

Голові разової
спеціалізованої вченої ради
Національного університету
«Львівська політехніка»
д.т.н., професору Віталію ЯКОВИНІ

РЕЦЕНЗІЯ

кандидата технічних наук, асистента

ПШЕНИЧНОГО Олександра Юрійовича

на дисертаційну роботу **ПОБЕРЕЙКА Петра Богдановича**

**«МЕТОДИ ТА ЗАСОБИ АНАЛІЗУ ВІДЕОПОТОКІВ
ТА ПОШУКУ ПОДІБНОСТЕЙ У КОНТЕНТНО-ОРІЄНТОВАНИХ
СИСТЕМАХ ВІДЕОІНФОРМАЦІЇ»,**

подану до разової спеціалізованої вченої ради
на здобуття ступеня доктора філософії в галузі 12 – Інформаційні технології
за спеціальністю 122 – Комп'ютерні науки

Актуальність теми дисертаційного дослідження. Дисертаційна робота присвячена актуальній і важливій науковій проблемі – розробленню методів та засобів аналізу відеопотоків і пошуку подібного контенту в контентно-орієнтованих системах відеоінформації. У роботі обґрунтовано необхідність ефективного аналізу та керування стрімко зростаючими обсягами відеоданих. Автор ретельно проаналізував сучасні підходи до контентного аналізу відео та методи пошуку схожих фрагментів, виявивши їх переваги й недоліки. Дисертація містить дослідження існуючих алгоритмів обробки відеопотоків, а також опис підходів до класифікації та індексації відеоконтенту для забезпечення швидкого пошуку за змістом. На основі проведеного аналізу запропоновано вдосконалену методику пошуку

подібностей у відео, що поєднує класичні методи обробки відеоданих із сучасними засобами глибинного навчання. Автор розробив інтегровану систему аналізу відео, провів її програмну реалізацію та експериментальне випробування для перевірки ефективності. Отримані результати демонструють підвищення точності пошуку та продуктивності аналізу відеопотоків, що має широкі перспективи практичного застосування і подальшого розвитку. В цілому, робота є завершеним самостійним науковим дослідженням, результати якого обґрунтовують доцільність впровадження запропонованих методів аналізу відео та пошуку подібностей у сучасних інформаційних системах, що працюють з відеоданими.

Зв'язок роботи з науковими програмами, планами, темами. У дисертаційній роботі вирішено актуальну науково-технічну задачу підвищення ефективності пошуку інформації у великих обсягах відеоданих за рахунок вдосконалення методів аналізу відеопотоків. У контексті стрімкого зростання обсягів відеоінформації в різних сферах – від систем відеоспостереження та мас-медіа до наукових досліджень і розважальних платформ – існує гостра потреба в ефективних засобах аналізу, пошуку та організації відеоданих. Відповідно, надзвичайно актуальним є завдання швидкого й точного пошуку потрібних відеоматеріалів за змістом, особливо в умовах обмежених обчислювальних ресурсів та великого обсягу даних. Запропоноване дисертантом рішення спрямоване на подолання цих викликів шляхом застосування прогресивних технологій комп'ютерного зору та штучного інтелекту для автоматизації аналізу відео. Використання глибоких нейронних мереж для розпізнавання та зіставлення фрагментів відео є новітнім підходом, що дозволяє адаптувати процес пошуку до особливостей контенту і значно підвищити його ефективність. Дана дисертаційна робота вирізняється комплексним підходом: поєднано аналіз існуючих методів контентного аналізу відео з розробкою інноваційної

системи, яка інтегрує методи глибинного навчання (зокрема, згорткові нейронні мережі та аналіз сцен) для покращення результатів. Більше того, в роботі не просто описано можливі рішення, а запропоновано конкретні, обґрунтовані та експериментально перевірені методи підвищення точності пошуку подібностей у відео. Результати дослідження мають широке практичне значення для різноманітних галузей (медіа-сервіси, відеоспостереження, цифрові архіви тощо), що підкреслює високу актуальність обраної теми для сучасної науки та практики.

Оцінка наукових результатів дисертації. Наукова цінність даного дослідження полягає в цілому ряді аспектів:

- Вдосконалення методів аналізу відеопотоків. Розроблено інтегрований підхід до аналізу відеоданих, який поєднує класичні алгоритми обробки відео із сучасними методами глибинного навчання. Це дозволило підвищити якість і точність виявлення та ідентифікації подібних відеофрагментів, а також зменшити час обробки, що має велике практичне значення при роботі з великими масивами відеоінформації.
- Удосконалення моделей та алгоритмів пошуку подібностей. У ході роботи було розроблено й формалізовано обчислювальні моделі для індексації відео та пошуку схожих фрагментів. Зокрема, запропоновано метод конструювання багаторівневих векторів ознак відео (що включають просторові та часові характеристики), які дозволяють більш повно описати вміст відео. Ці моделі стали основою для створення ефективного алгоритму пошуку подібностей у відеопотоках.
- Інтеграція глибинних нейронних мереж у систему пошуку. Удосконалено підходи до аналізу відеоконтенту шляхом використання сучасних глибоких нейронних мереж для екстракції ознак та класифікації відеофрагментів. Інтеграція згорткових

нейронних мереж та методів аналізу сцен (Scene Boundary Detection) відкриває нові можливості для автоматизації та оптимізації пошуку: забезпечено врахування як статичних ознак зображення, так і динаміки сцени, що підвищило точність і повноту пошуку.

- Експериментальне підтвердження та нові знання. У процесі дослідження автор отримав нові дані щодо ефективності різних нейронних архітектур (наприклад, використано LSTM, SlowFast для аналізу відео) та методів їх комбінування при пошуку подібностей. Розвинуто наукові уявлення про баланс між точністю пошуку та обчислювальними витратами при аналізі відеопотоків. Отримані результати можуть слугувати основою для подальших досліджень у галузі контентного пошуку відео.

Новизна отриманих результатів та можливість їх застосування в інших областях (наприклад, для аналізу відеоданих у суміжних сферах) роблять дане дослідження важливим і цінним для розвитку науки.

Практичні результати роботи, їх рівень та ступінь впровадження.

Практичне значення результатів дисертаційного дослідження полягає у підвищенні ефективності обробки та керування відеоінформацією: покращенні точності пошуку схожого відеоконтенту, зменшенні часу і ресурсів, необхідних на аналіз великих відеопотоків, та вдосконаленні процесів організації відеоданих у контентно-орієнтованих системах. Запропоновані в роботі рішення дозволяють автоматизувати рутинні операції аналізу відео, мінімізувати людський фактор при пошуку потрібних фрагментів і забезпечити більш ефективну роботу інформаційних систем, що працюють з відеоконтентом. Це особливо важливо для практичного застосування в таких сферах, як відеоспостереження (оперативний пошук подій на відео), медіабібліотеки

та стримінгові сервіси (пошук дублікатів або схожих матеріалів), система управління контентом тощо.

Такі практичні результати ґрунтуються на:

- Розробці концепції та архітектури інформаційної системи аналізу відеопотоків, яка враховує різноманітні параметри відеоданих (формати, роздільну здатність, тривалість сцен тощо) для ефективної індексації та пошуку. Здійснено проектування функціональної архітектури системи, що включає модулі сегментації відео на сцени, виділення ознак за допомогою нейронних мереж та пошуку подібностей за векторними представленнями відеофрагментів.
- Реалізації інтегрованої методики пошуку на основі глибинних нейронних мереж. Здійснено обґрунтований вибір і навчання нейронної мережі для аналізу відеоконтенту, зокрема використано згорткові нейронні мережі для виділення ознак зображення та рекурентні/інші мережі для врахування часової складової. Розроблений алгоритм пошуку подібностей використовує одержані ознаки для зіставлення фрагментів відео та формування списку релевантних результатів.
- Удосконаленні математичних моделей індексації та пошуку. Запропоновано модель представлення відеофрагментів у вигляді багатовимірних векторів ознак, а також критерії порівняння таких векторів для визначення ступеня схожості. Впроваджено алгоритмічні оптимізації (наприклад, використання додаткової нейронної мережі для контролю якості пошуку), що підвищують швидкодію та точність системи.
- Застосуванні методів машинного навчання та комп'ютерного зору для аналізу відеопотоків, що відкриває нові перспективи для автоматизації цього процесу, підвищення контрольованості

якості контенту та запобігання можливим проблемам в управлінні відеоданими (наприклад, виявлення дублікатів чи нерелевантного контенту, який перевантажує систему).

Результати експериментального впровадження розробленої системи показали її високу ефективність. Це відкриває можливості для широкого застосування запропонованих методів у практичних умовах – від корпоративних систем керування медіаконтентом до сервісів потокового відео та відеоархівів. Розроблені у ході дослідження методики і засоби можуть бути також використані в освітньому процесі закладів вищої освіти для підготовки висококваліфікованих спеціалістів у галузі комп'ютерних наук, інформаційних технологій та аналізу даних.

Повнота викладення результатів досліджень в опублікованих працях. Наукові положення та висновки дисертаційної роботи є теоретично обґрунтованими та підтвердженими експериментальними даними, отриманими в процесі виконання дослідження, що свідчить про їхню цілісність і логічність. Автор використав сучасний математичний апарат і загальноприйняті наукові методи (експериментальні дослідження, комп'ютерне моделювання, статистичний аналіз результатів) для підтвердження працездатності запропонованих методів. Всі етапи дослідження, від формулювання моделі до реалізації програмного прототипу, документовані й аргументовані.

Апробація результатів роботи на наукових конференціях та семінарах (у тому числі міжнародних) підтверджує практичну цінність та актуальність отриманих результатів. Основні положення і результати дисертації опубліковано у 6 наукових працях, з яких: 2 статті – у наукових фахових виданнях України, 3 – у міжнародних виданнях, 1 – у матеріалах конференцій. Це свідчить про високий рівень достовірності та визнання результатів дисертаційного дослідження науковою спільнотою.

Оцінка достовірності та обґрунтованості основних положень і висновків дисертації. Сформульовані в дисертаційній роботі положення й висновки є належно аргументованими та достовірними. Вони спираються на сучасні методології глибинного навчання, комп'ютерного бачення та контентно-орієнтованого пошуку відео. Для їх верифікації здобувач провів серію експериментів на відкритих датасетах UCF101 і HMDB51, а також на сформованій власній колекції відеофрагментів, що дало можливість оцінити ефективність розроблених моделей у різних сценаріях.

Отримані результати переконливо демонструють підвищення точності й продуктивності пошуку відеофрагментів порівняно з базовими підходами. Зокрема, інтеграція моделі DCNN + LSTM з адаптивним налаштуванням забезпечила стабільну роботу навіть за змін формату та якості відео. Достовірність знайдених рішень підтверджено їх успішним упровадженням у навчальний процес і виробничу практику компанії, що засвідчено відповідними актами.

Отже, висновки дисертації логічно впливають із поставлених завдань, базуються на репрезентативних експериментальних даних і засвідчують високу наукову надійність та вагому практичну цінність отриманих результатів.

Оцінка змісту, оформлення й обсягу дисертації. Основний зміст дисертаційної роботи викладено на 141 сторінці. Загальний обсяг роботи – ≈ 170 сторінок. Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел (91 найменування) та додатків. За структурою, мовою та стилем дисертація оформлена правильно, відповідно до чинних вимог МОН України. Виклад матеріалу є послідовним, логічним і науково аргументованим.

У вступі обґрунтовано актуальність теми дисертації, визначено мету і основні завдання дослідження, окреслено об'єкт і предмет дослідження,

наведено методи дослідження, а також описано наукову новизну та практичне значення отриманих результатів.

У першому розділі проаналізовано інформаційні джерела за тематикою дисертації та сучасні алгоритми пошуку у відеопотоках. Дисертант здійснив огляд і класифікацію підходів до аналізу та індексації відеоданих, розглянув методи обробки візуальних і часових характеристик відео. Проведений аналіз дозволив виявити недоліки існуючих методів та сформулювати вимоги до розробки власного підходу.

У другому розділі представлено розробку обчислювальних моделей для індексації відеоінформації. Докладно обґрунтовано вибір архітектури глибокої нейронної мережі для аналізу характеристик відеоконтенту. Запропоновано модель представлення відеоданих, що включає виділення ознак за допомогою нейронних мереж та структурування цих ознак у формат, придатний для швидкого пошуку. Сформовано концепцію багаторівневої індексації відео, яка враховує різні типи ознак (візуальні, часові) та їх вагомість.

У третьому розділі розроблено безпосередньо модель конструювання векторів ознак та пошуку подібностей у відео. Наведено методику автоматизованого виділення ключових кадрових сцен (SBD – визначення меж сцен) та отримання для них ознак за допомогою згорткових нейронних мереж і рекурентних підходів. Синтезовано алгоритми формування комбінованих векторів ознак, що описують кожну відеосцену, та алгоритм порівняння цих векторів для виявлення схожих фрагментів. Таким чином реалізовано ядро системи контентного пошуку відео, яке забезпечує зіставлення відеофрагментів за змістом.

У четвертому розділі описано розроблену архітектуру програмної системи та проведено експериментальну апробацію результатів. Представлено повноцінну інформаційну систему для аналізу відеопотоків, включаючи модулі завантаження відео, сегментації на сцени, обчислення

векторів ознак і пошуку схожих фрагментів. Наведено результати експериментів щодо ефективності запропонованого підходу: виміряно час обробки, точність (Precision), повноту (Recall) та обчислювальні витрати для різних конфігурацій нейронних мереж. Отримані дані демонструють переваги розробленої системи над традиційними методами (зокрема, покращення показників Precision/Recall при помірних затратах ресурсів). У розділі також обговорено можливості впровадження системи на практиці та надано рекомендації щодо подальшого удосконалення методики (наприклад, адаптація під реальний час, розширення набору ознак, оптимізація під конкретні апаратні платформи).

У загальних висновках дисертаційної роботи підсумовано основні результати дослідження. Сформульовано висновки, що повністю відповідають поставленій меті та завданням, підтверджують досягнення наукової новизни та вирішення актуальної наукової проблеми. Висновки відображають усі основні положення, отримані в кожному розділі, та вказують на перспективи подальшої роботи у цьому напрямі.

Зауваження до дисертаційної роботи.

- Доцільно більш детально описати процес навчання нейромережі FNN та механізм її коректувальної дії на основну мережу DCNN+LSTM, оскільки ефект такої корекції дуже значний: підвищення точності із 66% до 86%.
- Для більш строгого обґрунтування користі від додавання LSTM-шару до DCNN мережі варто було б квантифікувати ефект по аналогії із аналізом ефективності FNN.
- У розділі 2.3 згадується, як виділення просторових і тимчасових характеристик відео дозволяє відслідковувати об'єкти та їх взаємодії, і наводиться приклад, де у відео про спортивні події визначаються ключові моменти: голи, паси. Проте в подальшому, у розділі 2.4.1 описано, що виявлення ключових кадрів відбувається на основі

гістограм кольорів у просторі HSV. Варто детальніше пояснити, як саме векторне представлення сцен та взаємодій між об'єктами використовується при визначенні ключових кадрів.

- Для покращення масштабованості та надійності запропонованої системи рекомендується вдосконалити схему розгортання сервісів у Kubernetes для підвищення захисту від збоїв фізичного обладнання.

Висновки про відповідність дисертації встановленим вимогам.

Дисертаційна робота Поберейка Петра Богдановича на тему «Методи та засоби аналізу відеопотоків та пошуку подібностей у контентно-орієнтованих системах відеоінформації» є завершеною науковою працею, яка має важливе значення для розвитку методів контентного аналізу відеоданих та вдосконалення систем управління відеоінформацією шляхом впровадження сучасних інформаційних технологій. У роботі вирішено актуальну наукову задачу створення та впровадження нових методів аналізу відеопотоків і пошуку подібностей, що дозволяє ефективніше працювати з великими масивами мультимедійних даних. Отримані дисертантом результати сприяють підвищенню точності та швидкодії контентного пошуку відео, тим самим робота робить вагомий внесок у галузь комп'ютерних наук та інформаційних технологій.

Представлена дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки» (галузь знань 12 «Інформаційні технології») за своїм змістом, структурою, обсягом, науковою новизною та практичним значенням відповідає паспорту спеціальності 122 «Комп'ютерні науки», а також вимогам «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження доктора філософії», затвердженого постановою Кабінету Міністрів України від 12 січня 2022 р. № 44 (зі змінами). Таким чином,

Поберейко Петро Богданович заслуговує присудження наукового ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки».

Рецензент:

кандидат технічних наук, асистент,
кафедри систем штучного інтелекту
Національного університету
«Львівська політехніка»

Олександр ПШЕНИЧНИЙ

Підпис к. т. н., асистента
Олександра ПШЕНИЧНОГО засвідчую.
Вчений секретар
Національного університету
«Львівська політехніка»
к. т. н., доцент



Роман БРИЛИНСЬКИЙ