із фіксованої групи акордів, що звучить у даний момент часу.

Система розпізнавання послідовності акордів, що складають певний звук можуть використовуватися, наприклад, для навчання грі на музичному інструменті (автоматичний акомпанемент), при вивченні теорії музики (для ілюстрації залежностей у музичних композиціях). Можливо також застосувати систему розпізнавання акордів для пошуку композицій за їх вмістом.

Задачу розпізнавання послідовності акордів в звукозапису можна поставити таким чином: нехай задані звуковий сигнал  $x(t)$ ,  $t \in [t_{start}, t_{end}]$  і безліч можливих назв акордів  $Y$ . Необхідно для кожного моменту часу  $t \in [t_{start}, t_{end}]$  вказати акорд  $y \in Y$ , що звучить в цей момент.

Метою роботи є розробка методу розпізнавання послідовності акордів у звуковому сигналі над яким зроблено аналогово-цифрову трансформацію. З цієї точки зору, в роботі буде розглянуто розпізнавання акордів у цифровому сигналі, що містить частотно-часову конфігурацію звуку. Цілком цифрові файли музичних композицій – наприклад, MIDI файли містить конфігурацію нот та моментів часу, в які ноти повинні звучати.

Для реалізації мети необхідно вирішити такі задачі:

- 1) розглянути методи розпізнавання послідовності акордів у звуковому сигналі із файлів;
- 2) реалізувати систему для розпізнавання послідовності акордів у звуковому сигналі;
- 3) виконати тестування системи і оцінити якість розпізнавання.

Архітектура НМ не передбачатиме конфігурації для розпізнавання певного типу акордів, проте за результатами тестування буде визначено, які акорди одночасно розпізнаватиме система. Зокрема, буде протестована можливість одночасно розпізнавати тризвучні акорди та чотиризвучні септакорди.

## ВИСНОВКИ

Під час виконання роботи розглянуті та проаналізовані з використанням сучасних знань існуючі методи розпізнавання акордів у звуковому сигналі. За результатом аналізу запропонована і реалізована самодостатня система, яка реалізує рішення до поставленої задачі.

Система включає в себе генератор вхідних даних для навчання. Дані – WAV файл, у якому звучать акорди живого інструмента та файл розмітки звуку. Реалізований перетворювач (векторизатор) WAV файла так, щоб можливе була подальша робота у системі. Запропонований і реалізований метод віконного (з перекриттям) вибирання семплів зі звукового сигналу. ЗНС побудована із використанням сучасних знань для рішення поставленої задачі. Розроблений модуль транскрипції WAV файлів, які відповідають обмеженням задачі до послідовності акордів, яка в них звучить.

Натуральне звучання інструмента досягається завдяки використанню записів інструмента (технологія soundfont) та спеціального синтезатора (fluidsynth). Звуки акордів, на яких проводилось навчання і тестування, містять живі обертони, що характерні для натуральних інструментів і виражають самотність інструменту серед інструментів групи, до якої він належить.

Тестування системи показує, що система справляється із поставленою задачею. Про це свідчить можливість навченої системи запам'ятати велику кількість акордів, розпізнавати акорди близькі за звучанням, потенціал системи отримати ще кращий результат, можливість в рамках пакету утворити нові дані, використовувати доповнення даних, щоб рішення поставлені задачі.

Точність системи розпізнавання акордів становить 97.4% згідно з метрикою CSR (для набору з 36 акордів однієї октави). З використанням алгоритму трансформації згладжуванням, розпізнавання на тестових файлах (згенерованих на тому ж генераторі) становить близько 100%.

В результаті, система вирішує поставлену задачу і дає гідний результат приймаючи до уваги обмеження, що були описані разом із задачею. Модель готова для інтеграції у програми навчання музикальній теорії. Її можна застосувати як рішення підзадачі задачі транскрипції звукових файлів.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Что такое звук? [Электронный ресурс] – Режим доступа: <http://allfrets.ru/studiotheory/whatisound> – 29.09.2018
2. Интервали | Музична абетка [Электронный ресурс] – Режим доступа: <http://muzabetka.com.ua/interval.html> – 2.10.2018
3. Аккорды в музыке и их виды [Электронный ресурс] – Режим доступа: <http://muz-teoretik.ru/akkordy-v-muzyke> – 5.10.2018
4. Fujishima Takuya. Realtime chord recognition of musical sound: A system using common lisp music // Proc. ICMC. – Vol. 1999. – 1999. – P. 464–467.
5. De Haas W. Bas, Magalhaes Jos ~ e Pedro, Wiering Frans. Improving Audio Chord ´ Transcription by Exploiting Harmonic and Metric Knowledge // Proceedings of the 13th International Society for Music Information Retrieval Conference. – Porto, Portugal, 2012. – October 8-12. – [Электронный ресурс] – Режим доступа: <http://ismir2012.ismir.net/event/papers/295-ismir-2012.pdf> – 10.02.2019
6. Ian Goodfellow and Yoshua Bengio and Aaron Courville Deep Learning Book. MIT Press, 2016.
7. Гвоздев. В.Д., Шпинарева И.М. Распознавание аккордов в аналоговом звуковом сигнале с помощью нейронных сетей / Информатика, інформаційні системи та технології: тези доповідей шістнадцятої всеукраїнської конференції студентів і молодих науковців. Одеса, 19 квітня 2019р. – Одеса, 2019. – с.189-191
8. Структура WAV файла. [Электронный ресурс] – Режим доступа: <https://audiocoding.ru/articles/2008-05-22-wav-file-structure/> – 1.03.2019
9. Pychord documentation. [Электронный ресурс] – Режим доступа: <https://pychord.readthedocs.io/en/latest/> – 10.03.2019
10. Midiutil documentation. [Электронный ресурс] – Режим доступа: <https://www.emergentmusics.org/midiutil> – 14.03.2019

11. Fluidsynth | Software synthesiser based on Soundfont specification. [Електронний ресурс] – Режим доступу: <http://www.fluidsynth.org/> – 11.03.2019
12. Appendix G. SoundFont2 File Format. [Електронний ресурс] – Режим доступу: <https://csound.com/docs/manual/MiscSf2.html> – 15.03.2019
13. MIREX 2019: Audio Chord Estimation [Електронний ресурс] – Режим доступу: <https://www.musicir.org/mirex/wiki/> – 20.03.2019
14. Classification and Loss Evaluation – Softmax and Cross Entropy Loss [Електронний ресурс] – Режим доступу: <https://deeppnotes.io/softmax-crossentropy> – 20.03.2019
15. Building a Bayesian deep learning classifier. [Електронний ресурс] – Режим доступу: <https://towardsdatascience.com/building-a-bayesian-deep-learning-classifier-ece1845bc09> – 5.04.2019
16. Гвоздев. В.Д., Шпинарева И.М. Распознавание аккордов в аналоговом звуковом сигнале с помощью нейронных сетей / Информатика, інформаційні системи та технології: тези доповідей п'ятнадцятої всеукраїнської конференції студентів і молодих науковців. Одеса, 27 квітня 2018р. – Одеса, 2019. – с.38