

О. М. Кириленко¹
Р. Н. Квєтний¹
В. В. Гармаш¹
В. Ю. Коцюбинський¹

ЗАСТОСУВАННЯ АДАПТИВНИХ ГІПЕРМОДЕЛЕЙ У ЗАДАЧАХ РЕ-ІДЕНТИФІКАЦІЇ ЛЮДИНИ

¹Вінницький національний технічний університет

Проаналізовано наявні підходи до розробки та впровадження адаптивних гіпермоделей у сфері повторної ідентифікації людей. Вони базуються на порівняльному вивченні роботи моделей, застосовуваних на різних наборах даних, охоплюючи як лабораторні сценарії, так і реальні умови експлуатації. Основну увагу приділено оцінюванню ключових показників якості, зокрема точності розпізнавання (mAP, CMC) та швидкості обробки інформації, що дозволяє всебічно охопити ефективність застосовуваних методів.

Проведено дослідження, яке зосереджується на аналізі впливу динамічних змін параметрів на результати роботи моделей, а також на вивченні стратегій інкрементального навчання, які сприяють зниженню ризику катастрофічного забування при адаптації до нових умов без необхідності повного перенавчання. Завдяки такому підходу система здатна оперативно реагувати на зміни умов зйомки, наприклад, варіації в освітленні, ракурсах та інших характеристиках, що є критично важливим для систем відеоспостереження.

На основі проведеного аналізу окреслено перспективні напрямки подальших досліджень, спрямовані на вдосконалення алгоритмів адаптивного навчання, розробку нових архітектурних рішень та оптимізацію процесів масштабування. Це також сприятиме впровадженню надійніших та ефективніших технологій повторної ідентифікації в сучасних інформаційних системах.

Запропонований підхід поєднання гіпермоделі з основною глибокою нейронною мережею, ключовою перевагою якого є його висока адаптивність і стабільність навчання, забезпечена використанням динамічного коригування параметрів за допомогою гіпермоделей. А також поєднання крос-ентропійної та триплетної втрат, що дозволяє ефективно формувати компактні та роздільні ознаки для різних ідентичностей та підвищує здатність моделі ідентифікувати об'єкт навіть у випадках значної варіативності вхідних даних.

Результати дослідження демонструють перспективність інтеграції адаптивних механізмів у сучасні системи повторної ідентифікації, забезпечуючи підвищену стійкість до змін умов експлуатації та високий рівень продуктивності, що є необхідною умовою для успішного практичного застосування в інформаційних технологіях.

Ключові слова: повторна ідентифікація, гіпермоделі, динамічна адаптація, інкрементне навчання, CNN, Transformer, OSNet, Market-1501, DukeMTMC-ReID, MSMT17.

Вступ

Задача повторної ідентифікації людини сьогодні набуває особливої важливості, адже вона знаходить широке застосування у системах безпеки, реалізації концепції «розумних» міст, а також у аналізі поведінки відвідувачів як комерційних, так і громадських просторів. Суть проблеми полягає у здатності алгоритмів визначати, чи належить зображення або відеофрагмент до тієї ж особи, що представлена на інших кадрах або збережена в окремих базах даних, незважаючи на суттєві варіації в умовах зйомки, таких як різний кут огляду, зміни освітлення, а також можливі зміни зовнішнього вигляду через зміну одягу чи аксесуарів [1], [2].

Досягнення значного прогресу у вирішенні завдань повторної ідентифікації забезпечено завдяки використанню глибоких нейронних мереж, зокрема архітектур типу CNN, а також трансформерів, які тренувалися на наборах маркованих даних, як-от Market-1501 [4], DukeMTMC-ReID [5] та

MSMT17 [6]. Ці сучасні моделі демонструють високі показники точності, що вимірюються за допомогою таких метрик, як mAP та CMC Rank-1, і стабільно працюють у різних експериментальних умовах. Проте, у практичних системах відеоспостереження виникають додаткові специфічні ситуації. Дані надходять з нових камер, які можуть запропонувати зовсім інші ракурси зйомки, а умови експлуатації змінюються динамічно, враховуючи сезонні зміни, погодні умови чи особливості інтер'єрів. До того ж, стрімке зростання обсягів інформації ставить під сумнів можливості традиційних підходів до навчання моделей, оскільки повне перенавчання з нуля вимагає значних обчислювальних ресурсів та є затратним за часом [1].

Проблема ускладнюється також тим, що традиційні методи донавчання, які базуються на інкрементальному або постійному навчанні, часто стикаються з феноменом катастрофічного забування. Суть цього полягає в тому, що адаптація моделі до нових умов може призвести до втрати здатності коректно працювати на попередньо опрацьованих даних або сценаріях, оскільки стандартні методи регуляризації не завжди можуть забезпечити оптимальний баланс між накопиченими «старими» знаннями та новою інформацією. Це питання є особливо гострим у задачах повторної ідентифікації, де системі необхідно розрізняти потенційно тисячі осіб у середовищі, яке характеризується великою кількістю камер та різноманітним умовам зйомки.

Одним з перспективних напрямків вирішення зазначених проблем є застосування гіпермоделей (Hypernetworks). Основна ідея полягає в тому, що окрема «гіпермодель» здатна динамічно генерувати або коригувати параметри основної мережі (Target Network), що дозволяє зберігати інформацію про попередні налаштування навіть у разі оновлення моделі. Такий підхід робить процес адаптації значно гнучкішим і менш схильним до втрати попередніх накопичених знань, що вкрай важливо для систем, які повинні швидко «підлаштовуватися» під нові умови, такі як зміни ракурсів або специфічні характеристики середовища. Завдяки цьому гіпермоделі набувають високої привабливості як для реалізації масштабованих рішень у сфері відеоспостереження, так і для впровадження в інтелектуальні системи, що працюють у режимі реального часу [2], [3].

Метою роботи є підвищення точності повторної ідентифікації за рахунок підходу використання адаптивних гіпермоделей та глибоких нейронних мереж, а також оцінка ефективності моделей для подальшого використання в інформаційних технологіях.

Аналіз наявних архітектур гіпермоделей

Існує декілька основних підходів до побудови гіпермоделей, які відрізняються як структурою, так і принципами динамічного генерування параметрів основної мережі. Один з перших підходів запропоновано в роботі [2], де окрема мережа, що має набагато меншу складність, генерує вагові коефіцієнти для цільової мережі. Цей підхід дозволяє значно зменшити кількість параметрів, що потребують прямого навчання, та забезпечує можливість адаптації моделі до нових умов. Проте, незважаючи на свою універсальність, така архітектура має недолік нестабільності при генеруванні високорозмірних ваг, що особливо помітно під час роботи з великими базами даних чи за високої різноманітності вхідних зображень.

Іншим напрямком [3] є динамічні фільтрові мережі, які генерують параметри (фільтри) безпосередньо з вхідних даних для здійснення адаптивного згорткового перетворення. Цей підхід дозволяє моделі враховувати конкретні особливості кожного зображення. Проте динамічна генерація фільтрів зазвичай супроводжується додатковими витратами обчислювальних ресурсів і вимагає ретельного налаштування, щоб уникнути перенавчання або надмірної варіативності згенерованих параметрів.

Ще один варіант [7] — це умовні гіпермоделі, які генерують лише частину параметрів цільової мережі або ж працюють як модуль, що вносить корективи до проміжних ознак, отриманих основною мережею. Завдяки такому підходу можна зберегти більшу стабільність основної архітектури. Такий підхід дозволяє забезпечити кращий баланс між адаптивністю і обчислювальною ефективністю, що особливо важливо в задачах, де швидкість обробки відеопотоків та точність розпізнавання мають першочергове значення.

Також варто зазначити останні модифікації [8], які базуються на факторизації ваг або використанні низькорозмірних апроксимацій, що дозволяють значно зменшити навантаження на пам'ять при генерації параметрів. Ці підходи часто інтегруються з умовними гіпермоделями і націлені на збереження балансу між адаптивністю та обчислювальною ефективністю.

Аналізуючи переваги та недоліки кожного з підходів, можна зазначити, що традиційна гіпермодель, яка генерує повні вагові матриці для основної мережі, забезпечує велику гнучкість і може

врахувати широкий спектр варіацій даних, проте її застосування часто обмежене високою обчислювальною складністю та схильністю до нестабільності в роботі. Динамічні фільтрові мережі дозволяють моделі точніше адаптуватися до кожного конкретного вхідного зображення, що позитивно впливає на розпізнавання за значних варіацій умов зйомки, але при цьому вони потребують додаткових ресурсів для обчислення параметрів в режимі реального часу. Умовні гіпермоделі, що працюють як модулі корекції, пропонують оптимальний компроміс між адаптивністю та ефективністю, оскільки дозволяють оновлювати лише частину параметрів, зберігаючи стабільність основної архітектури та знижуючи ризик катастрофічного забування.

Враховуючи специфіку задачі повторної ідентифікації людини, де особлива увага приділяється здатності моделі адаптуватися до змін умов зйомки, варіацій ракурсів, освітлення та індивідуальних особливостей, найперспективнішим підходом є використання умовної гіпермоделі. Така архітектура дозволяє інтегрувати адаптивний модуль, який генерує коригувальні параметри для ключових проміжних шарів основної нейронної мережі, що безпосередньо відповідає за вилучення ознак зображень. Завдяки цьому підходу модель зберігає високу загальну продуктивність та точність, одночасно адаптуючись до нових умов без необхідності повного перенавчання. До того ж, така архітектура забезпечує оптимальний баланс між обчислювальними витратами та адаптивністю, що є критично важливим у системах відеоспостереження, де дані надходять у режимі реального часу з різних джерел.

Таким чином можна підсумувати, що з-поміж наявних архітектур гіпермоделей умовна гіпермодель, що застосовує модульну адаптацію параметрів цільової мережі, є найпридатнішою для задач повторної ідентифікації людини.

Математичний опис моделі умовної гіпермоделі

Нехай основна (цільова) мережа, яку будемо позначати як F , приймає зображення $x \in X$ та генерує його представлення $z \in R^d$. Ця мережа параметризована набором ваг W (які можуть бути поділені на параметри окремих шарів, наприклад $W = \{W_1, W_2, \dots, W_L\}$ для L шарів). Для того, щоб адаптувати модель до змінних умов (наприклад, відмінностей між камерами, умов освітлення, ракурсів тощо), вводиться умовний вхід $c \in C$, який кодує інформацію про ці умови (до прикладу, ідентифікатор камери, характеристики середовища, метадані) [9].

Умовна гіпермодель позначається як G і має параметри θ . Її основна задача полягає у генерації або модулюванні параметрів цільової мережі залежно від умовного сигналу c . Формально, можна записати, що для кожного шару l цільової мережі маємо

$$W_l(c) = \phi_l(c; \theta_l), \quad (1)$$

де $\phi_l(\cdot)$ — деяка функція, що перетворює умовний вхід c у модифіковані параметри або корекційні величини для l -го шару. Загальна система може працювати за різними схемами модифікації параметрів. Розглянемо два основні підходи:

Адаптація шляхом доповнення до базових ваг. Позначимо базові (статичні) параметри цільової мережі як W_l^0 . Тоді умовна гіпермодель генерує корекційний тензор $\Delta W_l(c)$, після чого ефективні параметри визначаються як

$$W_l(c) = W_l^0 + \Delta W_l(c); \quad (2)$$

$$\Delta W_l(c) = \phi_l(c; \theta_l). \quad (3)$$

Адаптація шляхом модулювання. У цьому випадку умовна гіпермодель генерує вектор, який масштабує коефіцієнти $a_l(c)$, після чого ваги кожного шару оновлюються так:

$$W_l(c) = \text{diag}(a_l(c))W_l^0, \quad (4)$$

де $\text{diag}(a_l(c))$ — діагональна матриця, що містить коефіцієнти модулювання, отримані так:

$$a_l(c) = \psi_l(c; \theta_l). \quad (5)$$

У загальному випадку, незалежно від вибору схеми адаптації, остаточно модель F набуває вигляду

$$z = F\left(x; \{W_l(c)\}_{l=1}^L\right), \quad (6)$$

де параметри кожного шару $W_l(c)$ є функцією від умовного вхідного сигналу c [10], [11].

Для задач повторної ідентифікації людини умовний вхід c може включати такі компоненти:

- ідентифікатор камери: дозволяє враховувати специфічні особливості зйомки з різних камер, які можуть мати різні кутові та оптичні характеристики.
- метадані про умови зйомки: інформація про освітлення, пору доби або навіть погодні умови.
- контекстуальні ознаки: додаткові дані про оточення, які можуть впливати на видимість чи контрастність зображення [12].

Завдяки такій структурі, умовна гіпермодель генерує корекційні параметри, які адаптують кожен шар цільової мережі до поточного контексту. Це дозволяє моделі точніше виловити релевантні ознаки та підвищувати загальну точність розпізнавання, що є особливо критичним в роботі з великими наборами даних, отриманими з різних джерел у реальному часі.

Оскільки цільова задача полягає у повторній ідентифікації людини, система зазвичай оптимізується за допомогою відповідної функції втрат, наприклад, комбінації крос-ентропійної втрати L_{CE} (для класифікації ідентичностей) та втрати на основі триплетів $L_{triplet}$ (для навчання відстаневих відношень)

$$L = \lambda_{CE} L_{CE}(F(x; W(c)), y) + \lambda_{triplet} L_{triplet}(F(x; W(c))), \quad (7)$$

де y — мітка ідентичності, а λ_{CE} та $\lambda_{triplet}$ — вагові коефіцієнти. Під час навчання відбувається спільна оптимізація як параметрів базової мережі W^0 , так і параметрів умовної гіпермережі θ . Це може бути організовано у вигляді спільної мінімізації [13]:

$$\min_{W^0, \theta} \mathbb{E}_{x, c, y} [L(F(x; W_l(c)), y)]. \quad (8)$$

Цей підхід дозволяє моделі динамічно адаптувати свої параметри під конкретні умови, забезпечуючи оптимальний баланс між збереженням попередньо набутого знання (зниження катастрофічного забування) та адаптацією до нових даних, що є ключовим для успішного застосування у системах повторної ідентифікації людини [14], [16].

Оцінювання ефективності моделі здійснюється за допомогою стандартних метрик, зокрема середньої точності (mean Average Precision, mAP) та кумулятивної характеристики зіставлення (Cumulative Matching Characteristic, CMC), з акцентом на Rank-1.

Cumulative Matching Characteristics (CMC) є широко використовуваним підходом для оцінювання результатів в задачах повторної ідентифікації людини. Це криві, які показують залежність ймовірності правильної ідентифікації від кількості найподібніших зразків у галереї, що враховуються під час аналізу запиту. В процесі обчислення CMC модель формує ранжування зразків за ступенем їхньої схожості до зображення запиту. Відповідно, показник точності CMC top-k визначається таким чином [1]:

$$A_{c_k} = \{1, \text{якщо зразки галереї з найвищим рейтингом містять ідентифікатор запиту}, 0 \text{ — інше}\},$$

якщо хоча б один зі зразків галереї, що має найвищий рейтинг подібності до запиту, відповідає шуканій особі, то точність дорівнює одиниці.

Ще однією важливою метрикою в задачах повторної ідентифікації є середня точність (*Mean Average Precision, mAP*). Вона широко застосовується для оцінки продуктивності моделей у завданнях виявлення та пошуку об'єктів, особливо у випадках, коли виникає потреба у точній ідентифікації осіб. Основу для її розрахунку складають такі показники, як точність (*Precision*) та відкликання (*Recall*). *Precision* розраховується як відношення числа істинно позитивних результатів (*TP*) до суми істинно позитивних і хибнопозитивних результатів (*TP + FP*). *Recall* визначається як співвідношення істинно позитивних результатів (*TP*) до суми істинно позитивних та помилково негативних (*TP + FN*).

Середня точність (*Average Precision, AP*) для конкретного класу обчислюється як середнє значення показника *Precision*, що вимірюється на різних рівнях відкликання, виражається так:

$$AP = \frac{1}{R} \sum_{r \in R} Precision(r). \quad (9)$$

Mean Average Precision (mAP), своєю чергою, визначає середнє значення середньої точності по всіх класах і є кінцевим агрегованим показником продуктивності моделі. Ця метрика активно ви-

користується для порівняння та оцінки продуктивності сучасних моделей у задачах повторної ідентифікації особи та виявлення об'єктів завдяки своїй здатності враховувати якість розподілу ознак по всіх представлених категоріях [1], [17].

$$\text{mAP} = \frac{1}{M} \sum_{m \in M} \text{AP}(m). \quad (10)$$

Опис експерименту та навчання моделей

У експерименті використовуються дві взаємодоповнювальні компоненти:

– Гіпермодель, що приймає на вхід вектор ознак розмірності `feature_dim` і генерує адаптивне представлення в ембеддинговому просторі розмірності `embedding_dim`. Архітектура побудована за принципом послідовності повнозв'язних шарів з використанням нормалізації, нелінійного перетворення за допомогою функції ReLU та регуляризації за допомогою Dropout. Важливою особливістю є попереднє нормування вхідних ознак (L2-нормалізація), що забезпечує стабільність подальшої генерації ембеддингів. Таким чином, модуль виконує роль адаптивного компонента, який може бути

використаний для коригування представлення ознак, отриманих від основної мережі.

– Основна мережа побудована на базі попередньо навченого backbone, з якого вилучено останній класифікаційний шар. Для посилення дискретних ознак використовується окремий модуль уваги над просторовими ознаками, отриманими з глибинних шарів мережі. Просторові ознаки передаються через блок адаптивного усереднення, лінійне перетворення і функцію активації ReLU для формування кінцевих ембеддингів. За необхідності застосовується BNNeck, що дозволяє поліпшити розподіл ознак для задач класифікації. Остаточний класифікатор реалізовано як лінійний шар, який відображає ембеддинг у простір кількості класів. У методі forward організовано послідовне вилучення просторових ознак з backbone з подальшим застосуванням модуля уваги, що дозволяє отримати релевантніше представлення для задачі повторної ідентифікації.

Під час навчання здійснюється спільна оптимізація параметрів базової мережі та умовної гіпермоделі згідно з функцією втрат, що враховує вимоги до високої точності ідентифікації та стійкості системи до змін умов зйомки.

На рис. 1 показана блок-схема роботи алгоритму умовної гіпермоделі для задачі повторної ідентифікації людини

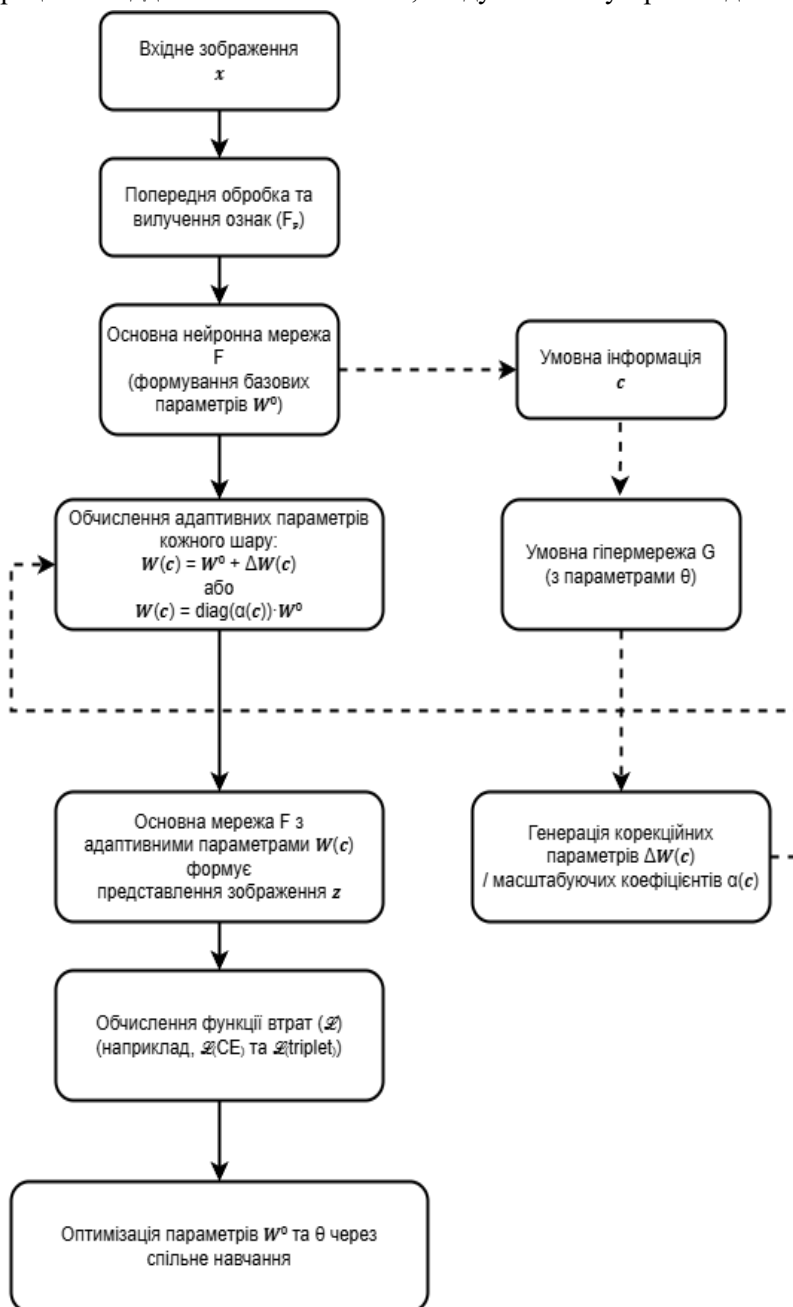


Рис. 1. Блок-схема роботи алгоритму умовної гіпермоделі для задачі повторної ідентифікації людини

Навчання та тестування нейромережових моделей проводилося з використанням сучасного фреймворка для глибокого навчання PyTorch [12]. У процесі розробки моделі для повторної ідентифікації людей основну увагу приділено інтеграції двох ключових компонентів: основної мережі, яка забезпечує ефективне вилучення просторових ознак із зображень за допомогою попередньо навченого backbone та механізму уваги, адаптивного модуля гіпермоделі, який перетворює отримані ознаки у компактний ембедінг із застосуванням лінійних шарів, нормалізації та функцій активації.

Основні особливості вибраного підходу включають:

- використання попередньо навчених моделей (OSNet) для вилучення базових ознак, що дозволяє прискорити процес навчання;
- застосування модуля уваги для підсилення релевантних просторових характеристик зображень;
- адаптивну трансформацію векторних представлень за допомогою гіпермережі, яка забезпечує динамічну корекцію параметрів залежно від умов зйомки;
- можливість інтеграції додаткових модулів нормалізації (BNNeck) для поліпшення розподілу ознак;
- підтримку навчання на кількох графічних процесорах та оптимізацію з використанням сучасних алгоритмів;
- аналіз якості моделей у режимі реального часу із застосуванням стандартних метрик (mAP, Rank-1, CMC) на відкритих наборах даних, таких як Market-1501 та DukeMTMC-reID.

Під час навчання основна мережа обробляє зображення, які спочатку проходять через шари попередньої обробки та вилучення просторових ознак, після чого модуль уваги виділяє найзначущі характеристики. Отримані ознаки передаються через блок spatial-to-embedding для формування фінального ембедінгу, який, за потреби, нормалізується та використовується для класифікації. Гіпермережа адаптивно коригує ці ознаки, що дозволяє моделі досягати високої точності при ідентифікації осіб.

Експериментальна перевірка проводилась із застосуванням стандартних наборів даних для задач повторної ідентифікації, що дозволило оцінити продуктивність моделі за показниками mAP, Rank-1 та іншими метриками. На рис. 2 показано динаміку зміни якості моделей під час навчання, що демонструє ефективність запропонованого підходу у порівнянні з традиційними методами.

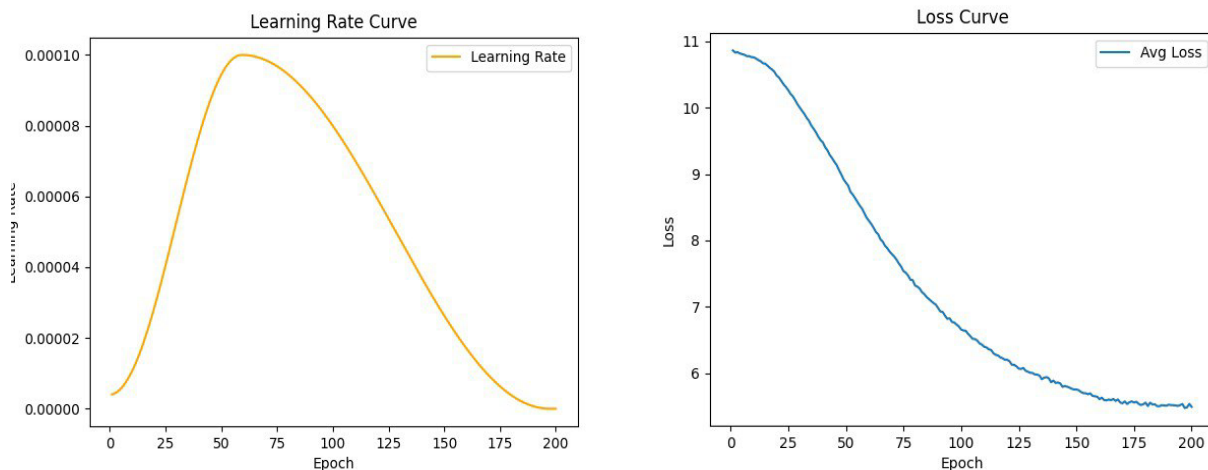


Рис. 2. Вплив динамічних змін параметрів на результати роботи моделей

У процесі навчання застосовується комбінована функція втрат, що містить крос-ентропійний компонент і триплетну втрату (на рис. 2 зображені криві зміни показників якості моделі в процесі навчання). Такий підхід сприяє максимізації відстаней між представленнями різних осіб та одночасно мінімізує відстані між представленнями однакових осіб. Результати проведених експериментів на датасетах Market-1501 та DukeMTMC-reID подано в табл. 1 та 2 відповідно.

Таблиця 1

Результати навчання моделей на наборі зображень Market-1501

ID	dataset	optim	lr	Base lr	epoch	mAP, %	Rank-1, %	Rank-5, %	Rank-10, %
ExpBN37	market-1501	adam	1e-4	1e-5	200	77,43	96,94	99,11	99,44
ExpBN38	market-1501	adam	1e-4	1e-5	200	77,00	96,56	99,17	99,44

Продовження табл. 1

ID	dataset	optim	lr	Base lr	epoch	mAP, %	Rank-1, %	Rank-5, %	Rank-10, %
ExpBN39	market-1501	adam	1e-4	1e-5	200	76,17	96,91	99,20	99,41
ExpBN40	market-1501	adam	1e-4	1e-5	200	77,03	96,70	99,38	99,64
ExpBN41	market-1501	adam	1e-4	1e-5	200	76,75	96,53	99,23	99,52
ExpBN42	market-1501	adam	1e-4	1e-5	200	78,25	96,32	99,23	99,52

Таблиця 2

Результати навчання моделей на наборі зображень DukeMTMC-reID

ID	dataset	optim	lr	Base lr	epoch	mAP, %	Rank-1, %	Rank-5, %	Rank-10, %
ExpBN49	dukeMTMC-reID	adam	1e-4	1e-5	200	67,40	84,50	92,00	94,80
ExpBN50	dukeMTMC-reID	adam	1e-4	1e-5	200	67,10	84,00	92,30	94,80
ExpBN51	dukeMTMC-reID	adam	1e-4	1e-5	200	66,80	84,00	92,20	94,90
ExpBN52	dukeMTMC-reID	adam	1e-4	1e-5	200	67,20	84,10	92,80	95,10
ExpBN53	dukeMTMC-reID	adam	1e-4	1e-5	200	66,90	83,90	92,30	95,00
ExpBN54	dukeMTMC-reID	adam	1e-4	1e-5	200	67,60	83,60	92,30	95,00

Експерименти проводились для оцінки ефективності моделей з використанням адаптивної гіпермоделі на двох стандартних наборах даних: Market-1501 та DukeMTMC-reID. Для обох наборів використовували оптимізатор Adam з початковою швидкістю навчання $1e-4$, базовою швидкістю $1e-5$ та warmup-періодом у епохах, загальна кількість епох становила 200.

На наборі Market-1501 моделі з адаптивною гіпермоделлю демонструють значне підвищення продуктивності порівняно з базовою моделлю OSNet без гіпермоделі [1]. При цьому середня точність (mAP) зросла з 73,5 % до діапазону 75,4...78,3 %, а показник Rank-1 — з 90 % до приблизно 96...97 %. Це свідчить про ефективне вилучення релевантних ознак та адаптивне коригування представлень, що дозволяє краще розрізняти різні ідентичності навіть за умов високої варіативності зображень.

Аналогічна тенденція спостерігається і на наборі DukeMTMC-reID, де базова мережа OSNet дає mAP 62,81 % і Rank-1 81,7 %. Моделі з гіпермоделлю досягають mAP в діапазоні 65,1...67,6 % та Rank-1 від 83,5 % до 84,5 %, що свідчить про підвищення здатності системи до розпізнавання осіб на складніших і різноманітніших даних. На рис. 3 показано приклад роботи моделі.

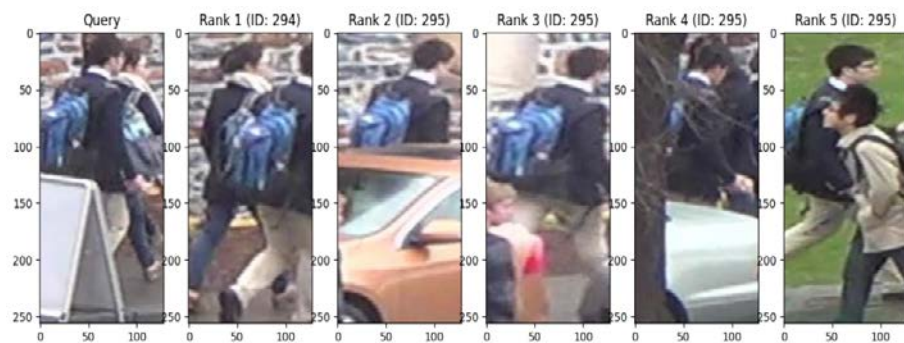


Рис. 3. Приклад роботи моделі

Таким чином, інтеграція адаптивної гіпермоделі суттєво поліпшує результати повторної ідентифікації як на Market-1501, так і на DukeMTMC-reID, забезпечуючи точнішу і стійкішу класифікацію осіб порівняно з традиційним підходом без гіпермоделі.

Висновки

Проведено аналіз відомих архітектур гіпермоделей: гіпермоделі динамічного генерування параметрів основної мережі, динамічні фільтрові мережі та умовні гіпермоделі. На підставі проведеного аналізу встановлено, що використання адаптивних гіпермоделей у задачах повторної ідентифікації людини має суттєві переваги порівняно з традиційними архітектурами. Запропонований підхід, який полягає в інтеграції умовної гіпермоделі з основною нейронною мережею, та дозволяє ефективно адаптуватися до змін умов експлуатації системи, таких як варіації в освітленні, ракурсах зйомки та параметрах камер спостереження.

Експериментальне дослідження проведено за допомогою мови програмування Python та бібліо-

теки глибокого навчання PyTorch. Експерименти на наборі Market-1501 показали, що інтеграція адаптивної гіпермоделі суттєво підвищує продуктивність порівняно з базовою моделлю OSNet без гіпермоделі. Середня точність (mAP) зросла з 73,5 % до діапазону 75,4...78,3 %, а показник Rank-1 — з 90 % до близько 96...97 %, що свідчить про ефективніше вилучення релевантних ознак та адаптивне коригування представлень, що дозволяє точніше розрізняти ідентичності навіть в умовах високої варіативності зображень. Подібну закономірність можна зазначити і на наборі DukeMTMC-reID, де базова модель OSNet демонструє mAP 62,81 % і Rank-1 81,7 %, тоді як моделі з гіпермоделлю досягають mAP у межах 65,1...67,6 % та Rank-1 — від 83,5 % до 84,5 %. Отримані результати підтверджують, що використання адаптивних гіпермоделей дозволяє підвищити здатність системи до розпізнавання осіб на складніших та різноманітніших даних, що відкриває перспективи для їхнього подальшого впровадження в реальні системи відеоспостереження та безпеки.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] О. М. Кириленко, «Розробка методу повторної ідентифікації людини», *Опт-ел. інф-енерг. техн.*, вип. 41, т. 1, с. 25-32, 2022.
- [2] D. Ha, A. Dai, and Q. V. Le, "Hyper Networks," *arXiv preprint*. 2016. [Electronic resource]. Available: <https://arxiv.org/abs/1609.09106>.
- [3] X. Jia, X. Wei, and X. Cao, "Dynamic Filter Networks," *arXiv preprint*. 2016. [Electronic resource]. Available: <https://arxiv.org/abs/1605.09673>.
- [4] *Market-1501 dataset*. [Electronic resource]. Available: <https://www.pkuvmc.com/dataset.html>.
- [5] *DukeMTMC-ReID dataset* [Electronic resource]. Available: https://github.com/layumi/DukeMTMC-reID_evaluation.
- [6] *MSMT17 dataset*. [Electronic resource]. Available: <https://arxiv.org/abs/1711.08565>.
- [7] K. Zhou, Y. Yang, A. Cavallaro, and T. Xiang, "Learning Generalisable Omni-Scale Representations for Person Re-Identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, т. 44, no. 10, pp. 7593-7609, 2021.
- [8] A. Vaswani, et al., "Attention is All You Need," *Advances in Neural Information Processing Systems*, pp. 5998-6008, 2017.
- [9] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Cambridge. MIT press, 2016, 775 с.
- [10] A. Hermans, L. Beyer, and B. Leibe, "In Defense of the Triplet Loss for Person Re-Identification," *arXiv preprint*, 2017. [Electronic resource]. Available: <https://arxiv.org/abs/1703.07737>.
- [11] E. Ristani, F. Solera, R. Zou, R. Cucchiara, and C. Tomasi, "Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking," *European Conference on Computer Vision (ECCV)*, pp. 17-35, 2016.
- [12] K. Zhou, and T. Xiang, *Torchreid: A Library for Deep Learning Person Re-Identification in Pytorch*.
- [13] L. Zheng, Y. Yang, and A. G. Hauptmann, "Person re-identification: Past, present and future," *arXiv preprint*, 2016. [Electronic resource]. Available: <https://arxiv.org/abs/1610.02984>.
- [14] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," *IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248-255.
- [15] K. He, X. Zhang, S. Ren, and J. Sun Deep, "Residual Learning for Image Recognition," *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770-778.
- [16] D. P. Kingma, and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv preprint*, 2014. [Electronic resource]. Available: <https://arxiv.org/abs/1412.6980>.
- [17] I. Loshchilov, and F. Hutter, "SGDR: Stochastic Gradient Descent with Warm Restarts," *arXiv preprint*, 2016. [Electronic resource]. Available: <https://arxiv.org/abs/1608.03983>.

Рекомендована кафедрою автоматизації та інтелектуальних інформаційних технологій ВНТУ

Стаття надійшла до редакції 8.04.2025

Кириленко Олександр Михайлович — асистент кафедри комп'ютерних систем управління, e-mail: sasha.kyrylenko@gmail.com ;

Квстний Роман Наумович — д-р техн. наук, професор кафедри автоматизації та інтелектуальних інформаційних технологій;

Гармаш Володимир Володимирович — канд. техн. наук, доцент, доцент кафедри автоматизації та інтелектуальних інформаційних технологій;

Коцюбинський Володимир Юрійович — канд. техн. наук, доцент, доцент кафедри автоматизації та інтелектуальних інформаційних технологій.

Вінницький національний технічний університет, Вінниця

O. M. Kyrylenko¹
 R. N. Kvyetnyi¹
 V. V. Garmash¹
 V. Yu. Kotsiubynskyi¹

Adaptive Hypermodels Usage in Person Re-Identification

¹Vinnitsia National Technical University

This article deals with existing development and implementation approaches to the adaptive hypermodels usage for person re-identification. The research is based on a comparative study of the models performance applied to different data sets, covering both laboratory scenarios and real operating conditions. Main attention is paid to the assessment of key quality indicators, in particular, recognition accuracy (mAP, CMC) and information processing speed, which allows comprehensive coverage of the effectiveness of the methods used.

The study focuses on analyzing the impact of dynamic parameter changes on the results of model work, as well as on the study of incremental learning strategies that helps reduce the risk of catastrophic (knowledge loss) forgetting when adapting to new conditions without the need for complete re-training. Thanks to this approach, the system is able to quickly respond to changes in shooting conditions, for example, variations in lighting, angles and other characteristics, which is critically important for the video surveillance systems.

Based on the analysis done, promising areas of further research are outlined, aimed at improving adaptive learning algorithms, developing new architectural solutions and optimizing scaling processes. This, in its turn, will contribute to the implementation of more reliable and effective re-identification technologies in modern information systems.

The proposed approach combines usage of a hypermodel with an updated deep neural network, the key advantage of which is its high adaptability and stability of learning, ensured by the usage of a dynamic parameter adjustment, using hypermodels. Combination of cross-entropy and triplet losses allows us to effectively form compact and separate features for different identities, as well as increase the model ability to identify an object even in cases of significant variability of input data.

The results of the study demonstrate the prospects for integrating adaptive mechanisms into modern re-identification systems, providing increased resistance to changes in operating conditions and a high level of productivity, which is a necessary condition for successful practical application in information technologies.

Keywords: re-identification, hypernetworks, dynamic adaptation, incremental learning, CNN, Transformer, OSNet, Market-1501, DukeMTMC-ReID, MSMT17.

Kyrylenko Olexandr M. — Assistant Professor of the Chair of Computer Control Systems, e-mail: sasha.kyrylenko@gmail.com ;

Kvyetnyi Roman N. — Dr. Sc. (Eng.), Professor of the Chair of Automation and Intellectual Information Technologies;

Garmash Volodymyr V. — Cand. Sc. (Eng.), Associate Professor, Associate Professor of the Chair of Automation and Intelligent Information Technologies;

Kotsiubynskyi Volodymyr Yu. — Cand. Sc. (Eng.), Associate Professor, Associate Professor of the Chair of Automation and Intelligent Information Technologies