

ВІДГУК ОФІЦІЙНОГО ОПОНЕНТА

доктора технічних наук, професора, професора кафедри штучного інтелекту,
навчально-наукового Інституту прикладного системного аналізу

Національного технічного університету України

«Київський політехнічний інститут імені Ігоря Сікорського»

Литвиненка Володимира Івановича

на дисертаційну роботу Швандта Максима Альбертовича

на тему

**«МОДЕЛІ ТА МЕТОДИ ДЕТЕКТУВАННЯ, ТРЕКІНГУ ТА НАСКРІЗНОЇ
ІДЕНТИФІКАЦІЇ ЛАБОРАТОРНИХ ТВАРИН У ВІДЕОДАНИХ»,**

представлену на здобуття наукового ступеня доктора філософії за спеціальністю

122 – Комп'ютерні науки, галузь знань 12 – Інформаційні технології

1 АКТУАЛЬНІСТЬ ТЕМИ ДИСЕРТАЦІЇ

Актуальність обраної теми дисертаційної роботи Швандта М. А. зумовлена потребою біологічних і медико-біологічних досліджень у достовірному, тривалому та мінімально інвазивному спостереженні за поведінкою лабораторних тварин. Відеофіксація дає змогу накопичувати значні обсяги об'єктивних даних, проте їх ручне опрацювання потребує великих часових витрат, залежить від кваліфікації спостерігача та ускладнює відтворюваність експериментів. Тому автоматизація виділення тварин у кадри, побудови їхніх траєкторій і збереження індивідуальної ідентичності є важливою передумовою переходу від епізодичних спостережень до кількісного аналізу поведінки.

Специфіка цієї задачі полягає в тому, що лабораторні відеозаписи часто отримують у складних для комп'ютерного зору умовах: фон змінюється, у кадри виникають відблиски, тіні й артефакти, а тварини мають близькі зовнішні ознаки, взаємно перекриваються, зливаються на зображенні та надалі розділяються. Навіть локальна помилка детектування може накопичуватися на наступних етапах оброблення і спричиняти розриви траєкторій або підміну ідентифікаторів. Отже,

окреме вдосконалення детектора без його узгодження з процедурами трекінгу та повторної ідентифікації не забезпечує необхідної надійності результатів.

У дисертаційній роботі запропоновано комплексний підхід до розв’язання зазначеної проблеми: від сегментації відеокадрів і формування масок об’єктів до асоціації треклетів у межах цілісного конвеєра. Поєднання удосконаленого методу SLIC, нейромережевої моделі Mask R-CNN із подвійною магістраллю екстракції ознак і концептуальної моделі Mask-TAUDL, орієнтованої на наскрізну ідентифікацію без ручного присвоєння ID-міток, відповідає сучасним тенденціям інтелектуального аналізу відеоданих. Актуальність дослідження зумовлена також необхідністю практичної реалізації розроблених моделей і методів, що в дисертаційній роботі забезпечується створенням спеціалізованого програмного застосунку для оброблення реальних лабораторних відеозаписів.

2 ЗВ’ЯЗОК РОБОТИ З НАУКОВИМИ ПРОГРАМАМИ, ПЛАНАМИ, ТЕМАМИ

Дисертаційну роботу виконано відповідно до планів науково-дослідної діяльності Національного університету «Одеська політехніка» у 2024–2026 рр. у межах держбюджетної теми № 264-73 «Методи та технології моделювання і забезпечення якості складних програмних систем у динамічних середовищах» (номер державної реєстрації 0125U003655), що підтверджує зв’язок дослідження з актуальними науковими завданнями університету.

3 ОЦІНКА ОБҐРУНТОВАНOSTІ НАУКОВИХ ПОЛОЖЕНЬ, ВИСНОВКІВ І РЕКОМЕНДАЦІЙ

Наукові положення, висновки та рекомендації, сформульовані в дисертаційній роботі Швандта М. А., є належно обґрунтованими. Дослідження побудовано послідовно: спочатку автор виявляє проблемні чинники лабораторних відеоданих, далі обґрунтовує вибір засобів сегментації, детектування і трекінгу, після чого переходить до розв’язання задачі збереження ідентичності об’єктів. Така логіка забезпечує змістовний зв’язок між поставленими завданнями,

запропонованими моделями та методами й одержаними результатами.

Обґрунтованість одержаних результатів підтверджується не лише теоретичним аналізом, а й порівняльними експериментами. Для удосконаленого SLIC показано приріст якості локалізації, а для Mask R-CNN із подвійною магістраллю ResNet18 і керованим SE-злиттям наведено покращення mAP50, Precision та Recall відносно базової конфігурації. Важливо, що оцінювання охоплює як рамки детекцій, так і сегментаційні маски, тобто відповідає багатокomпонентному характеру поставленої задачі.

Обґрунтованість результатів додатково підтверджується перевіркою інтегрованого конвеєра на різних за природою відеопослідовностях – синтетичній сцені та реальних записах із мишами й бичками. Використані показники сукупно характеризують стабільність ідентифікаторів, часову узгодженість масок і фрагментацію траєкторій. Відповідність кількісних результатів візуальним прикладам і результатам програмної реалізації дає підстави вважати основні висновки роботи належно підтвердженими.

4 ДОСТОВІРНІСТЬ НАУКОВИХ ПОЛОЖЕНЬ, РЕКОМЕНДАЦІЙ ТА ВИСНОВКІВ

Достовірність одержаних у дисертації результатів забезпечується обґрунтованим вибором методів дослідження, використанням кількісних критеріїв оцінювання та послідовною експериментальною перевіркою запропонованих рішень. Автор оцінює якість суперпіксельної сегментації, детектування та маскової сегментації, а також часову узгодженість траєкторій і стабільність збереження ідентичності об'єктів. Одержані значення показників зіставлено з результатами базових конфігурацій, а інтегрований конвеєр перевірено на синтетичній сцені та реальних відеозаписах із лабораторними тваринами. Узгодженість кількісних метрик із візуальними результатами та наявність програмної реалізації підтверджують обґрунтованість і достовірність сформульованих у роботі висновків.

5 НАУКОВА НОВИЗНА РЕЗУЛЬТАТІВ, ВИКЛАДЕНИХ У ДИСЕРТАЦІЇ

До основних наукових результатів дисертації слід віднести наступне:

1. Удосконалено метод суперпіксельної сегментації SLIC за рахунок використання багатопланового простору ознак, який базується на LAB-моделі зображення, враховує різницеве зображення та результати перевірки однорідності пікселів у локальному оточенні, що дозволило підвищити якість локалізації лабораторних тварин у відеоданих за умов нестаціонарного фону та наявності артефактів.
2. Удосконалено метод детектування лабораторних тварин у відеоданих шляхом використання методу суперпіксельної сегментації SLIC і контекстної фільтрації, що дозволило підвищити якість детектування лабораторних тварин за умов нестаціонарного фону.
3. Удосконалено нейромережеву модель Mask R-CNN за рахунок введення подвійної магістралі екстракції ознак із керованим SE-злиттям ($B \rightarrow A$), що дозволило підвищити якість детектування та сегментації лабораторних тварин у відеоданих за умов часткових перекриттів, злиття та подальшого розділення об'єктів зі збереженням прийнятної оперативності.
4. Вперше запропоновано концептуальну модель наскрізної ідентифікації лабораторних тварин у відеоданих без використання ручного маркування Mask-TAUDL, яка інтегрує механізм отримання ознак на основі удосконаленої нейромережевої моделі Mask R-CNN із подвійною магістраллю екстракції ознак, масково-зважені ROI-ознаки та механізми асоціації треклетів без використання ручного маркування TAUDL, що забезпечує формування нормалізованих векторів вбудовувань і стабільне збереження ідентичності лабораторних тварин протягом відеопослідовності.

6 ПРАКТИЧНЕ ЗНАЧЕННЯ ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

Наукова значущість одержаних результатів полягає у розвитку моделей комп'ютерного зору, орієнтованих не лише на покадрове виявлення об'єктів, а й на збереження їхньої ідентичності в часі. Запропоноване поєднання

багат шарової суперпiксельної сегментації, Mask R-CNN iз подвійною магістраллю екстракції ознак та Mask-TAUDL формує цілісний підхід до опрацювання відеопослідовностей за наявності оклюзій, злиття й подальшого розділення об'єктів, а також значної візуальної подібності тварин. Одержані результати мають також ширше методологічне значення для розв'язання задач багатоб'єктного трекінгу у контрольованих, але динамічно змінних сценах.

Практична значущість дослідження підтверджується створенням спеціалізованого застосунку, який забезпечує автоматизовану обробку лабораторних відеозаписів, відновлення траєкторій руху, оцінювання активності та аналіз контактів між тваринами. Розроблені рішення використовуються під час проведення лабораторних досліджень на біологічному факультеті Одеського національного університету імені І. І. Мечникова, а результати роботи впроваджено в освітній процес цього університету та Національного університету «Одеська політехніка».

7 ПОВНОТА ВИКЛАДУ РЕЗУЛЬТАТІВ У ПУБЛІКАЦІЯХ ТА ДОТРИМАННЯ АКАДЕМІЧНОЇ ДОБРОЧЕСНОСТІ

Результати дисертаційного дослідження достатньо повно представлено в наукових публікаціях. За темою дисертації підготовлено та опубліковано 12 наукових праць: 6 статей у наукових фахових виданнях, серед яких 3 публікації представлені у виданнях, що індексуються міжнародною наукометричною базою Scopus, а також 6 праць апробаційного характеру, оприлюднених у матеріалах наукових і науково-практичних конференцій.

У публікаціях послідовно висвітлено ключові складові дисертаційного дослідження: сегментацію та детектування лабораторних тварин, удосконалення нейромережових засобів локалізації, формування траєкторій і підходи до наскрізного збереження ідентичності об'єктів у відеопослідовностях. Для праць, виконаних у співавторстві, у дисертації конкретизовано особистий внесок здобувача.

Таким чином, оприлюднені матеріали відображають основні наукові положення, експериментальні результати та практичні розробки дисертації.

Кількість публікацій, повнота висвітлення в них результатів дисертаційного дослідження та характер видань відповідають вимогам пункту 8 Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії, затвердженого постановою Кабінету Міністрів України від 12.01.2022 р. № 44.

8 АНАЛІЗ ФОРМИ ТА ЗМІСТУ ДИСЕРТАЦІЙНОЇ РОБОТИ

Зміст дисертації відповідає поставленій меті та охоплює всі основні складові заявленої науково-практичної задачі. Побудова роботи є послідовною: від визначення вимог до автоматизованого відеоспостереження за лабораторними тваринами автор переходить до удосконалення окремих методів детектування й сегментації, розроблення засобів збереження ідентичності та об'єднання отриманих рішень у завершений програмно-алгоритмічний конвеєр.

Перший розділ має аналітико-постановочний характер. У ньому систематизовано класичні та нейромережеві підходи до детектування, трекінгу й повторної ідентифікації об'єктів, розглянуто специфіку лабораторних сцен та чинники, що спричиняють фрагментацію траєкторій і перемикання ідентифікаторів. Окремо сформульовано критерії вибору базових засобів комп'ютерного зору з урахуванням точності, швидкодії та можливості локального навчання, після чого визначено мету і задачі дослідження.

У другому розділі досліджено можливості ненеуромережевого детектування лабораторних тварин за умов складного фону. Автор описує процедури попередньої обробки, віднімання фону та формування різницевих представлень, а також удосконалює метод SLIC шляхом залучення багат шарового простору ознак і перевірки локальної однорідності. Контекстний аналіз у межах суперпікселів та подальша фільтрація дали змогу зменшити вплив артефактів і підвищити якість локалізації більш ніж на 6 % порівняно з базовим варіантом.

Третій розділ присвячено нейромережевій локалізації та сегментації екземплярів. Запропонована модифікація Mask R-CNN використовує дві магістралі ResNet18 для опрацювання базових і додаткових каналів та кероване

SE-злиття ознак у напрямі $B \rightarrow A$. Наведено процедуру навчання, налаштування і порівняння конфігурацій моделі; отримані кількісні показники підтверджують доцільність обраної архітектури за часткових перекриттів, злиття й подальшого розділення тварин.

У четвертому розділі розглянуто найскладнішу складову роботи – підтримання сталої ідентичності об'єктів у часі без ручного присвоєння ID-міток. Концептуальна модель Mask-TAUDL поєднує сегментовані треклети, масково-зважені ROI-ознаки, нормалізовані вектори вбудовувань та механізми TAUDL для асоціації фрагментів траєкторій. Запропоноване рішення орієнтоване на роботу за оклюзій і короточасних пропусків детектування та допускає використання в однокамерних і багатокамерних сценаріях.

У п'ятому розділі всі розроблені компоненти узгоджено в інтегрованому конвеєрі детектування, трекінгу та наскрізної ідентифікації. Описано взаємодію допоміжної сегментаційної та нейромережевої гілок, механізм арбітражу детекцій, побудову треклетів і відновлення траєкторій. Результати експериментів на синтетичній послідовності та реальних відеозаписах із мишами й бичками оцінено за метриками часової узгодженості, збереження ідентичності та безперервності спостереження. Розділ завершується описом спеціалізованого застосунку та результатів практичної апробації.

9 ЗАУВАЖЕННЯ ДО ДИСЕРТАЦІЙНОЇ РОБОТИ

По дисертаційній роботі Швандта М. А. можна зробити низку зауважень, які не знижують загальної наукової цінності дослідження, але можуть бути враховані в подальшій роботі автора.

1. У роботі недостатньо повно розкрито питання верифікації інтегрованого конвеєра на відеопослідовностях з еталонною покадровою розміткою траєкторій та ідентифікаторів об'єктів.

2. У роботі недостатньо досліджено чутливість інтегрованого конвеєра до вибору порогових значень, що використовуються під час детектування, міжкадрового зіставлення та завершення траєкторій.

3. У роботі недостатньо уваги приділено оцінюванню часу обробки та обчислювальної складності повного інтегрованого конвеєра.

4. У роботі недостатньо уваги приділено порівнянню інтегрованого конвеєра з поширеними методами багатооб'єктного трекінгу та повторної ідентифікації об'єктів.

5. У роботі недостатньо повно наведено кількісне оцінювання внеску допоміжної сегментаційної гілки, механізму арбітражу та модуля Mask-TAUDL у загальну якість роботи інтегрованого конвеєра.

6. У роботі недостатньо повно розкрито питання експериментальної перевірки реалізації запропонованої концептуальної моделі наскрізної ідентифікації Mask-TAUDL у багатокамерному сценарії.

7. У тексті дисертації мають місце окремі орфографічні, граматичні та стилістичні помилки, на які вказано автору.

Зазначені зауваження мають переважно рекомендаційний характер і можуть стати підґрунтям для подальших досліджень. Вони не зменшують наукової та практичної цінності дисертаційної роботи, яка в цілому виконана на високому рівні.

10 ЗАГАЛЬНІ ВИСНОВКИ ТА ОЦІНКА ДИСЕРТАЦІЇ

Дисертаційна робота Швандта Максима Альбертовича «Моделі та методи детектування, трекінгу та наскрізної ідентифікації лабораторних тварин у відеоданих» за своїм змістом відповідає спеціальності 122 – Комп'ютерні науки. Дисертація є завершеною науково-дослідною роботою, у якій отримано нові науково обґрунтовані результати щодо побудови комплексного конвеєра аналізу лабораторних відеозаписів. Запропонований конвеєр поєднує детектування, сегментацію, формування траєкторій та наскрізну ідентифікацію тварин, а його працездатність підтверджено програмною реалізацією та результатами експериментальної апробації.

Робота відповідає вимогам до дисертацій на здобуття наукового ступеня доктора філософії, а саме вимогам пунктів 6, 7, 8 і 9 Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої

ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії, затвердженого Постановою КМУ від 12.01.2022 р. №44, а здобувач Швандт Максим Альбертович заслуговує присудження наукового ступеня доктора філософії за спеціальністю 122 – Комп'ютерні науки.

Офіційний опонент,
доктор технічних наук, професор,
професор кафедри штучного інтелекту
навчально-наукового Інституту
прикладного системного аналізу
Національного технічного університету України
«Київський політехнічний
інститут імені Ігоря Сікорського»

Володимир ЛИТВИНЕНКО