DOI: https://doi.org/10.36910/6775-2524-0560-2025-60-16
УДК 004.85
Іванов Дмитро Анатолійович, аспірант
https://orcid.org/0000-0002-7386-4497
Державний університет «Житомирська політехніка», м. Житомир, Україна

МЕТОДИКА ПОЄДНАННЯ ПСЕВДОРОЗМІТКИ ТА СТИСКАННЯ МОДЕЛЕЙ ДЛЯ ОПТИМІЗАЦІЇ САМОНАВЧАННЯ В ЗАДАЧАХ ІДЕНТИФІКАЦІЇ ОБ'ЄКТІВ

Іванов Д. А. Методика поєднання псевдорозмітки та стискання моделей для оптимізації самонавчання в задачах ідентифікації об'єктів. У даній науковій роботі розглядається комбінований підхід до оптимізації навчання за допомогою самонавчання моделей ідентифікації об'єктів, що інтегрує механізм псевдорозмітки з технікою стискання моделей шляхом дистиляції знань (knowledge distillation). Основна мета даного методу полягає в покращенні точності класифікації об'єктів за умов обмеженої кількості маркованих даних та одночасному зниженні обчислювальних витрат, що пов'язані із навчанням глибоких нейронних мереж. Методика підходу передбачає автоматичне включення до навчального набору немаркованих зразків, для яких модель-вчитель демонструє високий рівень впевненості, з подальшим використанням цих зразків для навчання компактнішої моделі-студент. Навчання здійснюється із застосуванням «м'яких» ймовірнісних міток, що дозволяє моделі-вчителю краще засвоїти узагальнені знання. Експериментальні дослідження, проведені на наборах CIFAR-100 та ImageNet, показали, що даний запропонований підхід перевершує традиційні методи самонавчання за ключовими метриками, таким як accuracy, precision, recall, F1-score, та демонструє вищу стабільність і ефективність. Результати дослідження підтверджують потенціал комбінованого підходу для масштабованого застосування в реальних задачах комп'ютерного зору, зокрема в умовах нестачі маркованих даних або при розгортанні моделей на обмежених обчислювальних ресурсах. Додатковою перевагою є можливість зменшення частки помилкових псевдоміток завдяки застосуванню порогового контролю впевненості. Запропонована архітектура навчання також забезпечує кращу процес узагальнення результатів на нові вибірки. Отримані висновки можуть бути використані у наступних дослідженнях для подальшого вдосконалення методів навчання з частковою або нульовою анотацією.
Ключові слова: самонавчання, псевдорозмітка, knowledge distillation, стискання моделей, ідентифікація об'єктів, класифікація зображень.

Ivanov D. A Method for combining pseudo-labeling and model compression to optimize self-training in object Identification tasks. This scientific work examines a combined approach to optimizing training by means of self-training for object identification models, integrating a pseudo-labeling mechanism with a model compression technique through knowledge distillation. The main objective of the proposed method is to improve object classification accuracy under limited labeled data while simultaneously reducing the computational costs associated with training deep neural networks. The methodology involves automatically including unlabeled samples in the training set when the teacher model demonstrates high confidence, followed by using these samples to train a more compact student model. The training process employs soft probabilistic labels, enabling better transfer of generalized knowledge from teacher to student. Experimental results on the CIFAR-100 and ImageNet datasets show that the proposed method outperforms traditional self-training techniques across key metrics such as accuracy, precision, recall, and F1-score, and achieves higher training stability and efficiency. The findings confirm the potential of the combined approach for scalable deployment in real-world computer vision tasks, especially when labeled data is scarce or when deploying models on resource-constrained devices. An additional advantage is the reduced rate of incorrect pseudo-labels due to the application of a confidence threshold. The proposed training framework also improves generalization to unseen samples. These results can be used in future research aimed at enhancing training methods with partially or entirely unlabeled data.
Keywords: self-training, pseudo-labeling, knowledge distillation, model compression, object identification, image classification.

Постановка проблеми. У сучасних умовах широкого впровадження технологій комп'ютерного зору одним із ключових викликів залишається необхідність зменшення залежності від великих обсягів маркованих даних, необхідних для навчання глибоких нейронних мереж. Методи самонавчання, зокрема на основі псевдорозмітки, демонструють обнадійливі результати у задачах ідентифікації об'єктів. Проте їх ефективність суттєво знижується в умовах незбалансованих наборів даних або при роботі з новими доменами, що обмежує можливості їх практичного застосування. Окрім цього, потреба в багаторазовому перенавчанні великих моделей на оновлених вибірках призводить до значного зростання обчислювальних витрат, що ускладнює або унеможливлює їх використання в режимі реального часу чи на пристроях із обмеженими ресурсами.

У контексті вищезазначених проблем виникає науково-практичне завдання розробки комбінованих стратегій, здатних одночасно забезпечити контроль якості автоматично згенерованих міток і підвищити ефективність використання обчислювальних ресурсів. Одним із перспективних підходів є поєднання псевдорозмітки із технологією стискання моделей шляхом дистиляції знань (knowledge distillation), що дозволяє формувати компактні, продуктивні моделі без втрати точності. Розв'язання цієї проблеми має важливе значення як для наукового прогресу в галузі

напівавтоматизованого навчання, так і для прикладних застосувань комп'ютерного зору — від автономних систем до мобільних застосунків.

Стрімкий розвиток систем штучного інтелекту та комп'ютерного зору тісно пов'язаний із підвищенням ефективності навчання глибинних нейронних мереж, що розв'язують задачі ідентифікації об'єктів. Сучасні архітектури, зокрема згорткові нейронні мережі (Convolutional Neural Networks, CNN), трансформери для обробки візуальної інформації (Vision Transformers) та їх гібридні модифікації, демонструють високу точність за умови наявності великомасштабних розмічених наборів даних. Проте створення таких датасетів є ресурсоємним завданням, що потребує значних людських та фінансових витрат.

У цьому контексті особливого значення набувають підходи, спрямовані на оптимізацію процесу навчання, зокрема через впровадження механізмів самонавчання (self-training). Вони дозволяють ефективно використовувати великі обсяги немаркованих даних, тим самим знижуючи потребу в ручній анотації та розширюючи застосовність моделей у прикладних умовах. Самонавчання не лише зменшує залежність від маркованої вибірки, а й може виступати як інструмент підвищення узагальнювальної здатності моделі.

Однак класичні стратегії самонавчання характеризуються низкою недоліків, серед яких — накопичення помилок унаслідок використання ненадійних псевдоміток, а також значні обчислювальні витрати, пов'язані з багаторазовим донавчанням складних моделей. У зв'язку з цим перспективним напрямом є розробка комбінованих підходів до оптимізації процесу навчання, що поєднують псевдорозмітку з методами стискання моделей (knowledge distillation), з метою одночасного підвищення точності та скорочення ресурсних затрат. Такий підхід сприяє більш раціональному використанню даних і обчислювальних ресурсів, забезпечуючи ефективніше впровадження моделей у практичні сценарії.

Актуальність. В умовах зростаючого попиту на адаптивні та обчислювально ефективні моделі комп'ютерного зору, особливо в задачах класифікації, зберігається проблема пошуку методів, які дозволяють досягти високої продуктивності за мінімального залучення маркованих даних і ресурсів. Більшість сучасних моделей демонструють залежність не лише від обсягу даних, а й від обчислювальної складності, що ускладнює їх застосування в умовах обмежених можливостей — наприклад, у мобільних чи вбудованих системах [1, 2].

Методи самонавчання розглядаються як потенційне рішення цієї проблеми, однак у класичному вигляді вони схильні до накопичення похибок через включення до навчального процесу невпевнених або хибних псевдоміток [3]. Це знижує стабільність навчання та обмежує їх застосування у критичних системах. Водночас окреме використання знань попередньо навченої моделі (teacher) у форматі knowledge distillation дозволяє передавати інформацію меншій моделі (student), зменшуючи обчислювальні витрати, але без залучення нових даних така модель не має змоги адаптуватись до розширеного простору задач [4, 5].

Актуальність дослідження обумовлена необхідністю синергетичного підходу, який поєднує переваги обох підходів — використання немаркованих прикладів із контрольованою генерацією псевдоміток та ефективну передачу знань за допомогою дистиляції. Як продемонстровано у роботі RoyChowdhury et al. (2019), така комбінація значно покращує результати на нових доменах навіть за умов мінімального маркованого набору [9], що відкриває нові можливості для розробки універсальних і ресурсозберігаючих систем класифікації.

Метою даного дослідження є розробка та експериментальна перевірка ефективності комбінованої стратегії самонавчання моделей ідентифікації об'єктів, що об'єднує підхід псевдорозмітки з методом стискання моделей (knowledge distillation). Запропонована методика має на меті забезпечити підвищення точності класифікації, стабільність процесу самонавчання, а також зменшення обсягу необхідних маркованих даних і скорочення обчислювальних витрат, пов'язаних із повторним навчанням складних нейронних архітектур.

Огляд літератури. Один із ключових підходів до напівавтоматичного навчання нейронних мереж становить псевдорозмітка (pseudo-labeling), запропонована Lee D.-H [3]. Її суть полягає у використанні прогнозованих міток попередньо навченої моделі для немаркованих даних, за умови високої впевненості у таких прогнозах. Подальший розвиток цього методу реалізовано в підході Noisy Student, описаному в роботі Xie et al. [4], де псевдорозмітка доповнена зашумленням вхідних зображень та фільтрацією псевдоміток за порогом довіри. Метод FixMatch [5] у свою чергу впроваджує консистентну регуляризацію, що дозволяє стабілізувати навчальний процес і зменшити негативний вплив хибних міток.

Інший напрям пов'язаний з методами стискання моделей (knowledge distillation), які передбачають передачу інформації від великої моделі-вчителя до компактнішої моделі-учня через згладжені мітки з температурною нормалізацією [6]. Такі методи дозволяють зменшити обчислювальні витрати, водночас зберігаючи високі показники точності, що робить їх ефективними в контексті розгортання моделей на пристроях з обмеженими ресурсами.

Останні дослідження демонструють синергію між псевдорозміткою та knowledge distillation. Так, метод Label-Guided Distillation (LGD) [10] реалізує підхід внутрішньої дистиляції без зовнішнього вчителя, що робить його ефективним для роботи з CIFAR-100. Метод Smooth and Stepwise Self-Distillation (SSSD) [11] передбачає поетапну дистиляцію на рівні проміжних представлень, що підвищує стабільність навчання на великих наборах, таких як ImageNet.

Позитивні результати комбінування псевдорозмітки та дистиляції знань підтверджуються в дослідженнях, проведених на наборах CIFAR-100 [6] та ImageNet [7], зокрема у працях [4, 10, 11]. Водночас залишається відкритим питання оптимального налаштування гіперпараметрів і вибору архітектур, що забезпечують найкращий компроміс між точністю та ефективністю. З огляду на це, мета запропонованого дослідження полягає в розробці методики, яка забезпечить підвищення продуктивності самонавчання за рахунок поєднання зазначених підходів.

Методологія. У цьому розділі представлено загальні положення експериментальної частини дослідження, зокрема обґрунтування вибору датасетів, що використовуються для оцінювання ефективності запропонованого підходу. Вибрані набори даних відрізняються за обсягом, структурною складністю та типом зображень, що дає змогу здійснити комплексну перевірку методології самонавчання, яка базується на поєднанні псевдорозмітки та стискання знань.

Набір даних CIFAR-100 є одним із найпоширеніших бенчмарків для задач багатокласової класифікації. Він містить 60 000 кольорових зображень розміром 32×32 пікселі, що розподілені між 100 класами, по 600 зображень на кожен клас. З них 50 000 зображень призначено для навчання, а 10 000 — для тестування [6]. Завдяки збалансованій структурі та середньому рівню складності, CIFAR-100 дозволяє ефективно дослідити механізми генерації псевдоміток та вплив методів стискання моделей на стабільність і якість самонавчання. Крім того, широке використання цього датасету в попередніх дослідженнях забезпечує можливість коректного порівняння отриманих результатів із наявними методами.

Набір ImageNet (ILSVRC) є одним із найбільш репрезентативних ресурсів для задач глибинного навчання у сфері комп'ютерного зору. Він охоплює понад 1,2 мільйона зображень високої роздільної здатності, що належать до 1000 різних класів, і суттєво перевершує CIFAR-100 за обсягом, складністю об'єктів, варіативністю та присутністю шумових артефактів [7]. Застосування ImageNet дозволяє перевірити ефективність запропонованого методу в умовах, наближених до реального практичного використання. Крім того, цей набір вважається стандартом де-факто для оцінювання моделей, що робить його релевантною базою для зіставлення результатів із сучасними досягненнями в галузі.

Таким чином, поєднання наборів CIFAR-100 та ImageNet забезпечує всебічну перевірку якості, узагальнення та обчислювальної ефективності розробленого комбінованого методу самонавчання, що є ключовим завданням експериментальної частини дослідження.

У межах даного дослідження запропоновано комбінований підхід до самонавчання моделей ідентифікації об'єктів, який поєднує два ключові методи: псевдорозмітку та стискання моделей. Така інтеграція спрямована на одночасне зменшення залежності від вручну маркованих даних і скорочення обчислювальних витрат, пов'язаних із повторним навчанням глибинних нейронних мереж.

Перший етап передбачає автоматичне генерування псевдоміток за допомогою попередньо навченої нейромережевої моделі (teacher), яка виконує класифікацію немаркованих зразків. Якщо рівень впевненості моделі у певному передбаченні перевищує заданий поріг τ, відповідний зразок включається до розширеного навчального набору з присвоєною псевдоміткою. Відповідно до підходу, адаптованого з попередніх досліджень [3], формалізоване правило включення має вигляд:

$$\hat{y} = arg \max_{y} \mathrm{p}(x|y), \max_{y} \mathrm{p}(x|y) \geq \tau \quad (1)$$

де:

- x – немаркований вхідний зразок;
- $p(x|y)$ – ймовірність класу y;
- τ – поріг впевненості.

У ряді попередніх робіт використовувалося значення $\tau = 0{,}8$. Проте експериментальні результати, отримані в межах цього дослідження, засвідчили, що підвищення порогу до 0,9 дозволяє суттєво зменшити частку хибних псевдоміток, що, своєю чергою, позитивно впливає на стабільність і якість процесу самонавчання.

Після формування розширеного навчального набору здійснюється стискання моделей шляхом використання методики knowledge distillation. Цей підхід передбачає передачу знань від великої, продуктивної, але ресурсомісткої моделі-учителя до меншої та ефективнішої моделі-учня. У процесі дистиляції модель-учень навчається не лише на основі традиційних жорстких міток, а й з урахуванням згладженого розподілу ймовірностей, сформованого моделлю-учителем. Такий підхід дозволяє моделі-учню відтворити не лише остаточне передбачення, але й внутрішні залежності між класами, які містяться в «м'яких» мітках. Цей процес описується формулою (запозичена з роботи [5]):

$$q_i = \frac{\exp(z_i/T)}{\sum_i \exp(z_i/T)},\ L = -\sum_i q_i \log p_i \quad (2)$$

де:

- z_i – вихідні логіти моделі-вчитель для класу i;
- q_i – "м'який" розподіл моделі-вчитель;
- p_i – ймовірності, що видає модель-студент;
- T – температура згладжування (у даному дослідженні рекомендоване значення $T = 3$.

Таким чином, запропонований у межах дослідження комбінований метод самонавчання реалізує поєднання псевдорозмітки та знаннєвої дистиляції. Формули, наведені вище, запозичені з відповідних базових праць і адаптовані до умов поточного експерименту шляхом коригування ключових гіперпараметрів, зокрема порогу впевненості τ та температури згладжування T.

Проведення експерименту. Для реалізації запропонованого підходу були обрані дві сучасні архітектури глибокого навчання, які продемонстрували високу ефективність у задачах комп'ютерного зору. У якості базової (teacher) моделі використано архітектуру EfficientNet-B4, яка забезпечує високу точність класифікації завдяки застосуванню принципу комбінованого масштабування глибини, ширини та роздільної здатності. Здатність цієї моделі формувати достовірні передбачення робить її придатною для генерації псевдоміток у процесі самонавчання.

У ролі компактнішої (student) моделі застосовано архітектуру EfficientNet-B0, яка є спрощеною версією попередньої і характеризується меншою кількістю параметрів. Такий вибір дозволяє знизити обчислювальні витрати, прискорити навчання та забезпечити можливість розгортання моделі на пристроях з обмеженими ресурсами. Передача узагальнених знань від моделі-вчитель до моделі-студент за допомогою підходу knowledge distillation дає змогу досягти прийнятного рівня точності без необхідності повторного навчання складної архітектури.

Процес навчання здійснювався протягом 200 епох для датасету CIFAR-100 та 80 епох для ImageNet. Такий вибір зумовлений відмінностями у масштабах і складності цих наборів даних: CIFAR-100 є порівняно невеликим і дозволяє довготривале навчання без надмірного навантаження на ресурси, тоді як ImageNet потребує обмеження кількості епох для забезпечення ефективного використання обчислювальних потужностей.

З метою досягнення оптимального балансу між швидкістю збіжності та ефективністю використання ресурсів, розмір пакета було встановлено на рівні 128. У якості алгоритму оптимізації застосовано AdamW з початковою швидкістю навчання 1×10^{-3}. Цей оптимізатор є вдосконаленою версією класичного Adam і передбачає використання незалежного згасання ваг, що дозволяє покращити здатність моделі узагальнювати. Такий підхід рекомендовано в низці сучасних досліджень, присвячених глибинному навчанню, зокрема у контексті класифікації та напівавтоматичного навчання.

Порогове значення впевненості для включення немаркованих зразків до навчального набору було встановлено на рівні 0,9. Це дає змогу ефективно відсікати приклади з низькою достовірністю прогнозу, мінімізуючи ризик накопичення помилкових псевдоміток у процесі самонавчання. Доцільність вибору такого значення підтверджується як попередніми публікаціями, так і результатами власного експериментального аналізу, в якому порівняння з менш суворими порогами (наприклад, 0,8) засвідчило перевагу жорсткішого фільтрування.

Температурний коефіцієнт, що використовується в процесі knowledge distillation, встановлено на рівні 3. Таке значення сприяє згладженню вихідного розподілу ймовірностей

моделі-вчитель, надаючи моделі-студент більше інформації про ступінь впевненості у прогнозі для кожного класу. Даний підхід відповідає рекомендаціям оригінальних робіт, присвячених дистиляції знань.

Загалом, обрані гіперпараметри є результатом узгодження між емпіричними спостереженнями та існуючими теоретичними напрацюваннями, що дозволяє забезпечити високу ефективність реалізації запропонованої стратегії самонавчання.

Для оцінювання ефективності розроблених моделей у задачах класифікації об'єктів було застосовано чотири основні метрики: accuracy, precision, recall та F1-score. Використання цих показників забезпечує комплексне уявлення про якість навчання моделі та дозволяє оцінити її поведінку в умовах нерівномірного розподілу класів, що часто спостерігається у реальних наборах даних.

Метрика accuracy (загальна точність) відображає частку правильно класифікованих прикладів серед загальної кількості зразків. Хоча цей показник є релевантним у випадку збалансованих вибірок, його використання при суттєвому домінуванні окремих класів може призводити до упереджених висновків щодо ефективності моделі.

Precision (точність) характеризує частку істинно позитивних передбачень серед усіх позитивних прогнозів моделі. Цей показник є особливо важливим у задачах, де наявні рідкісні класи, оскільки велика кількість хибнопозитивних результатів може значно знизити практичну цінність моделі.

Recall (повнота) відображає здатність моделі виявляти всі наявні об'єкти певного класу. Високі значення recall свідчать про зменшення кількості пропущених об'єктів, що є критично важливим для систем виявлення у випадку довгохвостого розподілу.

Метрика F1-score розраховується як гармонічне середнє між precision та recall, забезпечуючи збалансовану оцінку в умовах, коли існує компроміс між точністю та повнотою. Це особливо важливо в контексті самонавчання, де наявна висока невизначеність і обмежене маркування даних.

Вибір зазначених метрик обумовлений необхідністю всебічної перевірки продуктивності моделей у напівконтрольованому режимі навчання із використанням псевдорозмітки. Крім того, широке застосування цих показників у сучасних дослідженнях забезпечує можливість коректного порівняння результатів із існуючими підходами.

З метою об'єктивної перевірки ефективності запропонованого комбінованого підходу було здійснено порівняльне оцінювання трьох стратегій: базового підходу, що використовує лише псевдорозмітку; ізольованого застосування knowledge distillation; а також комбінованої стратегії, яка поєднує обидва підходи. Експерименти проведено на наборах даних CIFAR-100 та ImageNet з використанням узгоджених метрик ефективності. Результати подано в таблиці 1.

Таблиця 1 – Порівняння ефективності різних методів самонавчання для моделі EfficientNet-B4

Датасет	Метод	Accuracy (%)	Precision	Recall	F1-score
CIFAR-100	Псевдорозмітка (Базовий)	71.8	0.70	0.68	0.69
	Лише knowledge distillation	73.5	0.72	0.70	0.71
	Комбінований (запропонований)	76.3	0.77	0.74	0.75
ImageNet	Псевдорозмітка (Базовий)	69.1	0.68	0.65	0.66
	Лише knowledge distillation	71.0	0.70	0.67	0.68
	Комбінований (запропонований)	74.0	0.74	0.71	0.72

Аналіз отриманих результатів засвідчує перевагу комбінованого підходу до самонавчання, що поєднує механізми псевдорозмітки та стискання знань. Найвищі значення за всіма метриками — точністю класифікації, precision, recall та F1-score — досягнуто саме при використанні комбінованої стратегії. Наприклад, для набору CIFAR-100 accuracy становить 76.3 %, precision – 0.77, recall – 0.74,

а F1-score – 0.75. Це свідчить про покращену здатність моделі до узагальнення та зменшення частки хибнопозитивних і хибнонегативних класифікацій.

Порівняно з базовою стратегією, яка базується лише на використанні псевдорозмітки та демонструє accuracy на рівні 71.8 %, запропонований підхід забезпечує приріст точності на 4.5 відсоткових пункти. Це є суттєвим досягненням для задачі багатокласової класифікації, особливо в умовах обмеженого обсягу маркованих даних. Використання лише knowledge distillation демонструє проміжні результати (accuracy = 73.5 %), що свідчить про позитивний вплив цього методу на стабільність навчання. Однак без залучення нових прикладів, які надає псевдорозмітка, модель обмежена у варіативності вхідних даних.

Аналогічна тенденція спостерігається і для набору ImageNet. Найкращі показники знову демонструє комбінований підхід (accuracy = 74.0 %, precision = 0.74, recall = 0.71, F1-score = 0.72), тоді як базовий метод показує найнижчі значення recall (0.65) та F1-score (0.66), що свідчить про труднощі з виявленням менш представлених класів.

Таким чином, результати проведених експериментів підтверджують доцільність об'єднання методів псевдорозмітки та knowledge distillation у рамках самонавчання. Запропонований комбінований підхід забезпечує кращий компроміс між якістю класифікації та ефективністю використання обчислювальних ресурсів, що робить його перспективним для застосування в реальних умовах комп'ютерного зору з обмеженим доступом до маркованих даних.

Висновки та перспективи подальших досліджень. У цьому дослідженні представлено комбінований підхід до самонавчання моделей ідентифікації об'єктів, що інтегрує механізми псевдорозмітки з контролем впевненості та стискання моделей за допомогою дистиляції знань. Запропонована методика орієнтована на зменшення залежності від маркованих даних та оптимізацію обчислювальних витрат під час перенавчання моделей.

Результати експериментального дослідження, проведеного на наборах CIFAR-100 та ImageNet із використанням архітектури EfficientNet-B4, засвідчили перевагу комбінованого підходу над ізольованими стратегіями. Було досягнуто покращення всіх ключових метрик — точності, точності класифікації, повноти та F1-міри — з максимальним приростом accuracy до 76.3 % для CIFAR-100 та 74.0 % для ImageNet. Також спостерігалося зниження частки хибних псевдоміток та підвищення стабільності процесу самонавчання.

Аналіз отриманих результатів підтверджує ефективність комбінованого підходу як інструмента для підвищення якості моделей за збереження або навіть скорочення обчислювальних ресурсів. Це робить його придатним для застосування в умовах обмеженого доступу до маркованих даних, а також перспективним для розгортання у реальних сценаріях комп'ютерного зору. Надалі доцільним є розширення дослідження шляхом тестування різних архітектур, адаптації методу до задач детекції та сегментації, а також подальша інтеграція результатів у дисертаційну роботу.

У подальших дослідженнях передбачається розширення запропонованої методики шляхом її адаптації до задач із незбалансованим розподілом класів, зокрема в умовах long-tail, а також розгляд впливу різних стратегій семплінгу псевдорозмічених прикладів. Перспективним напрямом є інтеграція елементів активного навчання та використання мультимодальних представлень для підвищення релевантності псевдоміток. Крім того, доцільно дослідити застосування підходу в низькоресурсних середовищах, зокрема на мобільних або вбудованих пристроях.

Список бібліографічного опису:

1. Gupta A., Dollar P., Girshick R. LVIS: A Dataset for Large Vocabulary Instance Segmentation // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019. P. 5356–5364. URL: https://openaccess.thecvf.com/content_CVPR_2019/papers/Gupta_LVIS_A_Dataset_for_Large_Vocabulary_Instance_Segmentation_CVPR_2019_paper.pdf.
2. Caesar H., Bankiti V., Lang A. та ін. nuScenes: A Multimodal Dataset for Autonomous Driving // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020. P. 11621–11631. URL: https://openaccess.thecvf.com/content_CVPR_2020/html/Caesar_nuScenes_A_Multimodal_Dataset_for_Autonomous_Driving_CVPR_2020_paper.html.
3. Lee D.-H. Pseudo-label: The Simple and Efficient Semi-Supervised Learning Method for Deep Neural Networks // ICML Workshop. 2013. URL: https://github.com/emintham/Papers/blob/master/Lee-%20Pseudo-Label%3A%20The%20Simple%20and%20Efficient%20Semi-Supervised%20Learning%20Method%20for%20Deep%20Neural%20Networks.pdf.
4. Xie Q., Luong M.-T., Hovy E., Le Q. V. Self-training with Noisy Student Improves ImageNet Classification // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020. P. 10687–10698. URL: https://openaccess.thecvf.com/content_CVPR_2020/html/Xie_Self-Training_With_Noisy_Student_Improves_ImageNet_Classification_CVPR_2020_paper.html.

5. Hinton G., Vinyals O., Dean J. Distilling the Knowledge in a Neural Network. 2015. URL: https://arxiv.org/abs/1503.02531.
6. Krizhevsky A. Learning Multiple Layers of Features from Tiny Images: Technical Report. 2009. URL: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf.
7. Deng J., Dong W., Socher R. та ін. ImageNet: A Large-Scale Hierarchical Image Database // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2009. P. 248–255. URL: https://ieeexplore.ieee.org/document/5206848.
8. Tan M., Le Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks // Proceedings of the 36th International Conference on Machine Learning (ICML), 2019. P. 6105–6114. URL: https://arxiv.org/abs/1905.11946.
9. RoyChowdhury A., Chakrabarty P., Singh A. та ін. Automatic Adaptation of Object Detectors to New Domains Using Self Training. 2019. URL: https://arxiv.org/abs/1904.07305.
10. Zhang P., Kang Z., Yang T. та ін. LGD: Label Guided Self Distillation for Object Detection. 2021. URL: https://arxiv.org/abs/2109.11496.
11. Deng J., Zhou X., Tian H. та ін. Smooth and Stepwise Self Distillation for Object Detection (SSSD). 2023. URL: https://arxiv.org/abs/2303.05015.

References:

1. Gupta A., Dollar P., Girshick R. LVIS: A dataset for large vocabulary instance segmentation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019. P. 5356–5364. URL: https://openaccess.thecvf.com/content_CVPR_2019/papers/Gupta_LVIS_A_Dataset_for_Large_Vocabulary_Instance_Segmentation_CVPR_2019_paper.pdf.
2. Caesar H., Bankiti V., Lang A. et al. nuScenes: A multimodal dataset for autonomous driving. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020. P. 11621–11631. URL: https://openaccess.thecvf.com/content_CVPR_2020/html/Caesar_nuScenes_A_Multimodal_Dataset_for_Autonomous_Driving_CVPR_2020_paper.html.
3. Lee D.-H. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. ICML Workshop, 2013. URL: https://github.com/emintham/Papers/blob/master/Lee-%20Pseudo-Label%3A%20The%20Simple%20and%20Efficient%20Semi-Supervised%20Learning%20Method%20for%20Deep%20Neural%20Networks.pdf.
4. Xie Q., Luong M.-T., Hovy E., Le Q. V. Self-training with noisy student improves ImageNet classification. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020. P. 10687–10698. URL: https://openaccess.thecvf.com/content_CVPR_2020/html/Xie_Self-Training_With_Noisy_Student_Improves_ImageNet_Classification_CVPR_2020_paper.html.
5. Hinton G., Vinyals O., Dean J. Distilling the knowledge in a neural network. 2015. URL: https://arxiv.org/abs/1503.02531.
6. Krizhevsky A. Learning multiple layers of features from tiny images: Technical report. 2009. URL: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf.
7. Deng J., Dong W., Socher R. et al. ImageNet: A large-scale hierarchical image database. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2009. P. 248–255. URL: https://ieeexplore.ieee.org/document/5206848.
8. Tan M., Le Q. V. EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning (ICML), 2019. P. 6105–6114. URL: https://arxiv.org/abs/1905.11946.
9. RoyChowdhury A., Chakrabarty P., Singh A. et al. Automatic adaptation of object detectors to new domains using self training. 2019. URL: https://arxiv.org/abs/1904.07305.
10. Zhang P., Kang Z., Yang T. et al. LGD: Label guided self distillation for object detection. 2021. URL: https://arxiv.org/abs/2109.11496.
11. Deng J., Zhou X., Tian H. et al. Smooth and stepwise self distillation for object detection (SSSD). 2023. URL: https://arxiv.org/abs/2303.05015.