

ЯВОРСЬКИЙ КОСТЯНТИН

Хмельницький національний університет

<https://orcid.org/0009-0000-7139-9723>e-mail: yavakostya@gmail.com**МАНЗЮК ЕДУАРД**

Хмельницький національний університет

<https://orcid.org/0000-0002-7310-2126>e-mail: eduard.em.km@gmail.com**СКРИПНИК ТЕТЯНА**

Хмельницький національний університет

<https://orcid.org/0000-0002-8531-5348>e-mail: tkskripnik1970@gmail.com**ПАСІЧНИК ОЛЕКСАНД**

Хмельницький національний університет

<https://orcid.org/0000-0002-8760-4688>e-mail: o.a.pasichnyk@gmail.com

ВИЗНАЧЕННЯ ОБСЯГУ ДАНИХ ДЛЯ ЕФЕКТИВНОЇ КЛАСИФІКАЦІЇ НОМЕРІВ АВТОМОБІЛІВ

В роботі наведено результати дослідження впливу обсягу даних, зокрема за розміром датасету та за кількістю класів, на ефективність класифікації номерів автомобілів за допомогою нейронної мережі. Тестування проводиться на нейронній мережі YOLOv11.

Ключові слова: розпізнавання номерних знаків, нейронні мережі, комп'ютерний зір.

YAVORSKYI KOSTIANTYN, MANZIUK EDUARD, SKRYPNYK TETIANA, PASICHNYK OLEKSANDR
Khmelnitskyi National University

DETERMINATION OF THE DATA VOLUME FOR EFFECTIVE CLASSIFICATION OF VEHICLE LICENSE PLATES

This paper presents the results of a study investigating the impact of dataset size and class count on vehicle license plate classification using neural networks. The research was conducted using the YOLOv11 model, a state-of-the-art architecture for object detection tasks. The study comprises four main steps, each designed to assess how different configurations of training data affect the model's performance.

The first step involved training the model on a medium-sized dataset consisting of four classes over 15, 30, and 50 epochs. The second step focused on training the model using the same medium-sized dataset but limited to a single class for 30 epochs. The third step evaluated the model's performance on a smaller dataset over the same set of epochs (15, 30, 50). After each step, detailed analysis of training graphs and performance metrics was conducted to assess how the model's accuracy and loss evolved with different data sizes and epoch counts.

In the final step, a comprehensive comparison of all models was performed, taking into account the overall performance across the various dataset sizes and class configurations. This comparison aimed to identify the optimal training configuration, with a focus on maximizing accuracy and minimizing overfitting, to determine the best approach for vehicle license plate classification.

Through this study, the research highlights the influence of dataset volume and class distribution on model performance, providing insights into how to effectively train neural networks for classification tasks with varying data complexities. The findings underscore the importance of selecting an appropriate dataset size and class count to balance computational efficiency and classification accuracy in real-world applications.

Keywords: license plate recognition, neural networks, computer vision.

Вступ та постановка проблеми

Кількість автомобілів на дорогах України, Європи, США збільшується, що створює потребу в вдосконаленні методів детекції номерів. Ефективне розпізнавання об'єктів, таких як автомобільні номери, значною мірою залежить від якості та структури навчального набору даних. Грамотно сформований датасет може значно підвищити точність та продуктивність моделі. Автомобільні номери мають різноманітні характеристики, що їх відрізняють. Зокрема, існують унікальні, індивідуалізовані номери, які створені за спеціальними замовленнями та мають нестандартний дизайн. Водночас більшість номерів є стандартизованими, що дозволяє їх класифікувати за країною походження [1, 2].

Поділ датасету на класи може відбуватись на основі цих особливостей. Один клас може включати номери з унікальним дизайном, а інші – стандартизовані номери, розподілені за географічною ознакою. Такий підхід до структуризації даних сприяє покращенню адаптації моделі до різних типів номерів та підвищує її здатність до розпізнавання у реальних умовах.

З наукової точки зору, дослідження ефективності класифікації номерів допомагає покращити розуміння впливу різних параметрів навчальних даних на продуктивність моделей. Аналіз структурованих датасетів із різними класами (наприклад, унікальні чи стандартизовані номери) сприяє розробці більш універсальних алгоритмів глибокого навчання, здатних адаптуватися до складних і неоднорідних умов. Це також відкриває нові перспективи для вивчення адаптивності моделей до багатокласових задач, що є актуальним напрямом у машинному навчанні.

З практичної точки зору, розв'язання цієї проблеми має безпосередній вплив на реальні сценарії застосування, такі як автоматичне розпізнавання номерних знаків у транспортних системах, безпечних

рішеннях, автоматизації процесів паркування та митного контролю. Використання грамотно підбраного датасету дозволяє підвищити точність роботи систем розпізнавання, знизити кількість помилкових визначень та покращити продуктивність навіть у складних умовах, таких як різні погодні умови чи змішання стилів номерних знаків.

Метою роботи є: дослідження впливу обсягу даних на ефективність класифікації за допомогою нейронних мереж, зокрема вивчення взаємозв'язку між розміром датасету та кількістю класів. Це дослідження включає аналіз того, як збільшення обсягу даних та варіації в класах можуть вплинути на точність класифікації в задачах детекції, зокрема для класифікації номерів автомобілів. Одним з ключових аспектів є визначення оптимального розміру датасету та кількості класів, що дозволяє досягти найкращих результатів при тренуванні моделей нейронних мереж, таких як YOLO, для завдань, що вимагають високої точності та ефективності. У рамках роботи також вивчаються питання адаптації моделей до змін в обсязі даних та оптимізація навчальних процесів для забезпечення кращої узагальнюючої здатності при роботі з реальними умовами.

Дослідження

Для проведення дослідження була використана нейромережа YOLOv11, яка є однією з найбільш ефективних архітектур для задач детекції об'єктів. Процес навчання моделі за допомогою YOLOv11, зображений на рисунку 1. Зображення датасету потрапляють до основних компонентів архітектури YOLO, які можна поділити на три головні частини: каркас, шия і голова.

Каркас відповідає за вилучення ознак із зображення – це глибока згортокова нейронна мережа, яка аналізує зображення і формує набір характеристик. Далі ці ознаки передаються до "ший", яка об'єднує інформацію з різних рівнів ієрархії ознак і готує їх до остаточного передбачення. "Голова" моделі виконує фінальні операції детекції та ідентифікації об'єктів, формуючи відповідні бокси навколо об'єктів та визначаючи їх категорії.

На останніх етапах передбачення, отримані від "головы", проходять процеси загального об'єднання та вирівнювання, що забезпечує точність і стабільність виявлених об'єктів. Після цього модель зберігається у вигляді ваг і архітектури для подальшого використання, тестування чи розгортання [3].

YOLOv11 включає нові архітектурні інновації та оптимізації в порівнянні з минулими версіями, зокрема заміну блоку C2f на C3k2 для покращення продуктивності детекції. C3k2 оптимізує обчислення за рахунок використання двох малих згорток замість однієї великої, як у YOLOv8, що підвищує швидкість обробки без втрати якості. Рівняння що описує цей блок:

$$C3k2(X) = Conv(Split(X)) + Conv(Merge(Split(X))), \quad (1)$$

де Conv – це операція згортки, Split – операція розділення карти ознак на 2 частини, Merge – операція об'єднання вихідних результатів.

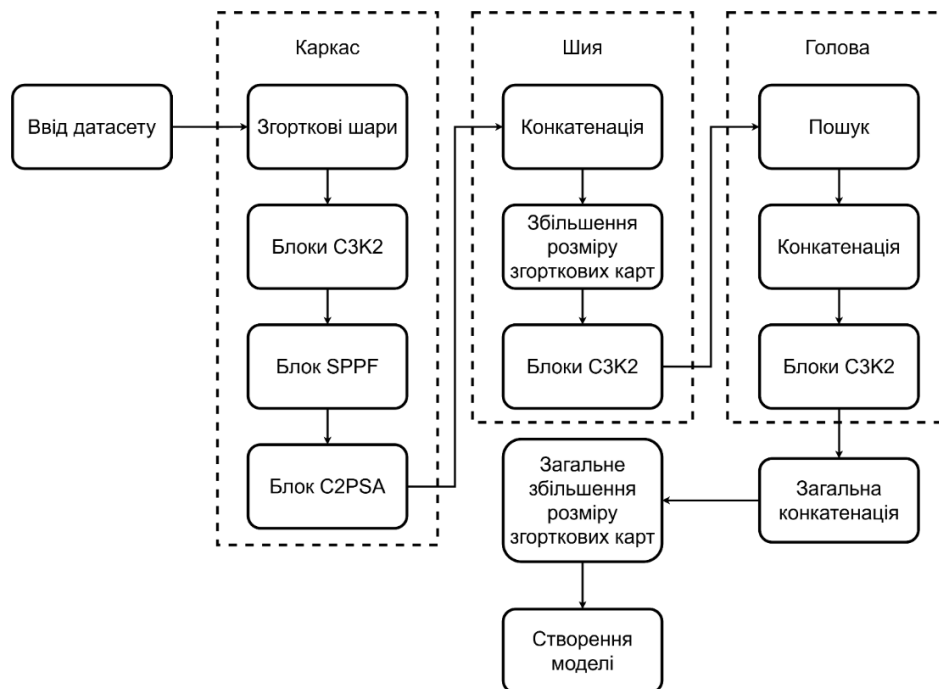


Рис. 1. Схема процесу навчання нейромережі YOLOv11

У YOLOv11 також зберігається використання блоку просторової піраміди (SPPF), але додано новий блок C2PSA. Цей блок дозволяє моделі краще фокусуватися на важливих ділянках зображення, покращуючи точність виявлення об'єктів різного розміру та положення завдяки механізмам просторової уваги. Рівняння що описує цей блок:

$$C2PSA(X) = Attention(Concat(X_{path1}, X_{path2})), \quad (2)$$

де Attention – це механізм, що дозволяє фокусуватися на важливих частинах, Concat – операція, яка з'єднує карти ознак.

Рівняння, що описує блок SPPF:

$$SPPF(X) = Concat(MaxPool(X, 5), MaxPool(X, 3), MaxPool(X, 1)), \quad (3)$$

де MaxPool(X) – це операція “максимального пулінгу” [4].

Для проведення дослідження були підготовлені датасети з малою та середньою кількістю зображень для навчання. Утворено дві версії середнього, за кількістю зображень, датасету. В першій, зображення розділились за чотирма наступними класами

- номери машин (Car plates);
- номери машин України (Car plates Ukraine);
- номери машин ЄС (Car plates EU);
- номери машин США (Car plates USA).

В другій версії, зображення належали одному об'єднаному класу. Розділення датасета на класи зображено на рисунку 2.

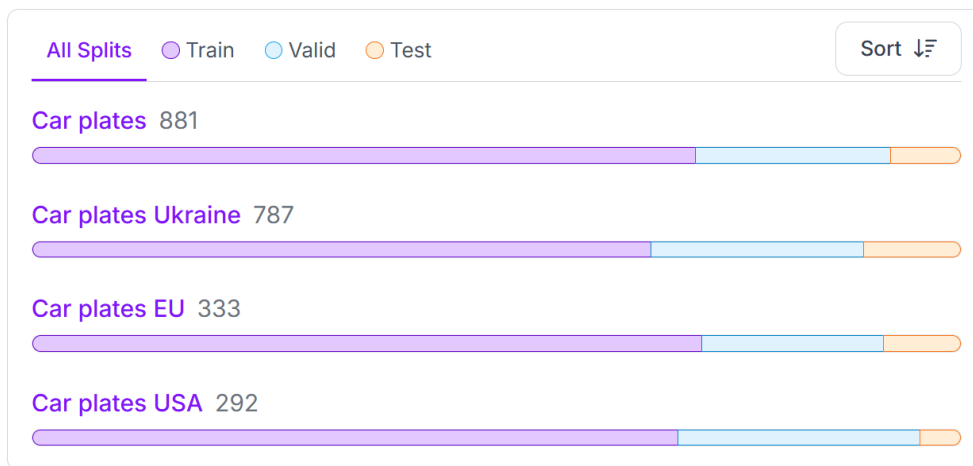


Рис. 2. Кількість анотацій кожного класу

Малий датасет складається з 330 зображень, тоді як середній датасет, після попередньої обробки, що включала дублювання зображень із шумом та їх поворот за осями, налічує 3829 зображень.

Оптимізатор, для створення моделі, налаштований за допомогою AdamW – вдосконаленої версії Adam, який використовує адаптивні моменти для кожного параметра, що дозволяє ефективніше коригувати ваги в процесі навчання. Вибраний оптимізатор розв’язує одну з проблем класичного рішення – не зовсім коректне застосування вагового спаду (weight decay). У AdamW процес регуляризації та корекція параметрів за допомогою адаптивних моментів (перших і других моментів) відбуваються незалежно один від одного.

Для навчання використовується початкова швидкість навчання (learning rate) 0.0001 – досить низьке значення, що дозволяє моделі поступово коригувати свої параметри, запобігаючи надмірним коливанням під час навчання. Параметр weight_decay встановлений на значення 0.0005, що додає до функції втрат штраф за великі значення ваг, сприяючи покращенню узагальнення моделі і зменшенню ймовірності перенавчання.

Використання цих налаштувань дозволяє збалансувати швидкість навчання і ефективність оптимізації, забезпечуючи стабільний процес навчання та мінімізацію втрат.

Дослідження складається з чотирьох кроків: першим кроком відбулось навчання моделі на середньому датасеті з чотирма класами протягом 15, 30, 50 епох, другим кроком відбулось навчання, на середньому датасеті з одним класом на 30 епохах, третім кроком відбулось навчання моделі на малому датасеті протягом 15, 30, 50 епох. В кінці кожного кроку був проведений аналіз графіків та метрик. Четвертим кроком відбулось загальне порівняння моделей, та вибір оптимальної.

Під час навчання на першому кроці спостерігається швидке зростання mAP на початкових етапах (до 2000 кроків), що свідчить про ефективне тренування на початку. Після досягнення 4000-5000 кроків модель виходить на плато – приріст точності незначний, але стабільний. Після 6000 кроків mAP залишається майже незмінним, що може свідчити про досягнення межі продуктивності даної моделі на конкретному датасеті. Оптимальною моделлю на цьому кроці можна вважати ту, яка навчена на тридцяти епохах. Ефективність навчання моделі на середньому датасеті протягом 15, 30, 50 епох зображена на рисунку 3.

Другим кроком було проведено навчання на середньому датасеті протягом оптимальних 30 епох, без розділу на різні класи. На рисунку 4 зображено порівняння “mAP50-95” на крок” двох моделей, навчених на середньому по величині датасеті. Обидві моделі показують схожий приріст продуктивності на початкових етапах. Однак модель, навчена на 4 класах, демонструє повільніший темп зростання mAP у порівнянні з моделлю, навченою на об’єднаному класі. На середніх кроках модель з об’єднаними класами має стабільно

вищі показники точності порівняно з моделлю на чотирьох класах. Це може свідчити про те, що об'єднана модель краще справляється з узагальненням даних і швидше навчається розпізнавати основні характеристики об'єктів. На фінальних етапах різниця між моделями зберігається, причому об'єднана модель досягає вищого mAP на останніх етапах тренування.

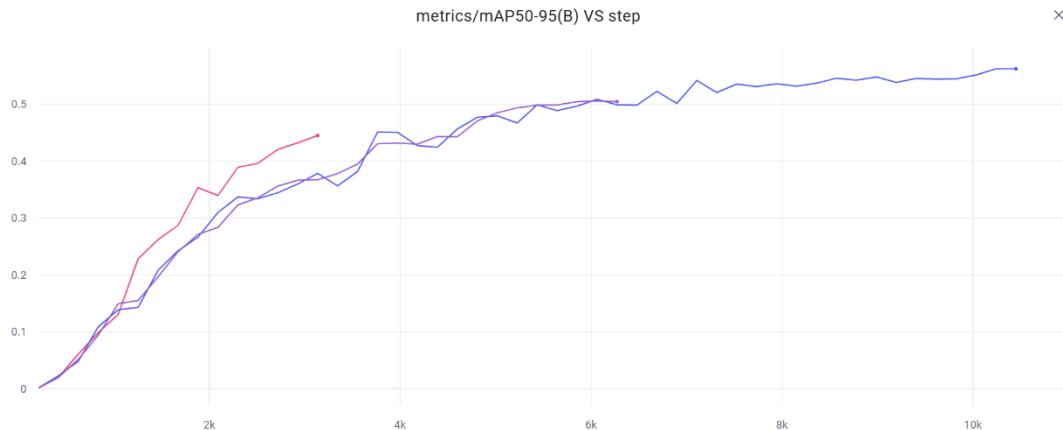


Рис. 3. Графік “mAP50-95 на крок” під час навчання нейромережі

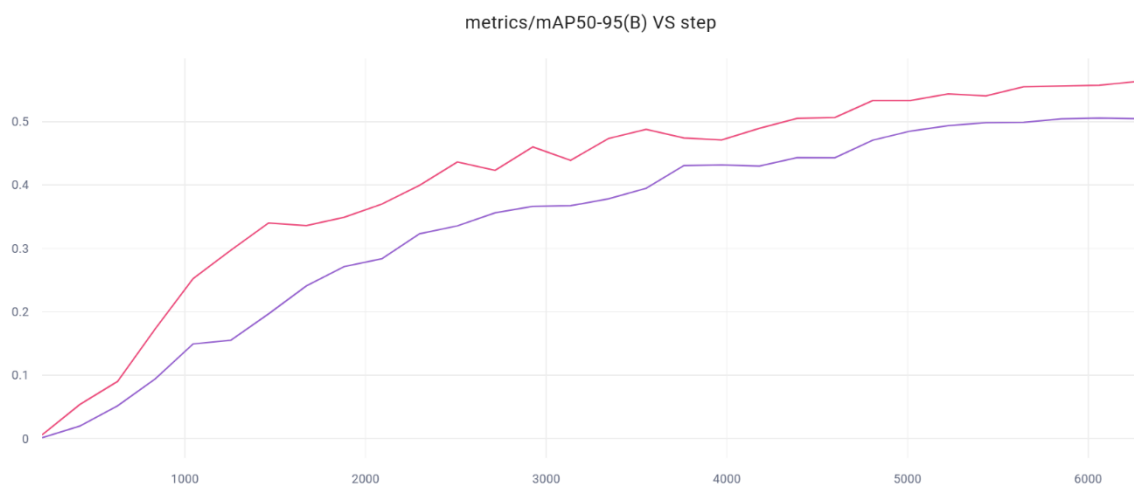


Рис. 4. Графік “mAP50-95 на крок” під час навчання нейромережі

На третьому кроці, після навчання моделей на малому датасеті проведемо їх порівняння. На графіку, зображеному на рисунку 5, можна побачити ефективність процесу навчання. На початкових етапах навчання всі моделі показують схожу результативність, проте згодом їх зростання стає більш стрімким. Модель, навчена на 50 епохах, демонструє найвищий темп поліпшення показників. Наприкінці навчання не спостерігається значного спаду темпу, що свідчить про доцільність продовження навчання на більшій кількості епох.

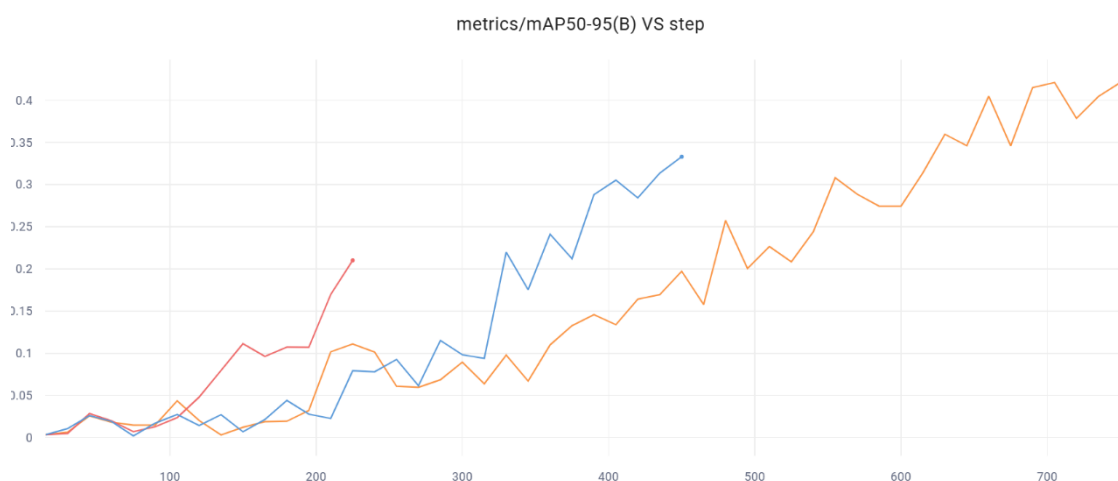


Рис. 5. Графік “mAP50-95 на крок” під час навчання нейромережі

Останнім кроком проведемо загальне порівняння моделей. На рисунку 6 зображено порівняння оптимальних моделей, обраних на минулих етапах, за mAP 50-95 на крок.

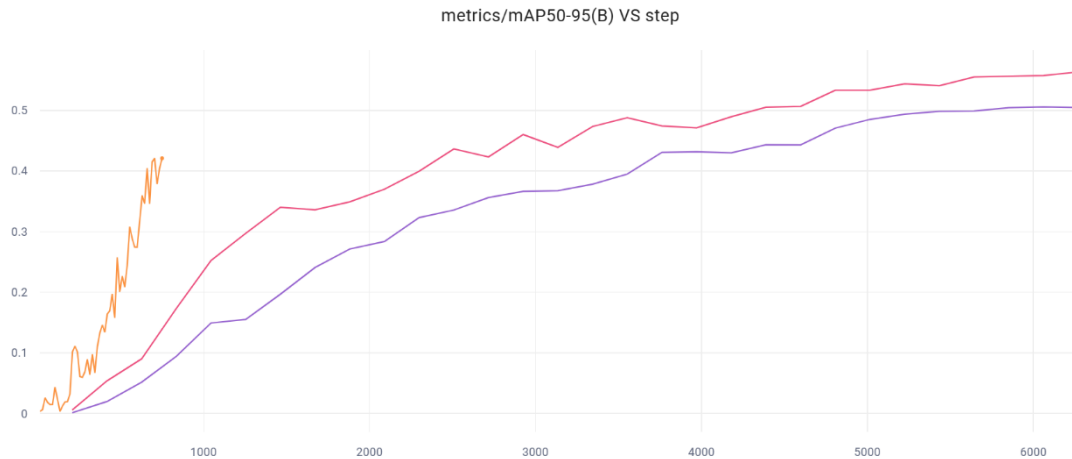


Рис. 6. Графік “mAP50-95 на крок” під час навчання нейромережі

На графіку спостерігається значна різниця між моделями, навченими на середньому та малому датасетах. У випадку малого датасету видно недостатню кількість кроків для досягнення стабільної роботи системи у порівнянні з іншими двома. Отже, оптимальною можна вважати модель, навчену на середньому датасеті з одним класом, протягом тридцяти епох, так як вона показала найкращі результати.

Отриманні під час навчання метрики представлені в таблиці 4.3.

Таблиця 1

Метрики навчених моделей

	Сер. датасет, 4 кл			с.датасет 1 кл	Малий датасет		
Кількість епох	15	30	50	30	15	30	50
Прецензійність	0.642	0.6973	0.724	0.921	0.671	0.925	0.8
Повнота	0.598	0.6687	0.725	0.743	0.3	0.483	0.6
F1-міра	0.619	0.682	0.724	0.822	0.41	0.634	0.68
Втрати	58.953	58.953	49.524	48.267	154.734	102.673	75.799
mAP50	0.654	0.738	0.778	0.829	0.368	0.613	0.672
mAP50-95	0.445	0.504	0.562	0.563	0.210	0.333	0.420

Порівняння трьох різних датасетів – середнього з 4 класами, середнього з 1 класом та малого – виявляє відмінності в продуктивності моделей залежно від розміру і складності даних, а також кількості епох навчання.

Середній датасет з 4 класами демонструє стабільне покращення продуктивності зі збільшенням кількості епох. Прецензійність підвищується з 0.642 до 0.724, а повнота – з 0.598 до 0.725. F1-міра, що відображає баланс між цими метриками, також зростає до 0.724 після 50 епох. Втрати знижуються з 58.953 до 49.524, що свідчить про ефективне навчання. mAP50 значно покращується (з 0.654 до 0.778), тоді як mAP50-95 зростає помірно, досягаючи 0.562. Це вказує на те, що модель ефективно справляється зі складними завданнями детекції, особливо при достатній кількості епох.

Середній датасет з 1 класом має найвищу прецензійність, що досягає 0.921 на 30 епохах, але після 50 епох вона знижується до 0.8, що може вказувати на перенавчання. Повнота досягає максимуму на 30 епохах (0.743), а F1-міра – 0.822. Втрати знижуються значно (від 154.734 до 48.267), що показує стабілізацію навчання. mAP50 досягає найвищого значення (0.829) на 30 епохах, після чого покращення стає мінімальним. Це свідчить про те, що модель швидко навчається на менш складних даних, але подальше навчання може призвести до деградації.

Малий датасет демонструє нестабільні результати через обмежену кількість даних. Прецензійність зростає до 0.925 на 30 епохах, але знижується до 0.8 на 50 епохах. Повнота починається з дуже низького значення (0.3), але поступово зростає до 0.6. F1-міра також зростає, досягаючи максимуму 0.68 на 50 епохах. Втрати залишаються високими порівняно із середніми датасетами, хоча і знижуються зі 154.734 до 75.799. mAP50 та mAP50-95 демонструють поступове покращення, досягаючи 0.672 та 0.420 відповідно. Цей датасет вимагає більше епох для досягнення прийнятних результатів, але через обмеженість даних перенавчання може настати швидше.

Таким чином, навчання на середніх датасетах забезпечує більш стабільні та високі результати, тоді як малий датасет потребує ретельного налаштування кількості епох для уникнення перенавчання і досягнення збалансованої продуктивності.

Висновки

Експериментальне дослідження впливу обсягу даних на ефективність класифікації зокрема за розміром датасету та за кількістю класів номерів автомобілів за допомогою моделі YOLOv11 виявило кілька важливих висновків.

По-перше, результати дослідження підтверджують, що розмір датасету та кількість класів суттєво впливають на точність класифікації. Великі датасети як правило, дають кращі результати, хоча це супроводжується збільшенням обчислювальних витрат. Навчання на малому датасеті часто призводить до зниження точності, що підкреслює важливість різноманіття даних та їх обсягу.

По-друге, навчання на середньому датасеті з більшою кількістю класів виявилось менш ефективним, ніж навчання моделі на датасеті, що містить лише один клас, для задачі детекції автомобільних номерів.

По-третє, кількість епох також має критичне значення в процесі навчання моделі. Аналіз показав, що більша кількість епох не завжди призводить до лінійного покращення результатів. Оптимальна кількість епох залежить від розміру та складності датасету. Менші датасети показують краще збереження результату при навчанні на п'ятдесяти епохах та більше. З іншої сторони, збільшується ризик перенавчання моделі.

Перспективи подальшого розвитку полягають в експериментах із більшими за об'ємом датасетами. Розширення обсягів даних дає можливість покращити та розширити отримані результати. Також варто зрівняти данні результати з моделями навченими на старіших версіях YOLO, та на інших нейромережах.

Література

1. Global EV outlook 2024 – analysis - IEA [Electronic resource] // IEA. – Mode of access: <https://www.iea.org/reports/global-ev-outlook-2024> (date of access: 01.11.2024). – Title from screen.
2. New car registrations: +4.3% in June 2024; battery electric 14.4% market share [Electronic resource] // acea. – Mode of access: <https://www.acea.auto/pc-registrations/new-car-registrations-4-3-in-june-2024-battery-electric-14-4-market-share/> (date of access: 01.11.2024). – Title from screen.
3. Sapkota R. Comparing yolov11 and yolov8 for instance segmentation of occluded and non-occluded immature green fruits in complex orchard environment [Electronic resource] / Ranjan Sapkota, Manoj Karkee // <https://arxiv.org/>. – Mode of access: <https://arxiv.org/pdf/2410.19869> (date of access: 30.10.2024). – Title from screen.
4. Alif M. Yolov11 for vehicle detection: advancements, performance, and applications in intelligent transportation systems [Electronic resource] / Mujadded Alif // <https://arxiv.org/>. – Mode of access: <https://arxiv.org/pdf/2410.22898> (date of access: 01.11.2024). – Title from screen.