

ПРОГРАМНІ КОМПОНЕНТИ СИСТЕМИ АУДІОСИНТЕЗУ ТА РОЗПІЗНАВАННЯ МУЗИЧНИХ ЗВУКІВ

Вінницький національний технічний університет

Анотація

Розглянуто основні програмні складові системи аудіосинтезу та розпізнавання музичних звуків. Ключовим компонентом є модуль аудіообробки, що реалізує основну логіку системи а також використовує нейронну мережу для забезпечення одного з декількох підходів генерації музичних послідовностей.

Ключові слова: аудіосинтез, користувацький інтерфейс, нейронна мережа.

Abstract

Primary software components of the system of audio analysis and music recognition are considered. The key component is an audio handling module. It implements the core logic of a system and uses a neural network to provide one of the multiple approaches for music sequence generation.

Keywords: audio synthesis, user interface, neural network.

Вступ

Система аудіосинтезу та розпізнавання музичних звуків реалізує комбінований підхід для синтезу музичних послідовностей, що подана клієнтськими додатками на базі операційних систем Android та iOS. Серед основного функціоналу системи є можливість створення власних музичних композицій з використанням декількох підходів. Перший – надає можливість відігравати бажану мелодію, награну голосом, за допомогою обраного музичного інструмента. Крім того, такі вхідні дані можна також використовувати як шаблон для мелодії, згенерованої за допомогою штучного інтелекту на її основі. Маючи результат, що поданий на двовимірній площині, користувач має змогу ручного редагування та паралельного використання вищевказаних підходів [1].

Результати дослідження

Основним компонентом розроблюваної системи є модуль аудіообробки, що реалізує основний функціонал генерації та розпізнавання музичних композицій [2]. Один з варіантів створення композицій дозволяє використовувати награні голосом мелодії як шаблон для синтезу схожих послідовностей. Для його імплементації використовується нейронна мережа, розроблена з використанням фреймворку Tensorflow.

Іншим ключовим компонентом розроблюваної системи є користувацький інтерфейс. Оскільки такий програмний продукт, у першу чергу, націлений на мобільні платформи, такі як Android та iOS, важливим є використання правильного набору технологій для забезпечення задовільного досвіду користування додатком. Існує низка доступних технологічних рішень на ринку для реалізації сучасного користувацького інтерфейсу для мобільних додатків. Одним з таких рішень для платформи Android є Jetpack Compose, фреймворк для побудови декларативного користувацького інтерфейсу. Основною відмінністю та перевагою декларативного підходу в програмуванні є те, що на відміну від імперативного, де відбувається модифікування мутованого стану, рендеринг інтерфейсу відбувається в результаті контрольного потоку в коді. Таким чином, можна уникнути велику кількість помилок та підвищити читабельність програмного коду. Користувацький інтерфейс у такому випадку описується у вигляді лінійного переліку команд, а стан додатку визначається на ходу, з виконанням if-else конструкцій та циклів. Аналогічно SwiftUI реалізує цей підхід для додатків на базі платформи iOS. Крім того, існує кросплатформне рішення Flutter, що забезпечує аналогічний функціонал, надаючи змогу написання єдиної кодової бази, що працює на обох мобільних платформах.

Структурна схема розроблюваної системи зображена на рисунку 1.

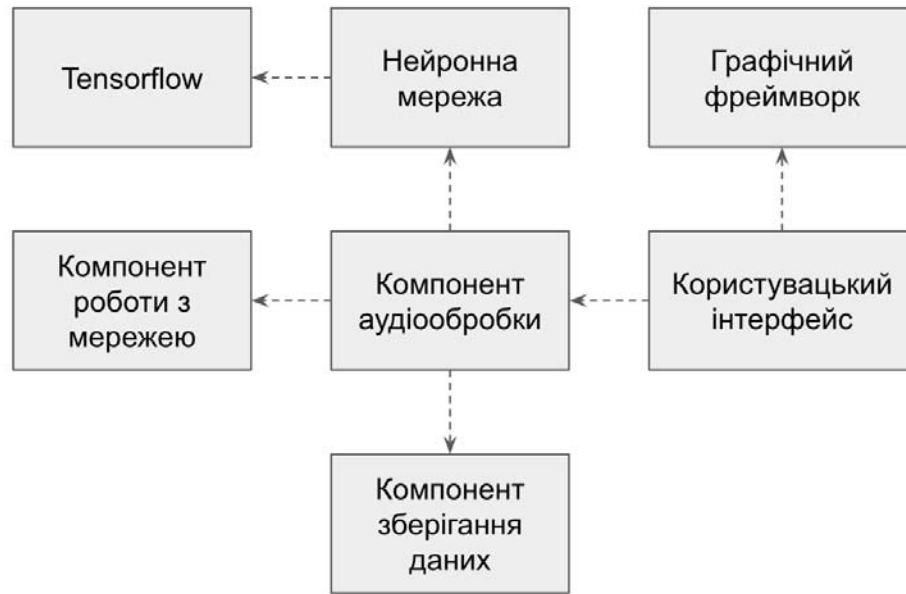


Рис.1. Структурна схема розроблюваної системи

Двовимірна площина, що відображає відношення висоти нот у часі, потребує реалізації окремого компонента користувацького інтерфейсу. Крім того, такий функціонал повинен забезпечувати високу швидкодію, а тому перевага повинна бути надана саме нативним рішенням та фреймворкам для побудови інтерфейсів. Таким чином, обрано нативні рішення Jetpack Compose та SwiftUI для розробки візуального подання системи.

Висновки

Таким чином, було розглянуто основні структурні компоненти системи аудіосинтезу та аналізу музичних композицій. Така система надає можливість генерувати музичні послідовності, відіграні за допомогою обраних музичних інструментів на основі награних мелодій голосом. Основними складовими такої системи є компонент аудіообробки, що реалізує логіку аудіогенерації, а також використовує нейронну мережу, створену з використанням Tensorflow. Крім того, слід приділити особливу увагу користувацькому інтерфейсу, що повинен досягти високої швидкодії при взаємодії з двовимірною площиною редагування музичних послідовностей.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Viktoriia V. Voitko, Svitlana V. Bevez, Sergii M. Burbelo, Pavlo V. Stavytskyi, Oleksandr M. Khoshaba, Natalia O. Rysynets, Olena Yu. Teplova, Andrzej Smolarz, Saule Smailova, Assel Mussabekova, Bakhyt Yerallyeva, "Analysis of the development approaches of the system of audio synthesis and recognition with the option of using photonic processors," Proc. SPIE 12040, Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2021, 120400N (3 November 2021); <https://doi.org/10.1117/12.2611464>
2. Войтко В. В., Бевз С.В., Бурбело С.М., Ставицький П.В. Підходи до реалізації компонента аудіосинтезу системи синтезу та розпізнавання музичних звуків / Електронні інформаційні ресурси: створення, використання, доступ. Пам'яті О.П.Стахова. Збірник матеріалів МНПІК 9-10 листопада 2021 р. – Суми/Вінниця: НІКО/ВНТУ. 2021. – 223 с. – ISBN 978-617-7422-16-6. – С. 40-44.

Войтко Вікторія Володимирівна – кандидат технічних наук, доцент кафедри програмного забезпечення, Вінницький національний технічний університет, м. Вінниця, email: dekanfki@i.ua.

Ставицький Павло Валерійович – аспірант кафедри програмного забезпечення, Вінницький національний університет, м. Вінниця, email: pavlo.stavytskyi@gmail.com.

Viktoriia Voitko – Ph.D., Associate Professor of Software Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: dekanfki@i.ua.

Pavlo Stavytskyi – postgraduate student, Vinnytsia National Technical University, Vinnytsia, email: pavlo.stavytskyi@gmail.com.