

УДК 004.8

Кок І.А., Поліщук Д.С., Кліменко В.І., Бармак О.В.

*Хмельницький національний університет*

## **НЕЙРОМЕРЕЖЕВА СИСТЕМА АВТОМАТИЗОВАНОГО КОНТРОЛЮ ЗА ТЮТЮНОПАЛІННЯМ У ПУБЛІЧНИХ МІСЦЯХ**

*У роботі розглянуто нейромережеву систему автоматизованого контролю за тютюнопалінням у публічних місцях, орієнтовану на підтримку реалізації антитютюнової політики в умовах сучасної урбанізованої інфраструктури. Ядром рішення є модель на основі Vision Transformer, попередньо навчена на великомасштабному наборі зображень і донавчена методом трансферного навчання на спеціалізованому корпусі кадрів із курцями та некурцями. Для підвищення прозорості роботи системи та придатності результатів до використання в адміністративних процедурах інтегровано модифікований Grad-CAM, що забезпечує візуалізацію найбільш інформативних ділянок кадру. Показано, що запропонований підхід створює підґрунтя для інтеграції контролю за тютюнопалінням у платформи «розумного міста» та подальшого розвитку мультимодальних систем моніторингу поведінкових порушень.*

*The paper presents the neural network-based system for automated control of smoking in public spaces, designed to support tobacco control policy within modern urban environments. Its core is a Vision Transformer model pre-trained on a large-scale image dataset and fine-tuned via transfer learning on a domain-specific corpus of images containing smokers and non-smokers. To ensure decision transparency and suitability of the outputs for administrative procedures, a modified Grad-CAM method is integrated, providing visual explanations that highlight the most informative regions of each frame. The proposed approach forms a basis for integrating smoking control into smart city platforms and for further development of multimodal AI systems for monitoring behaviour-related violations in public spaces.*

Актуальність застосування методів комп'ютерного зору та глибокого навчання для автоматизованого контролю за тютюнопалінням у публічних місцях зумовлена зростаючими вимогами до дотримання антитютюнового законодавства та необхідністю оперативного моніторингу поведінкових порушень у багатолюдних урбанізованих просторах [1]. Поширення мереж відеоспостереження в містах створює передумови для впровадження інтелектуальних систем відеоаналітики [2], які здатні автономно виявляти ризиковані чи заборонені дії, однак традиційні алгоритми комп'ютерного зору, засновані на ручних ознаках, виявляються недостатньо стійкими до варіацій освітлення, ракурсів, динамічних сцен та поведінкової різноманітності [3]. Тютюнопаління характеризується поєднанням дрібних локальних об'єктів (сигарета, дим) і складних позно-жестових патернів, що

робить задачу виявлення нетривіальною та потребує використання високорівневих моделей, здатних до глибокої інтерпретації просторово-часових даних.

Сучасні конволюційні нейронні мережі [4] та трансформерні архітектури [5] відкривають можливість моделювання складних візуальних залежностей, інтегруючи як локальні текстури, так і глобальні контексти сцени [6]. CNN, зокрема ResNet, EfficientNet або MobileNet, демонструють високу продуктивність у завданнях детекції дрібних об'єктів та класифікації поведінкових ознак [7], тоді як Vision Transformer забезпечує ширшу контекстуальну узагальненість і стійкість до шумів відеопотоку [8]. Комбінація цих підходів дозволяє створювати моделі, здатні виявляти сигарету навіть за умов часткового перекриття об'єкта, низької якості кадру або нерегулярної появи диму, що суттєво підвищує точність системи порівняно з класичними методами [9].

У контексті цієї роботи ключовою є можливість масштабування архітектур глибокого навчання до режиму безперервного аналізу відеопотоків у реальному часі [10, 11]. Оптимізовані CNN дозволяють виконувати інференс з високою пропускну здатністю, що є критичним для використання в міських системах відеоспостереження, де кількість камер може сягати тисяч. Додаткові переваги створює застосування трансферного навчання, яке дає змогу швидко адаптувати попередньо навчені моделі до спеціалізованих корпусів відеоданих, що містять унікальні патерни тютюнопаління, притаманні конкретним середовищам спостереження. Такі моделі здатні урахувати контекст локації, характерні жести користувачів, типи сигаретних виробів, що підсилює їхню робастність та зменшує кількість хибних спрацьовувань [12].

Не менш важливим аспектом є питання інтерпретованості, яке вирішується завдяки використанню теплових карт уваги, зокрема модифікованих варіантів Grad-CAM. Це забезпечує прозорість прийняття рішень системою та робить можливим використання результатів у юридично значущих процедурах. Візуалізація інформативних ділянок кадру дає змогу операторам підтвердити або спростувати факт порушення, а також коригувати роботу моделі за рахунок аналізу помилкових класифікацій.

Таким чином, застосування комп'ютерного зору та CNN у задачі виявлення тютюнопаління формує технологічну основу для створення масштабованих, стійких і пояснюваних систем контролю поведінки в публічних місцях. У поєднанні з інфраструктурою «розумного міста» такі системи можуть стати ключовим інструментом у реалізації політик громадського здоров'я, а також у підвищенні ефективності міських сервісів безпеки та превентивного моніторингу.

Нейромережева система автоматизованого контролю за тютюнопалінням у публічних місцях розглядається як складний програмно-апаратний комплекс, що поєднує технології комп'ютерного зору, глибокого навчання та інструменти пояснюваного штучного інтелекту для підтримки реалізації антитютюнової політики в умовах сучасного міського середовища. Її призначення полягає у безперервному аналізі відеопотоків із камер спостереження, виявленні ознак

тютюнопаління в реальному часі, формуванні машинних рішень щодо факту порушення та наданні оператору інтерпретованих візуальних доказів, придатних для подальших адміністративних дій. Концепція такої системи спирається на попередньо розроблений метод автоматизованого контролю за тютюнопалінням з використанням архітектури Vision Transformer та Grad-CAM-візуалізацій, удосконалюючи його до рівня повноцінної прикладної інфраструктури відеоаналітики.

Проблематика тютюнопаління в публічних місцях має комплексний характер і виходить за межі суто медичного виміру. Вона охоплює аспекти громадського здоров'я, правозастосування, міського планування та цифрової трансформації сервісів безпеки. Попри багаторічну практику нормативного обмеження куріння, більшість держав, у тому числі Україна, стикається із систематичними порушеннями заборони на використання тютюнових виробів, електронних сигарет та пристроїв нагрівання тютюну у місцях масового скупчення людей. Пасивне куріння, яке неминуче супроводжує такі порушення, асоціюється з підвищеним ризиком серцево-судинних, онкологічних та респіраторних захворювань, причому особливо вразливими є діти, літні люди та особи з хронічними хворобами дихальної системи. За умов інтенсивної урбанізації та збільшення пасажиропотоків у громадському транспорті, торговельно-розважальних центрах, на транспортних вузлах та територіях освітніх закладів навіть епізодичні випадки куріння створюють значний сукупний ризик для населення.

Традиційна модель контролю за дотриманням антитютюнового законодавства ґрунтується на чергуванні працівників поліції, інспекторів муніципальних служб чи локального персоналу безпеки. Такий підхід демонструє обмежену ефективність у масштабі сучасного міста, оскільки фізична присутність контролюючих осіб у кожній потенційно проблемній точці є економічно недоцільною, а людський фактор неминуче призводить до пропуску значної частки порушень. Розгортання мереж відеоспостереження формує технічну основу для переходу до автоматизованих стратегій моніторингу поведінки людей, проте без інтелектуальної обробки відеоданих такі системи залишаються пасивним засобом фіксації подій, а не інструментом активного запобігання.

Нейромережева система автоматизованого контролю за тютюнопалінням пропонує якісно інший рівень функціонування інфраструктури відеоспостереження. Вона перетворює камери з традиційних сенсорів зображення на активні вузли мережі штучного інтелекту, що здатні самостійно розпізнавати ситуації куріння на основі аналізу просторово-візуальних патернів. Відповідно, навантаження на персонал зміщується від ручного перегляду багатогодинних відеозаписів до вибіркової роботи з уже попередньо відфільтрованими та інтерпретованими сповіщеннями, що містять кадри з високою ймовірністю виявленого порушення. Це не лише підвищує оперативність реагування, але й створює прозорий, задокументований механізм фіксації інцидентів, який знижує суб'єктивність оцінок.

З науково-технічної точки зору, завдання розпізнавання фактів тютюнопаління має низку особливостей, які відрізняють його від класичних задач комп'ютерного зору. По-перше, об'єкти інтересу є дрібними, часто частково перекритими і візуально подібними до фонового шуму: сигарета має невеликі розміри, а дим може маскуватися під атмосферні ефекти, відблиски чи шуми сенсора. По-друге, ключовими індикаторами є не лише самі об'єкти, а й поведінкові жести – характерне положення руки, піднесення її до рота, певна конфігурація обличчя. По-третє, середовище спостереження характеризується значною варіативністю: різкі зміни освітлення, контрасту, ракурсу, типу камери, щільності навігпу. Ці фактори обумовлюють необхідність застосування архітектур глибокого навчання, здатних моделювати глобальний контекст сцени та вловлювати високорівневі залежності між просторово рознесеними ознаками.

Класичні згорткові нейронні мережі, які домінували в комп'ютерному зорі протягом останнього десятиліття, демонструють високу ефективність у задачах класифікації та детекції об'єктів, однак їхня локальна природа обробки поступово виявляє обмеження в подібних задачах. Згортки будують подання на основі локальних рецептивних полів, і хоча використання глибоких архітектур, залишкових зв'язків та багатомасштабних блоків частково компенсує цей недолік, глобальний контекст сцени все ще відновлюється опосередковано. На противагу цьому трансформерні підходи, такі як Vision Transformer, безпосередньо моделюють взаємодію між усіма фрагментами зображення через механізм самоуваги, що особливо важливо для ситуацій, коли вирішальними є відносні взаєморозташування об'єктів та поведінкові патерни.

У запропонованій нейромережевій системі центральне місце посідає модель на основі Vision Transformer, попередньо навчена на великомасштабному наборі ImageNet та донавчена за допомогою трансферного навчання на спеціалізованому корпусі зображень із курцями та некурцями. Вхідний кадр з відеопотоку перетворюється у нормалізоване зображення фіксованого розміру, після чого розбивається на рівновеликі патчі, які проєктуються у векторний простір ознак. Послідовність цих векторів доповнюється позиційними кодуваннями, що зберігають інформацію про розташування фрагментів у вихідному кадрі, а також спеціальним класифікаційним токеном, який акумулює глобальну інформацію. Далі послідовність проходить через багатопаровий енкодер із багатоголовою самоувагою та блоками багатопарового перцептрона, де формується високорівневе подання сцени.

Стратегія трансферного навчання, покладена в основу системи, орієнтована на ефективну роботу з відносно невеликими доменно-специфічними датасетами. Основні параметри трансформерного енкодера залишаються замороженими, що дозволяє зберегти узагальнюючу здатність моделі, накопичену на базовому наборі даних, тоді як донавчання зосереджується на фінальному класифікаційному шарі, адаптуючи його до дихотомії «куріння / відсутність куріння». Такий підхід мінімізує ризик перенавчання, зменшує обчислювальні витрати та дає змогу

впроваджувати систему на обладнанні з обмеженими ресурсами, зокрема на вбудованих рішеннях у мережесх відеореєстраторах або edge-пристроях.

Важливим концептуальним компонентом нейромережевої системи є пояснюваність прийнятих рішень. Оскільки результати автоматизованого контролю можуть бути підставою для адміністративних стягнень, оператори та уповноважені органи мають отримувати не лише бінарний висновок моделі, а й наочне обґрунтування. Для цього в систему інтегровано модифікований підхід Grad-CAM, адаптований до трансформерної архітектури. Після отримання класифікаційного результату алгоритм обчислює градієнти за відношенням до проміжних представлень патчів, агрегує їх та формує теплову карту важливості, яка накладається на оригінальний кадр. Таким чином, оператор бачить, які області зображення стали вирішальними: це може бути зона в руках людини, обличчя, згусток диму або комбінація кількох ділянок. Така візуалізація підвищує довіру до системи, спрощує виявлення помилкових спрацювань та сприяє аналізу граничних випадків. Діаграма класів розробленої інтелектуальної системи наведено на рисунку 1.

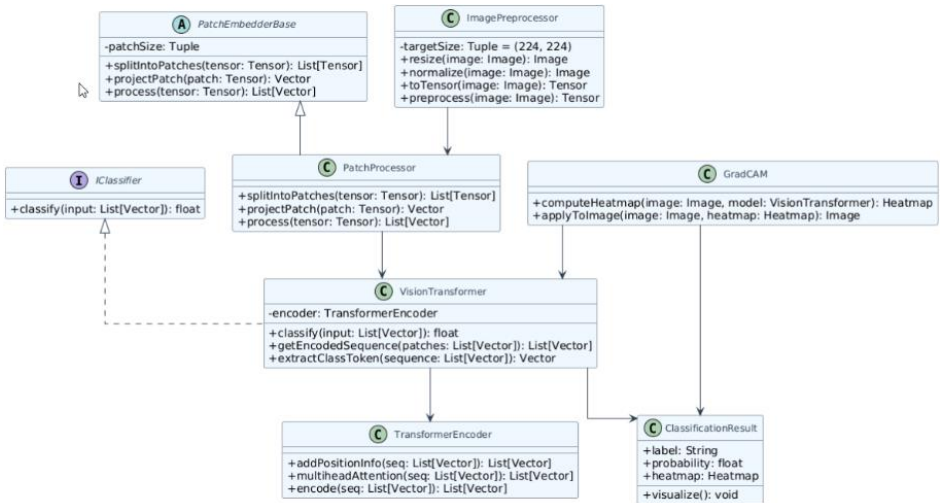


Рисунок 1 – Діаграми класів

Система функціонує в кількох логічно пов'язаних режимах. У режимі безперервного моніторингу відеопотік із камер розщеплюється на окремі кадри із заданою періодичністю або подією (наприклад, за зміни сцени чи виявлення людей зовнішніми детекторами). Кожен кадр або серія ключових кадрів передається до модуля попередньої обробки, де виконується масштабування, колірна нормалізація, при потребі – стабілізація та усунення артефактів стиснення. Після цього

зображення потрапляє до нейромережевого ядра, яке повертає ймовірність наявності тютюнопаління. Якщо ця ймовірність перевищує наперед визначений поріг, модуль прийняття рішень створює об'єкт події, до якого включаються кадр-джерело, теплові карти, часові мітки та ідентифікатор камери. Подальші дії залежать від політики експлуатації: система може автоматично формувати сповіщення оператору, зберігати фрагмент відео у спеціальному сховищі, інтегруватися з міськими центрами обробки даних або інформаційними системами правоохоронних органів.

Другий важливий режим – аналітико-дослідницький. Він призначений для розробників, адміністраторів та експертів у сфері громадського здоров'я і дозволяє працювати з архівними даними, оцінювати динаміку порушень у певних локаціях, час доби чи днях тижня, а також аналізувати помилки нейромережі. У рамках цього режиму забезпечується розширена візуалізація результатів, можливість переглядати серії кадрів до і після події, а також побудова статистичних звітів. Така функціональність відкриває перспективи інтеграції нейромережевої системи контролю за тютюнопалінням у ширші платформи «розумного міста», де дані про куріння розглядаються поряд з іншими індикаторами безпеки та поведінки населення.

Формування навчального і тестового корпусу є критичним етапом, що визначає якість роботи системи. Для початкового навчання використовується відкритий збалансований датасет зображень, у якому категорії курців та некурців представлені приблизно в рівних пропорціях. У класі «курців» переважають сцени з чітко видимими сигаретами або пристроями для вейпінгу, у той час як клас «некурців» включає як нейтральні зображення, так і кадри з подібними до куріння жестами, але без фактичної наявності тютюнових виробів. Цей прийом ускладнює задачу й змушує модель шукати тонкі відмінності, не спираючись лише на грубі контури руки біля обличчя. Після базового навчання система може донавчатися на локальних даних з конкретних камер або інфраструктурних об'єктів, що дає змогу враховувати особливості реального середовища – типові ракурси, освітлення, локальні культурні патерни поведінки.

Експериментальне дослідження нейромережевої системи проводиться на основі розділення корпусу зображень на навчальну, валідаційну та тестову підмножини. У процесі тренування відстежуються такі метрики, як точність класифікації, повнота, точність у вузькому розумінні, F1-міря та площа під ROC-кривою. Аналіз кривих навчання й валідації дозволяє своєчасно зупиняти навчання при досягненні плато, запобігаючи перенавчанню. За результатами тестування на відкладеній вибірці можна зробити висновок про здатність моделі узагальнювати знання на нових сценах. Особливу увагу приділяють аналізу помилкових позитивних та помилкових негативних спрацювань, оскільки перші створюють надлишкове навантаження на операторів, а другі безпосередньо знижують ефективність контролю.

З погляду системної інтеграції важливою характеристикою є масштабованість нейромережевої системи. Вона повинна коректно працювати як у форматі пілотного проекту на декількох камерах, так і в умовах міської інфраструктури з сотнями чи тисячами джерел відео. Для цього передбачається модульна мікросервісна архітектура, у якій окремі вузли відповідають за отримання відеоданих, буферизацію, розподіл кадрів між інференс-модулями, зберігання результатів та візуалізацію. За наявності графічних прискорювачів або спеціалізованих нейропроцесорів система може виконувати обробку в режимі, близькому до реального часу, що дозволяє оперативно реагувати на порушення в публічних місцях. Рисунок 2 наводить приклад роботи розробленої системи.

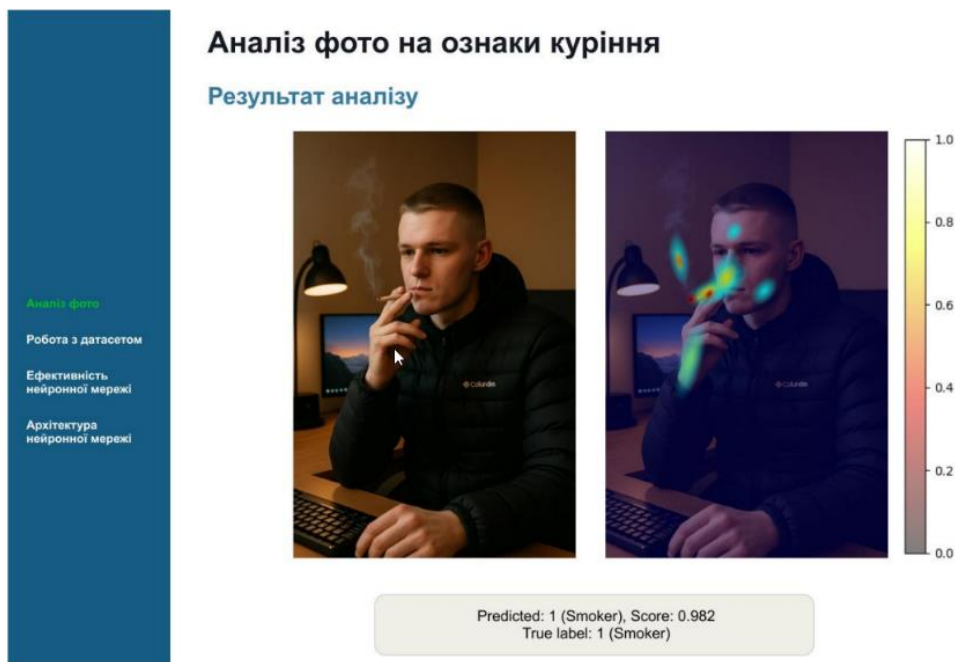


Рисунок 2 – Результат виявлення тютюнопаління

Разом з технічними аспектами не можна ігнорувати питання етики та захисту персональних даних. Нейромережева система автоматизованого контролю за тютюнопалінням працює з відеозображеннями людей, а отже повинна відповідати вимогам законодавства щодо обробки персональних даних і принципам мінімізації втручання в приватність. Це передбачає обмеження доступу до повних відеозаписів, використання ролей та прав доступу, протоколювання дій операторів, а також, за можливості, застосування технологій деперсоналізації або розмиття

облич у тих сценаріях, де ідентифікація конкретної особи не є необхідною. Технічні рішення мають бути доповнені організаційними регламентами, що визначають, хто, скільки часу та з якою метою може працювати з даними системи.

Незважаючи на перспективність запропонованого підходу, нейромережева система має низку обмежень, які визначають напрями подальших досліджень. Одним із ключових є залежність від якості та репрезентативності даних: якщо у навчальному наборі недостатньо прикладів певних сценаріїв, система може демонструвати нестабільну поведінку в реальних умовах. Ще одним викликом є адаптація до нових форм курильних пристроїв, дизайни яких швидко змінюються під впливом ринку. Це вимагає регулярного оновлення корпусу зображень та періодичного донавчання моделі. Окремою лінією розвитку є перехід від класифікації на рівні кадру до повноцінної просторово-часової детекції, коли система здатна враховувати динаміку жестів у коротких відеофрагментах, а не лише статичні зображення.

Перспективним напрямом є також використання мультимодальних моделей, які поєднують аналіз відео з іншими каналами даних. Наприклад, у закритих приміщеннях система може доповнюватися сенсорами якості повітря, що фіксують концентрації продуктів згоряння, а на відкритих майданчиках – даними з соціальних сервісів або платформи «розумне місто», які вказують на масові заходи чи аномальні скупчення людей. Інтеграція таких сигналів із нейромережею дозволить формувати більш надійні композитні рішення, зменшуючи кількість хибних спрацювань.

Підсумовуючи, нейромережева система автоматизованого контролю за тютюнопалінням у публічних місцях поєднає низку сучасних технологічних складових: архітектуру Vision Transformer для високоточного розпізнавання візуальних ознак куріння, методи трансферного навчання для роботи з обмеженими датасетами, Grad-CAM-інструменти пояснюваності, модульну інфраструктуру для обробки відеопотоків та аналітичні засоби для підтримки прийняття управлінських рішень. Її впровадження здатне підвищити ефективність виконання антитютюнового законодавства, зменшити ризики, пов'язані з пасивним курінням, та сприяти побудові більш безпечного й здорового міського середовища. Одночасно система створює підґрунтя для подальших досліджень на стику глибокого навчання, комп'ютерного зору, цифрової безпеки та політики громадського здоров'я, де кожне нове покоління моделей може забезпечувати ще більш точний, стійкий і прозорий контроль за поведінковими порушеннями у публічному просторі.

### Перелік посилань

1. Куріння, що повільно вбиває. До Всесвітнього дня без тютюну. НАМН України. НАМН України. URL: <https://amnu.gov.ua/kurinnya-shho-povilno-vbyvaye/>
2. Аналіз ринку цигарок в Україні. Інвестиції в Україні. InVenture. URL: <https://inventure.com.ua/uk/analytics/investments/analiz-rinku-cigarok-v-ukrayini>

3. Vezha Smoking Detection module allows to identify a smoker with a cigarette. Milestone Systems. URL: <http://www.milestonesys.com/technology-partner-finder/incoresoft-llc/vezha-smoking-detection/>
4. Didur V., Molchanova M., Mazurets O. Research on the effectiveness of neural network detection of plots with the destroyed buildings remains. Modern technologies and science: problems, new and relevant developments. Proceedings XXI International Scientific and Practical Conference. May 26, 2025. Zaragoza, Spain. Pp. 245-251.
5. Bas I.S., Kadynska V.D., Klimenko V.I., Mazurets O.V. Convolutional Neural Network Transfer Learning Method for Aircraft Image Classification. Scientific method: reality and future trends of researching. Proceedings of VI International Scientific and Theoretical Conference. June 6, 2025. Montreal, Canada. Pp. 147-155.
6. Mushtyn O., Sobko O., Molchanova M., Mazurets O. Convolutional Neural Network Architecture for Image-Based Architectural Style Recognition. Evolving Science: Theories, Discoveries and Practical Outcomes. Proceedings of 4th International Scientific and Practical Conference. June 9-11, 2025. Zurich, Switzerland. Pp. 130-143.
7. Malaydakh V., Molchanova M., Shevchuk P., Mazurets O. Deep learning neural network architecture for determining sunflower growth stage from visual data. Modern Scientific Research: Theoretical and Practical Aspects. Proceedings II International Scientific and Practical Conference. May 26-28, 2025. Riga, Latvia. Pp. 143-148.
8. Hladun O.V., Molchanova M.O., Zalutska O.O., Mazurets O.V. Effectiveness research of using ViT neural network architecture for classifying the destroyed buildings remains. Achievements of Science and Applied Research. Proceedings of 2nd International Scientific and Theoretical Conference. May 19-21, 2025. Dublin, Ireland. Pp. 96-100.
9. Krak I., Mazurets O., Zalutska O., Molchanova M., Barmak O., Approach for Postural Disorders Recognizing from Visual Data Using Deep Neural Network, CEUR Workshop Proceedings. 4049. 2025. Pp. 244-255.
10. Кок І.А., Мазурець О.В., Кліменко В.І., Петровський С.С. Метод автоматизованого визначення оцінки ступеня співвіднесення графічних зображень до актуальних категорій із застосуванням згорткової нейронної мережі. Збірник наукових праць за матеріалами XVI Всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2024». 15-16 листопада 2024. Хмельницький, 2024. с. 277-281.
11. Kok I.A., Kadynska V.D., Zalutska O.O., Mazurets O.V. Object-Oriented Intelligent System for Automated Control of Smoking by Video Data. Current scientific goals, approaches and challenges. Proceedings of IV International Scientific and Theoretical Conference. June 13, 2025. Dresden, Federal Republic of Germany. Pp. 156-164.
12. Кок І.А., Мазурець О.В., Молчанова М.О. Пояснений підхід на основі трансферного навчання Vision Transformers до виявлення куріння у публічних просторах. Матеріали XIII Міжнародної науково-практичної конференції «Інформаційні управляючі системи та технології ІУСТ-ОДЕСА-2025». 24-26.09.2025. Одеса. 2025. С.147-149.