

ВІДГУК
ОФІЦІЙНОГО ОПОНЕНТА

доктора технічних наук, професора

Худова Геннадія Володимировича

на дисертаційну роботу Шупилюка Микити Володимировича

на тему: «Модель та методи інформаційної технології автоматизованого

графологічного аналізу рукописного тексту»,

подану до захисту на здобуття наукового ступеня доктора філософії

за спеціальністю 126 Інформаційні системи та технології

галузі знань 12 Інформаційні технології

1. Актуальність теми дисертаційної роботи.

Автоматизований графологічний аналіз рукописного тексту є важливим напрямом розвитку сучасних інформаційних систем і технологій, що активно застосовується в задачах криміналістичної експертизи, автентифікації особи, забезпечення безпеки та дослідження документів. У межах спеціальності 126 Інформаційні системи та технології актуальними є питання побудови моделей інформаційних технологій, розроблення методів інтелектуальної обробки візуальних даних та автоматизації аналітичних процесів опрацювання зображень.

Практична потреба сьогодні полягає не лише у порівнянні двох зразків письма, а у пошуку автора серед великих колекцій рукописних документів, коли кількість потенційних кандидатів заздалегідь невідома, а еталонні зразки представлені нерівномірно. Класичні дескриптори письма чутливі до умов отримання зображення, зокрема до роздільної здатності сканування, типу письмового інструменту, кольору та зашумленості основи, через що показники пошуку помітно погіршуються при переході від однієї колекції документів до іншої. У цьому контексті перспективним є формування компактних ознакових описів, інваріантних до масштабу та умов реєстрації зображення і придатних для індексації та швидкого пошуку за подібністю у великих сховищах даних. Водночас відомі рішення потребують трудомісткого ручного налаштування під

конкретну колекцію та розраховані на локальні обчислення, що ускладнює їх тиражування й масштабування в межах єдиної інформаційної системи.

Тому дослідження, присвячене розробленню моделі та методів інформаційної технології автоматизованого графологічного аналізу рукописного тексту, є своєчасним, науково обґрунтованим і відповідає актуальним завданням галузі знань 12 Інформаційні технології та спеціальності 126 Інформаційні системи та технології.

2. Методологія наукових досліджень.

У першому розділі дисертації сформовано методологічне підґрунтя дослідження. Проведено аналіз сучасного стану досліджень у сфері графологічного аналізу рукописного тексту, розглянуто основні процеси розпізнавання й аналізу почерку та особливості графологічного аналізу. Здійснено огляд і порівняльний аналіз методів ідентифікації автора та методів визначення особистості, проаналізовано психометричні моделі особистості, методи нанесення цифрового водяного знака та технології безпечного зберігання даних. За результатами аналізу сформульовано мету та задачі дослідження.

У другому розділі представлено метод визначення психологічних характеристик особистості за рукописним текстом, що набув подальшого розвитку. Обґрунтовано вибір психометричної моделі «Велика п'ятірка» та встановлено кореляцію її рис із відповідними графологічними особливостями почерку. Запропоновано функціональну модель визначення особистості, яка включає етапи попередньої обробки вхідних даних із застосуванням масштабно-інваріантного ознакового перетворення, а безпосереднє визначення особистості реалізовано на основі архітектури зорового трансформера. Обґрунтовано вибір програмних засобів і набору даних CVL для навчання; результати експериментального дослідження підтверджують ефективність методу та високий рівень точності класифікації психологічних характеристик.

У третьому розділі представлено удосконалений метод ідентифікації та пошуку автора за рукописним текстом. Обґрунтовано вибір масштабно-інваріантних центрально-окружних детекторів (CenSurE) як методу виявлення ознак зображень та запропоновано підхід, що інтегрує їх з архітектурою зорового

трансформера. Побудовано функціональну модель ідентифікації автора та здійснено її формалізацію шляхом побудови математичної моделі. Експериментально підтверджено суттєве збільшення загальної кількості вилучених ознак, скорочення середнього часу попередньої обробки та підвищення показників ідентифікації за метриками Hard Top K, Soft Top K і mAP.

У четвертому розділі представлено модель інформаційної технології автоматизованого графологічного аналізу рукописного тексту. Побудовано функціональну та структурну моделі інформаційної технології, здійснено математичну формалізацію задачі. Розроблено архітектуру інформаційної системи на базі технологій хмарних обчислень, яка охоплює повний цикл обробки даних, починаючи від завантаження зображення і завершуючи формуванням комплексного звіту. Інтегровано цифрові водяні знаки як метод захисту зображень та реалізовано механізми забезпечення цілісності даних на основі криптографічного хешування.

Достовірність та дієвість запропонованих моделей і методів підкріплено проведеними обчислювальними експериментами на еталонних наборах даних рукописного тексту, порівнянням отриманих показників із відомими рішеннями, а також актом впровадження результатів дослідження.

Отже, обрана методологія дослідження є цілісною, логічно завершеною та належно обґрунтованою. Одержані результати підтверджують достовірність сформульованих висновків, а також придатність розробленої інформаційної технології до практичного застосування.

3. Наукова новизна дисертаційної роботи.

Достовірність наукових положень підтверджено узгодженістю теоретичних побудов із результатами обчислювальних експериментів, а також апробацією основних результатів на міжнародних науково-технічних конференціях. До основних наукових результатів, що визначають новизну дисертаційної роботи, належать такі:

1. Отримала подальший розвиток модель інформаційної технології автоматизованого графологічного аналізу рукописного тексту, яка, на відміну від існуючих рішень, інтегрує в єдиному функціональному циклі метод визначення

особистості на основі зорового трансформера та метод ідентифікації автора на основі масштабно-інваріантних центрально-окружних детекторів і зорового трансформера, що забезпечує комплексне виконання задач ідентифікації автора та його психологічного профілювання в межах однієї інформаційної системи.

2. Набув подальшого розвитку метод визначення психологічних характеристик особистості за рукописним текстом, який, на відміну від існуючих підходів на основі згорткових нейронних мереж, використовує архітектуру зорового трансформера та психометричну модель «Велика п'ятірка», що забезпечує одночасне врахування локальних і глобальних графологічних ознак почерку та дозволяє підвищити точність класифікації психологічних характеристик.

3. Удосконалено метод ідентифікації та пошуку автора, в якому, за рахунок інтеграції масштабно-інваріантних центрально-окружних детекторів отримується більша кількість ознак з зображення рукописного тексту. Запропонований метод забезпечив збільшення значень точності ідентифікації, а також кількісне скорочення як середнього часу на елемент, так і загальної тривалості попередньої обробки, а також збільшення загальної кількості вилучених ознак.

Таким чином, отримані в дисертаційній роботі результати характеризуються науковою новизною, розширюють існуючі підходи до автоматизованого графологічного аналізу рукописного тексту та мають практичну спрямованість у задачах поведінкової біометрії, криміналістичної експертизи й психологічного профілювання.

4. Практична цінність дисертаційного дослідження.

Дисертаційна робота має істотну практичну і теоретичну цінність, оскільки запропоновані в ній модель та методи інформаційної технології автоматизованого графологічного аналізу рукописного тексту орієнтовані на застосування в реальних задачах криміналістичної експертизи, поведінкової біометрії, автентифікації особи та опрацювання архівних і історичних документів.

Розроблені підходи дозволяють підвищити точність визначення психологічних характеристик особистості та точність ідентифікації і пошуку автора за рукописним текстом, збільшити кількість виділених ознак зображення, а також зменшити середній час обробки одного елемента та загальну тривалість попередньої обробки зображень рукописного тексту.

Практична цінність результатів підтверджується розробленою програмною реалізацією інформаційної системи на базі хмарних технологій, яка забезпечує повний цикл обробки даних, починаючи від автентифікації користувача та завантаження зображення і завершуючи формуванням комплексного звіту, що містить результати ідентифікації автора та його психологічного профілю. Застосування цифрових водяних знаків і криптографічного хешування забезпечує захист та контроль цілісності оброблюваних зображень, що є критично важливим для експертної практики.

Результати роботи можуть бути застосовані в подальших наукових дослідженнях, а також впроваджені у прикладних системах, що потребують підвищення точності аналізу рукописного тексту та захисту цілісності зображень, зокрема в системах криміналістичної експертизи, поведінкової біометрії, автентифікації особи та опрацювання архівних і історичних документів.

5. Апробація результатів дослідження та наукові публікації.

Апробація результатів дисертаційного дослідження здійснювалася шляхом їх оприлюднення у наукових публікаціях та представлення на міжнародних науково-технічних конференціях.

Загалом за тематикою дисертації опубліковано 10 наукових праць, зокрема чотири статті, з них дві у наукових фахових виданнях України категорії Б, та шість тез доповідей у матеріалах міжнародних науково-технічних конференцій, серед яких дві публікації індексуються в наукометричній базі Scopus, одна з них у виданні третього квартиля.

Опубліковані наукові праці відображають основні етапи та результати проведеного дослідження, зокрема аналіз сучасного стану проблеми графологічного аналізу рукописного тексту, розроблення методу визначення

психологічних характеристик особистості, удосконалення методу ідентифікації та пошуку автора, побудову моделі інформаційної технології, а також експериментальну оцінку їх ефективності. У публікаціях висвітлено питання застосування архітектури зорового трансформера, методів виявлення ключових точок, попередньої обробки зображень рукописного тексту та забезпечення цілісності даних.

Основні положення дисертаційної роботи доповідалися та обговорювалися на міжнародних науково-технічних конференціях, що дозволило отримати зауваження й рекомендації фахівців і врахувати їх у подальших дослідженнях.

Сукупність опублікованих результатів підтверджує достатній рівень апробації основних положень дисертації, їх наукову новизну, обґрунтованість і достовірність, а також відповідність кількості та якості публікацій установленим вимогам.

6. Дотримання норм академічної доброчесності.

Ознайомлення з дисертаційною роботою та науковими публікаціями здобувача дає підстави стверджувати, що дослідження виконано з дотриманням принципів академічної доброчесності та встановлених вимог наукової етики.

Автором забезпечено коректне використання результатів попередніх досліджень, як власних, так і інших науковців із обов'язковим посиланням на відповідні джерела. Запозичення ідей, підходів і результатів здійснено належним чином зазначаючи джерела запозичень та з дотриманням правил цитування, що виключає можливість академічного плагіату.

Подані в дисертації наукові результати мають оригінальний характер, відображають особистий внесок здобувача та спрямовані на розв'язання актуальної науково-прикладної задачі у сфері інформаційних технологій автоматизованого графологічного аналізу рукописного тексту.

За результатами аналізу дисертаційної роботи порушень принципів академічної доброчесності не виявлено.

7. Зауваження.

Аналіз змісту та структури дисертаційної роботи, викладених у ній теоретичних положень, науково-методичних і прикладних розробок дає підстави

загалом оцінити роботу позитивно. Дисертація за своїм змістом відповідає сформульованій меті та поставленим завданням дослідження, матеріал викладено у логічній послідовності, науковим стилем українською мовою та оформлено відповідно до чинних вимог. Разом із тим, позитивно оцінюючи дисертаційну роботу, доцільно звернути увагу на окремі дискусійні положення, зокрема:

1. У третьому розділі при обґрунтуванні вибору масштабно-інваріантних центрально-окружних детекторів недостатньо детально розкрито вплив параметрів детектора (порогових значень відгуку, кількості масштабів, розміру фільтрів) на кількість вилучених ознак та час попередньої обробки. Доцільно було б навести аналіз чутливості запропонованого методу до цих параметрів.

2. У роботі ефективність ідентифікації оцінюється за метриками Hard Top K, Soft Top K та mAP, проте не наведено обґрунтування вибору значення K та пояснення, як ці показники співвідносяться з вимогами експертної практики, де важливою є ціна помилки першого та другого роду.

3. У роботі формується комплексний звіт із результатами ідентифікації та психологічного профілювання, проте недостатньо розкрито, як саме результат має інтерпретуватися кінцевим користувачем і яка міра впевненості супроводжує сформований висновок.

Проте наведені зауваження мають дискусійний характер, не є принциповими та не зменшують загальної позитивної оцінки дисертаційного дослідження, а окреслюють можливі напрями подальшого розвитку отриманих результатів.

8. Загальна оцінка роботи.

Дисертаційна робота Шупилюка Микити Володимировича «Модель та методи інформаційної технології автоматизованого графологічного аналізу рукописного тексту» є завершеною науковою працею, у якій розв'язано актуальне науково-прикладне завдання підвищення точності та автоматизація аналізу зображень рукописного тексту. Робота відзначається належним рівнем наукової зрілості, обґрунтованості та методологічної цілісності.

У дисертації здійснено ґрунтовний аналіз сучасного стану наукових підходів у сфері ідентифікації автора та визначення особистості за рукописним текстом, чітко визначено проблематику дослідження, а також окреслено напрями власних досліджень автора. У дисертаційній роботі простежується вдале поєднання теоретичних досліджень (розроблення моделей, математична формалізація методів) з експериментальною перевіркою отриманих результатів на еталонних наборах даних, що забезпечує її як наукову новизну, так і практичну спрямованість.

Структура роботи є логічно послідовною: кожен розділ містить постановку задач, обґрунтування методів дослідження, виклад отриманих результатів та їх аналіз. Наукові положення, висновки та рекомендації є достатньо обґрунтованими, підтвердженими результатами обчислювальних експериментів та актом впровадження.

З урахуванням наукової новизни отриманих результатів, рівня їх обґрунтованості та достовірності, а також практичної значущості сформульованих положень і висновків, можна зробити висновок, що дисертаційна робота відповідає вимогам пунктів 6, 7, 8, 9 Порядку присудження ступеня доктора філософії, затвердженого постановою Кабінету Міністрів України від 12 січня 2022 року № 44 (зі змінами).

Шупилюк Микита Володимирович заслуговує на присудження наукового ступеня доктора філософії у галузі знань 12 «Інформаційні технології» за спеціальністю 126 «Інформаційні системи та технології».

Офіційний опонент:

Начальник кафедри тактики радіотехнічних
військ факультету радіотехнічних військ
протиповітряної оборони Харківського
національного університету Повітряних Сил
імені Івана Кожедуба

доктор технічних наук, професор

Геннадій ХУДОВ