

Ярченко Д. С.,
здобувач вищої освіти спеціальності 122 Комп'ютерні науки
Пархоменко О. Ю., к.ф.-м.н., доцент кафедри економічної кібернетики,
комп'ютерних наук та інформаційних технологій
Миколаївський національний аграрний університет
м. Миколаїв

ПІДХОДИ ДО АВТОМАТИЧНОЇ КЛАСИФІКАЦІЇ ПІДЖАНРІВ ЕЛЕКТРОННОЇ МУЗИКИ (TECHNO, HOUSE, TRANCE) З ВИКОРИСТАННЯМ MFCC-ОЗНАК ТА АРХІТЕКТУР ГЛИБОКОГО НАВЧАННЯ

Стрімке зростання обсягів цифрових аудіоколекцій та потокових сервісів робить автоматичну категоризацію музики критично важливим завданням. Якщо класифікація широких категорій (рок, джаз, класика) наразі досягла високих показників якості, то розрізнення піджанрів всередині одного стилю залишається відкритим викликом. Особливої складності задачі надає висока варіативність композицій всередині жанру та розмитість меж між суміжними стилями.

Електронна танцювальна музика (EDM) є типовим прикладом такої проблеми. Такі піджанри, як Techno, House та Trance, мають спільне коріння та використовують подібний інструментарій, проте суттєво відрізняються за темпоритмічною структурою, щільністю аранжування та частотними характеристиками. Традиційні методи, що покладаються на ручне аналізування, є надто трудомісткими. На противагу цьому, сучасні підходи на основі глибокого навчання, зокрема згорткові нейронні мережі (CNN), здатні автоматично виявляти складні спектральні закономірності.

Метою роботи є аналіз та узагальнення існуючих методів автоматичної класифікації піджанрів електронної музики та визначення перспективних напрямів підвищення точності моделей. У роботі також окреслено підхід до підвищення ефективності класифікації шляхом комбінування спектральних ознак та використання сучасних архітектур глибокого навчання.

Мел-кепстральні коефіцієнти (MFCC) залишаються одними з найпоширеніших акустичних ознак у задачах музичної інформатики. Їх популярність зумовлена здатністю компактно описувати короткостроковий спектр потужності, що імітує нелінійне сприйняття висоти людським вухом. Оскільки звукова палітра Techno, House та Trance значною мірою формується синтезованими тембрами (барабанні машини, басові партії, синтезаторні педи), MFCC виявляються особливо чутливими до таких параметрів, як атака звуку, спектральний центроїд та нахил форманти – характеристик, що утворюють впізнаване "забарвлення" жанру.

Для Techno, що характеризується щільною партією великого барабана, мінімалістичними басовими лініями та шумовими ефектами, MFCC-профілі демонструють високу щільність низькочастотних компонентів та відносно стабільний кепстральний патерн. Натомість для House музики, з її свінговим грувом, багатшими гармоніями та вокальними семплами, характерна більша

варіабельність коефіцієнтів, що відповідають за середні частоти. Trance вирізняється протяжними хвилями синтезаторних підкладок (pads), яскраво вираженими мелодичними лініями та суттєвими перепадами спектральної енергії, що може бути формалізовано через аналіз динаміки зміни MFCC у часі.

Втім, використання лише MFCC-векторів для розрізнення цих піджанрів стикається з обмеженнями, адже деякі композиції на перетині стилів (наприклад, Tech House або Progressive Trance) можуть утворювати майже ідентичні акустичні відбитки. Для подолання цього дослідники все частіше комбінують MFCC із додатковими спектральними ознаками або використовують їх як вхідні дані для глибоких архітектур.

Протягом останнього десятиліття методологія класифікації музики зазнала суттєвої трансформації. Ранні роботи спиралися на алгоритми опорних векторів (SVM) та випадкового лісу (RF), які використовували усереднені MFCC-ознаки. Результати дослідження на основі класифікації EDM (2022) показали, що випадковий ліс здатен досягти точності близько 83% на тестовій вибірці, випереджаючи SVM, однак цей показник суттєво варіюється залежно від якості попередньої обробки даних.

Сучасний етап характеризується домінуванням глибокого навчання, зокрема згорткових нейронних мереж (CNN). На відміну від методів, що потребують ручного конструювання ознак, CNN здатні навчатися ієрархічним представленням безпосередньо зі спектрограм або матриць MFCC. У роботі за 2025 рік показано, що модель CNN із декількома згортковими шарами та регуляризацією досягає точності 95%, що значно перевершує RF (77%) при класифікації широких жанрів [3]. Основні характеристики сучасних підходів до класифікації музичних жанрів узагальнено в табл. 1.

Таблиця 1. Порівняльний аналіз сучасних підходів до класифікації музичних жанрів

Підхід (Модель)	Джерело (Рік)	Ознаки	Точність	Специфіка застосування
CNN	(2024)	MFCC (матриця)	98%	Класифікація 10 жанрів (GTZAN)
CNN + ВНО Optimization	(2024)	MFCC + STFT	95,7%	Гібридний підхід, підвищена здатність до узагальнення
RF	(2022)	Нормалізовані ознаки	83%	Фокус на EDM (танцювальній музиці)
CNN (Adagrad)	(2023)	MFCC	75%	Специфічний набір даних корінної музики
CNN (Оптимізована)	(2023)	MFCC (3-сек. фрагменти)	72%	Класифікація за початковими секундами треку

Попри високі абсолютні показники точності на загальних наборах даних, застосування CNN до суміжних піджанрів EDM є більш складним завданням. Тонкі відмінності в аранжуванні потребують аналізу довготривалих залежностей у сигналі, що ставить під сумнів ефективність стандартних двовимірних згорток на коротких фреймах. Вузькоспеціалізовані дослідження показують, що на

обмежених і специфічних наборах даних точність CNN може падати до позначки 75%, що свідчить про значний вплив розміру, збалансованості та репрезентативності навчальної вибірки.

Ключова наукова проблема при розрізненні Techno, House та Trance полягає у високому ступені їх семантичної подібності. На відміну від задачі "рок проти класики", де відмінності в інструментарії очевидні, EDM піджанри можуть використовувати ідентичні звукові пресети, а відмінності часто полягають у ритмічних нюансах або формі хвилі басової партії.

Аналіз літератури вказує на те, що для підвищення дискримінативності моделі необхідно відходити від унімодальних MFCC-представлень. Перспективним напрямом видається об'єднання MFCC з ознаками на основі короткочасного перетворення Фур'є (STFT) та використання паралельних ансамблів CNN, де гіперпараметри оптимізуються під конкретний набір ознак. Крім того, оскільки EDM характеризується чіткою ритмічною сіткою, доцільним є застосування механізмів уваги (attention) для виокремлення темпорально значущих ділянок треку – моментів вступу басу або зміни малюнка ударних.

Проблема нестачі великих, професійно розмічених датасетів саме за піджанрами EDM залишається гострою, що відкриває широке поле для застосування методів навчання з перенесенням (transfer learning) та аугментації даних. У подальших дослідженнях планується використання відкритих музичних датасетів (зокрема GTZAN) для експериментальної перевірки запропонованих підходів.

Проведений огляд показує, що поєднання MFCC-ознак зі згортковими нейронними мережами є ефективним підходом у задачах класифікації музики. Водночас для розрізнення близьких піджанрів електронної музики (Techno, House, Trance) класичні підходи потребують удосконалення. Перспективними напрямками є використання комбінованих спектрально-часових представлень, механізмів уваги та методів навчання з перенесенням. Практичне значення роботи полягає у можливості застосування отриманих результатів при розробці систем автоматичної класифікації музичного контенту.

Список використаних джерел

1. Deep Learning-Based Music Genre Classification using Convolutional Neural Network / S. Mohanapriya et al. 2024 *First International Conference on Software, Systems and Information Technology (SSITCON)*, Tumkur, India, 18-19 October 2024. 2024. P. 1-6. URL: <https://doi.org/10.1109/ssitcon62437.2024.10796579>.

2. Z. Dai and X. Huang. Electronic Dance Music Classification Based on Machine Learning Methods. 2022 *International Conference on Electronics and Devices, Computational Science (ICEDCS)*.Marseille. France. 2022. pp. 351-354, URL: <https://doi.org/10.1109/ICEDCS57360.2022.00084>.

3. Olojede, Ojo Abraham, Stephen Olatunde, Olabiyisi, Oluwaseun. O Alo. Performance Evaluation of Some Machine Learning Models for Music Genre Classification. *International Journal of Latest Technology in Engineering Management & Applied Science*. 2025. Vol. 14, no. 2. P. 18-24. URL: <https://doi.org/10.51583/ijltemas.2025.1402003>.