

# АНАЛІЗ ВИКОРИСТАННЯ ЗГОРТКОВОЇ НЕЙРОМЕРЕЖІ ДЛЯ ОПТИЧНОГО РОЗПІЗНАВАННЯ МУЗИЧНОЇ ТРАНСКРИПЦІЇ

Вінницький національний технічний університет;

## *Анотація*

*Робота аналізує застосування згорткової нейромережі для оптичного розпізнавання музичної транскрипції, описує етапи розпізнавання та підготовки зображення.*

**Ключові слова:** нейромережа, музична транскрипція, оптичне розпізнавання, згорткова нейромережа.

## *Abstract*

*The work analyzes the application of a convolutional neural network for the optical recognition of musical transcription, describes the stages of image recognition and preparation.*

**Keywords** neural network, musical transcription, optical recognition, convolutional neural network.

## Вступ

Транскрипція музичної партитури у машинночитаний формат може бути виконана вручну. Однак складність нотної нотації неминуче призводить до розробки важкодійного програмного забезпечення для редагування музичних партитур, що ускладнює весь процес, роблячи його часо- та ресурсо-затратним та схильним до помилок. Таким чином, автоматизовані системи транскрипції для музичних документів представляють цікаві інструменти. Галузь, присвячена вирішенню цієї задачі, відома як оптичне розпізнавання музики (ОРМ) [1]. Зазвичай система ОРМ отримує зображення нотного партитури та автоматично експортує її в якусь символічну структуру, таку як MEI або MusicXML. Процес розпізнавання вмісту музичної партитури є складним, і тому робочий процес системи ОРМ є дуже обширним [2]. Дослідження фокусується на розпізнаванні вмісту, який з'являється на окремому персональному рядку (наприклад, ноти для скрипки, флейти тощо), аналогічно більшості досліджень розпізнавання тексту, які спрямовані на розпізнавання слів на зображенні рядка. Це не повинно бути проблемою, оскільки існують успішні алгоритми як для виділення персональних рядків, так і для розділення музики та тексту (супровідний текст)[3].

Для вирішення цієї конкретної задачі було запропоновано використання згорткових нейромереж (CNN), які успішно застосовуються у багатьох послідовних завданнях розпізнавання, таких як емоції або текст.

## Результати дослідження

Звичайною практикою є передобробка зображення для полегшення подальшого аналізу. У цьому випадку ми маємо справу з фотографіями, зробленими камерою, які часто мають певні види спотворень, таких як коливання рівня освітлення та перспективні деформації. Нижче ми перераховуємо всі процедури передобробки, які застосовуються до вхідного зображення:

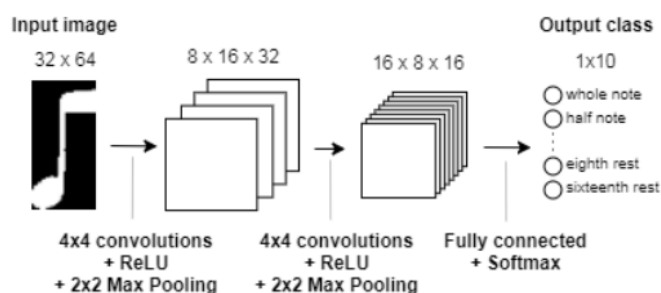
1. Бінаризація
2. Корекція перспективи
3. Виявлення місцезнаходження
4. Виявлення музичних символів
5. Сегментація заголовків нот

Підхід полягає у тому, щоб окремо подавати нейронній мережі індивідуальні ноти та символи для їх визначення, а потім відновлювати музику, використовуючи цю інформацію. З цієї причини ноти, з'єднані вертикальними лініями, потрібно розділяти. Для досягнення цього всі чорні нотні крапки виявляються, і у випадку, якщо дві або більше з них знаходяться на одній лінії(з попереднього кроку), коробка розділяється горизонтально, рівномірно ділячи всі нотні крапки. Для виявлення чорних нотних голівок застосовується операція морфологічного відкриття до зображення

з еліптичним ядром, знаходяться всі з'єднані компоненти, і визначається, які з них є нотними крапками за деякими характеристиками (розмір, відношення сторін, площа та близькість до лінії).

Після того, як кожний символ був сегментований, розпочинається етап розпізнавання. Спочатку вертикальні лінії визначаються за їхніми особливими властивостями (дуже мала ширина і дуже мале відношення сторін), кінцева вертикальна лінія також визначається окремо. Щодо інших символів, вони надсилаються до нейронної мережі для розпізнавання. У наступних підрозділах ми пояснимо архітектуру нейронної мережі та спосіб її навчання.

Враховуючи успіх, ми використовували згорткову нейронну мережу вони виконували багато завдань із розпізнавання графіки. Ми вибрали просту загальну архітектуру, оскільки класифікація завдання не дуже вимогливе. У майбутньому це може змінитися було проведено більше тестувань. Зараз лише 10 класів розглядається: ноти і паузи від цілого до шістнадцятого. В найближче майбутнє, більше символів, таких як випадковості, ключі та час підписи будуть додані. Схема архітектури мережіпоказано на рисунку 1.



Рисунку 1 – Схема архітектури згорткової нейромережі

Розроблена згорткова нейронна мережа була навчена за допомогою 2500 зразків нотних творів.

## Висновки

Використання згорткових нейронних мереж у роботі забезпечило високу точність розпізнавання, що виявлено під час оцінювання моделі на реальних даних. Робота відкриває перспективи для подальших досліджень у сфері оптичного розпізнавання музичних записів та можуть сприяти подальшому розвитку автоматизованих систем обробки музичної транскрипції.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. D. Bainbridge and T. Bell, "The challenge of optical music recognition," *Computers and the Humanities*, vol. 35, no. 2, pp. 95–121, 2001.
2. A. Rebelo, I. Fujinaga, F. Paszkiewicz, A. R. S. Marc, al, C. Guedes, and J. S. Cardoso, "Optical music recognition: State-of-the-art and open issues," *IJMIR*, vol. 1, no. 3, pp. 173–190, 2012.
3. J. A. Burgoyne, Y. Ouyang, T. Himmelman, J. Devaney, L. Pugin, and I. Fujinaga, "Lyric extraction and recognition on digital images of early music sources," in *ISMIR*, 2009, pp. 723–727.

**Хібовський Денис Олегович** — студент групи 1KN-20б, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: alekseystepka5589@gmail.com.

Науковий керівник: **Колесницький Олег Константинович** — к.т.н., доцент, професор кафедри комп'ютерних наук.

**Khibovsky Denys Olehovych** - student of group 1KN-20b, Faculty of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: alekseystepka5589@gmail.com.

Scientific adviser: **Oleh Kostiantynovych Kolesnytskyi** — candidate of technical sciences, associate professor, professor of the Computer Sciences.Department.