

Долгополов Сергій Юрійович

Аспірант кафедри інформаційних технологій,

<https://orcid.org/0000-0001-9418-0943>

Київський національний університет будівництва і архітектури, Київ

Крук Павло Михайлович

Аспірант кафедри інформаційних технологій,

<https://orcid.org/0000-0002-6786-452X>

Київський національний університет будівництва і архітектури, Київ

Жук Роман Володимирович

Аспірант кафедри інформаційних технологій,

<https://orcid.org/0009-0007-1657-6281>

Київський національний університет будівництва і архітектури

Гончаренко Тетяна Андріївна

Докторка технічних наук, професорка, завідувачка кафедри інформаційних технологій,

<https://orcid.org/0000-0003-2577-6916>

Київський національний університет будівництва і архітектури, Київ

МЕТОДИ ПОПЕРЕДНЬОЇ ОБРОБКИ ЗОБРАЖЕНЬ ДЛЯ АВТОМАТИЗОВАНИХ СИСТЕМ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ

Анотація. У дослідженні пропонується метод попередньої обробки зображень для застосування в автоматизованих системах розпізнавання об'єктів, який дає змогу підвищити точність, надійність та швидкість розпізнавання за різних умов освітлення та інших зовнішніх впливів. Запропонований метод базується на адаптивному вирівнюванні гістограми, обмеженому контрастом, гамма-корекції та CLAHE (contrast limited adaptive histogram equalization). У процесі покращення освітлення метод використовує нову адаптивну точку відсікання для вирівнювання гістограми на основі інформації про текстуру блоків зображення. Для вирішення проблем, таких як артефакти ореолів, недостатнє та надмірне підсилення, запропоновано впровадити в структуру функцію першої гамма-корекції, яка дає змогу покращити загальну варіацію відтінків сірого на зображенні. Запропоновано ввести другу функцію гамма-корекції, щоб запобігти негативному впливу різного освітлення на темних і світлих ділянках. Для розроблення нового методу попередньої обробки зображень на основі згорткової нейронної мережі запропонована унікальна компоновка шарів, яка містить декілька шарів згортки, активації, об'єднання та нормалізації. Експериментальні випробування довели ефективність розробленого методу, яка була оцінена шляхом визначення коефіцієнта точності (93,74%) у порівнянні з іншими методами попередньої обробки зображень в автоматизованій системі розпізнавання обличчя на основі локальних бінарних шаблонів (англ. Local Binary Patterns, LBP) та методу опорних векторів (англ. support vector machine, SVM). Апробація результатів підтвердила, що розроблений метод може удосконалити існуючі системи розпізнавання об'єктів або стати основою для створення нових автоматизованих програмних додатків, що вимагають високої надійності у розпізнавання зображень.

Ключові слова: згорткові нейронні мережі; CNN; методи попередньої обробки зображень; CLAHE; LBP; SVM; автоматизовані системи розпізнавання об'єктів

Актуальність та аналіз проблеми

Розпізнавання зображень є одним із найважливіших завдань у галузі комп'ютерного зору та біометричних технологій. Воно знаходить широке застосування в різних сферах, таких як безпека,

контроль доступу, мобільні додатки, соціальні мережі та багато інших. За останні роки значний прогрес у цьому напрямі було досягнуто завдяки розвитку глибокого навчання та згорткових нейронних мереж (CNN), що дозволило значно підвищити точність і надійність систем розпізнавання зображень.

Конволюційні нейронні мережі (англ. convolutional neural network, CNN) займають провідне місце у вирішенні завдань розпізнавання облич завдяки своїй здатності ефективно обробляти й аналізувати великі обсяги даних із зображень. Дослідження українського колективу дослідників О. В. Зінченко, О. С. Звенігородського та Т. М. Кисіля демонструє, що використання CNN дозволяє значно підвищити точність класифікації зображень, зокрема при розпізнаванні облич, завдяки здатності алгоритмів розрізняти важливі ознаки в структурі облич [1]. Водночас дослідники J. Zhang, Z. Su, L. Liu розробили мережу Median Pixel Difference Convolutional Network (MeDiNet), яка є релевантним прикладом, що демонструє покращену стійкість до шумів на зображеннях і здатна точно розпізнавати обличчя в реальному світі [2]. Відповідно до дослідження N. Lakshmi та M. P. Arakeri, оптимізована нейронна мережа може ефективно вирішувати «проблеми освітлення», що є ключовим аспектом при розпізнаванні облич [3]. В той же час колектив учених у складі I. Yildirim, M. Belledonne, W. Freiwald, J. Tenenbaum розробили модель, яка демонструє можливості нейромереж у біологічній обробці образів облич [4].

Більшість алгоритмів розпізнавання зображень у системах розпізнавання облич працюють за такими базовими етапами: попередня обробка, виявлення обличчя, вилучення ознак і класифікація. Ці етапи зазвичай виконуються послідовно, за рідкісними винятками. На етапі попередньої обробки алгоритм усуває будь-які шуми, присутні на зображенні. Крім того, попередня обробка може бути використана для виправлення проблем з освітленням на зображенні та нормалізації його відповідно до вимог алгоритму. Таким чином, логічно включати попередню обробку на початкових етапах процесу. У контексті попередньої обробки зображень у системах розпізнавання облич особливу увагу заслуговує дослідження S. Sarkar, K. B. Ajitha Shenoy, яке описує гібридний підхід до використання штучних нейронних мереж для розпізнавання облич. Автори розробили систему, що комбінує дві нейронні мережі для покращення точності розпізнавання: одну для вилучення ознак із використанням головних компонент (PCA-ANN); другу для лінійного дискримінантного аналізу (LDA-ANN). Ці мережі «співпрацюють» в ефективній нормалізації та корекції освітлення зображень задля виявлення облич у різних умовах освітлення [5]. Gamma intensity correction (GIC) є ще однією популярною технікою для корекції проблем з освітленням у системах розпізнавання облич. Дослідження L. Zhuowen, W. Kejun, Z. Guofeng, Y. Lei запропонувало вдосконалений метод GIC, який

використовується для покращення якості зображень облич за нерівномірного освітлення [6].

Метод Difference of Gaussians (DoG) ефективно працює з тінями, присутніми на зображенні, тому його можна розглядати як фільтр, що застосовується до вхідного зображення для кращого розуміння цього методу [7]. CLAHE (Contrast Limited Adaptive Histogram Equalization) – адаптивне гістограмне вирівнювання з обмеженим контрастом, є вдосконаленою версією гістограмного вирівнювання. На відміну від глобальної функції HE, CLAHE працює з блоками зображення та обмежується точкою відсікання. Цей підхід має на меті не лише підвищити контрастність, а й зберегти природність зображень. Завдяки автоматичному встановленню точки відсікання на основі текстури блоку та подвійній гамма-корекції цей метод ефективно підвищує яскравість у темних ділянках, одночасно контролюючи артефакти надмірного підсилення [9]. Методи підпростору на основі зовнішності, такі як аналіз головних компонент (PCA) [9], лінійний дискримінантний аналіз (LDA) [10] та сингулярна декомпозиція (SVD) [11], спрямовані на зменшення розмірності. Підходи до цілісного виділення ознак другої категорії використовують частотні перетворення, такі як дискретне косинусне перетворення (DCT) [12] або вейвлети Габора [13], щоб захопити унікальні особливості зображення. Хоча методи підвищення контрастності дають прийнятні результати в багатьох випадках, вони все ще стикаються з недоліками, коли вхідне зображення має сильні темні ділянки й глибокі тіні.

Розпізнавання облич стало одним з найпопулярніших методів у біометрії завдяки таким його властивостям, як ненав'язливість, відсутність необхідності участі людини та технологічність. Хоча нещодавні вдосконалення технології дозволили досягти прийнятного рівня точності, все ще залишається місце для вдосконалення, особливо при вирішенні проблем, спричинених неконтрольованими ситуаціями. Однією з таких проблем, яка може вплинути на точність системи, є зміна освітленості на зображенні. Зазвичай це спричинено такими факторами, як погані сенсори камери, неправильні джерела світла, цикл затвора та погане навколишнє освітлення. Відкинуті тіні, переекспоновані ділянки, занадто темні або занадто яскраві зображення можуть знизити швидкість розпізнавання системи, погіршуючи інформацію та характеристики зображення. Один із способів подолати ці варіації та спотворення – відкоригувати вхідне зображення перед тим, як подавати його в систему. Цей крок називається попередньою обробкою, і його можна знайти майже в кожній програмі для обробки зображень як одну з найважливіших частин.

Мета публікації

Метою цієї роботи є дослідження існуючих методів попередньої обробки зображень для розроблення нового методу для автоматизованих систем розпізнавання об'єктів на основі згорткових нейронних мереж, який надасть змогу підвищити точність, надійність та швидкість розпізнавання за різних умов освітлення та інших зовнішніх впливів.

Виклад основного матеріалу

За результатами опрацювання існуючих наукових робіт [7–16] методи попередньої обробки зображень доцільно поділити на три основні групи:

1. Методи на основі нелінійної передавальної функції.
2. Методи корекції гістограм.
3. Методи в частотній області.

Нелінійні методи здебільшого модифікують окремі значення пікселів зображення. Основними характеристиками цих методів є простота використання та налаштування, що робить їх ідеальними для корекції контрасту. Широко використовуваним прикладом нелінійних методів є гамма-корекція інтенсивності. Гамма-корекція працює з дискретними значеннями зображення для покращення освітлення на темних зображеннях.

Методи гістограмної корекції, з іншого боку, прагнуть виправити контрастність зображення шляхом уніфікації рівнів сірого по всьому зображенню. Чудовою властивістю методів гістограмногo вирівнювання є те, що вони вимагають відносно низьких обчислювальних витрат. Гістограму зображення можна загалом визначити як співвідношення рівнів сірого та частот. Наступне рівняння демонструє гістограму зображення у відтінках сірого:

$$H(i) = \frac{P_i}{N},$$

$$i = 0, 1, \dots, L - 1, \quad (1)$$

де i визначає рівень сірого зображення; P_i вказує на кількість пікселів цього рівня; N – загальна кількість пікселів на зображенні.

Вирівнювання гістограми для зображення, засноване на визначенні $H(i)$ має вигляд:

$$HE_m = M(j) = (L - 1) \sum_{i=0}^j H(i), \quad (2)$$

де $M(j)$ визначає поняття функції відображення; L – динамічний діапазон зображення.

Метод гістограмногo вирівнювання стикається зі своїм основним недоліком на зображеннях із помітними та відносно великими однорідними темними областями, де має тенденцію показувати надмірне посилення на виході. Рис. 1 демонструє цю поведінку як приклад.

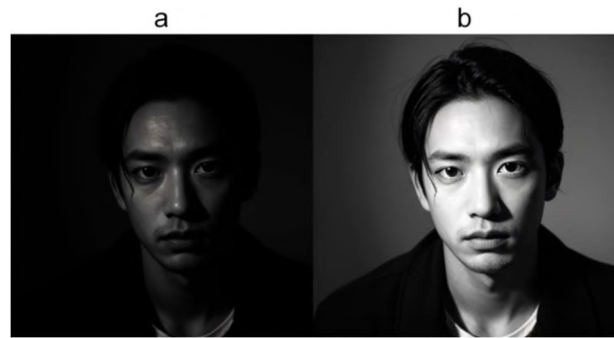


Рисунок 1 – Використання методу гістограмногo вирівнювання: а – оригінальне зображення; б – зображення після застосування методу

Таким чином, на світлих ділянках спостерігається надмірне підсилення, що призводить до втрати текстури. Крім того, недостатнє підсилення видно на темних ділянках, особливо якщо вони прилягають до добре освітлених частин обличчя. Це особливо негативно впливає на швидкість розпізнавання системи розпізнавання обличчя, усуваючи частину цінної інформації, прихованої в обрізаних текстурах зображення обличчя.

Було докладено певних зусиль для подолання цих проблем. Деякі з цих удосконалень вводять різні розподіли гістограми для зображення, такі як вирівнювання гістограми зі збереженням яскравості (англ. bi-histogram equalization, BBHE), вирівнювання гістограми дуалістичного підзображення рівної площі (англ. sub-image histogram equalization, DSIHE), вирівнювання гістограми з мінімальною середньою похибкою яскравості (англ. minimum mean brightness error bi-histogram equalization, MMBEBHE) та модель гауссової суміші (англ. Gaussian mixture model, GMM). На противагу цьому, деякі інші методи намагаються модифікувати інші аспекти традиційного методу, такі як зміна типу досліджуваної інформації на просторовий, використання подвійної ймовірності тощо. Прикладами таких методів є двостороння крива Безьє (англ. bilateral Bezier curve, BBC) та контекстне посилення контрасту.

Інші методи покращення традиційного HE пов'язані з сегментацією вхідного зображення. Найпомітнішим недоліком HE є погіршення характеристик зображення з відносно великими темними або світлими ділянками. Поділ зображення на різні сегменти та області допомагає запобігти цій проблемі, розбиваючи ці області. Крім того, робота з контрастом, пов'язаним з конкретною ділянкою, дає алгоритму кращий контроль над зображенням порівняно з однорідним контрастом на вході. Деякі приклади цих методів – експонетричне вирівнювання гістограми підзображення (ESIHE), вирівнювання гістограми підзображення на основі

середнього значення (MMSICHE), селективне динамічне вирівнювання гистограми (SSDHE) та сегментно-залежне динамічне вирівнювання мультигістограми (SDDMHE). Основна відмінність між цими методами полягає в тому, як вони вирішують проблему розподілу та корекції посилення контрасту.

Іншим важливим методом підвищення контрастності є адаптивне гістограмне вирівнювання з обмеженою контрастністю (англ. contrast limited adaptive histogram equalization, CLAHE), яке ефективно вирішує проблему надмірного посилення при гістограмному вирівнюванні. Основна ідея методу CLAHE полягає в тому, щоб розбити зображення на рівні підсекції і виконати гістограмне вирівнювання на кожній з цих секцій. У багатьох дослідженнях [7], метод CLAHE застосовується для покращення контрастності зображення, зберігаючи при цьому інформацію та уникаючи надмірного підсилення. Однак, у випадках, коли темні ділянки містять більшу частину зображення, CLAHE не здатний суттєво покращити загальну яскравість зображення. Цей недолік, однак, можна усунути, якщо застосувати нелінійну техніку, таку як гамма-корекція, на виході CLAHE. Поєднання цих двох методів дозволяє підвищити загальну освітленість зображення, зберігаючи при цьому текстурну інформацію в темних і світлих ділянках. Такі методи, як адаптивна гамма-корекція з розподілом ваг (AGCWD), адаптивна гамма-корекція (AGC) та комбінація AGC і вирівнювання бігістограми з обмеженим діапазоном (RLBBHE), використовують принцип поєднання двох методів. Проте на зображеннях з різною освітленістю ефективність

методів адаптивної гамма-корекції викликає сумніви. Y. Chang та ін. [9] запропонували комбінацію CLAHE з подвійною гамма-корекцією, але оскільки запропонована методика призначена для покращення зображення для візуальних цілей, тонкість загальної функції не дозволяє їй бути ідеальною для покращення освітленості зображення при розпізнаванні облич.

Основна відмінність між адаптивним вирівнюванням гистограми з обмеженням контрастності та звичайним вирівнюванням гистограми полягає в методі, який CLAHE використовує для обмеження точки відсікання. На рис. 2 зображено загальні етапи CLAHE.

Як видно з рис. 2, CLAHE складається з п'яти основних етапів. На першому етапі метод ділить зображення на рівні блоки. Потім виконується покращення гистограми для кожного блоку окремо. Покращення гистограми складається з трьох етапів: створення гистограми, обрізання гистограми та перерозподіл обрізаної частини. Останнім етапом у CLAHE є білінійна інтерполяція, яка допомагає зменшити артефакти між блоками і згладити вихідні дані.

Точку відсікання пікселів у методі CLAHE пропонується визначати такою формулою:

$$\beta = \frac{N}{D} = (1 + \frac{\alpha}{100} G_m), \quad (3)$$

де N – кількість пікселів у блоці; D – динамічний діапазон блоку; α – коефіцієнт відсікання; G_m – максимальний градієнт.

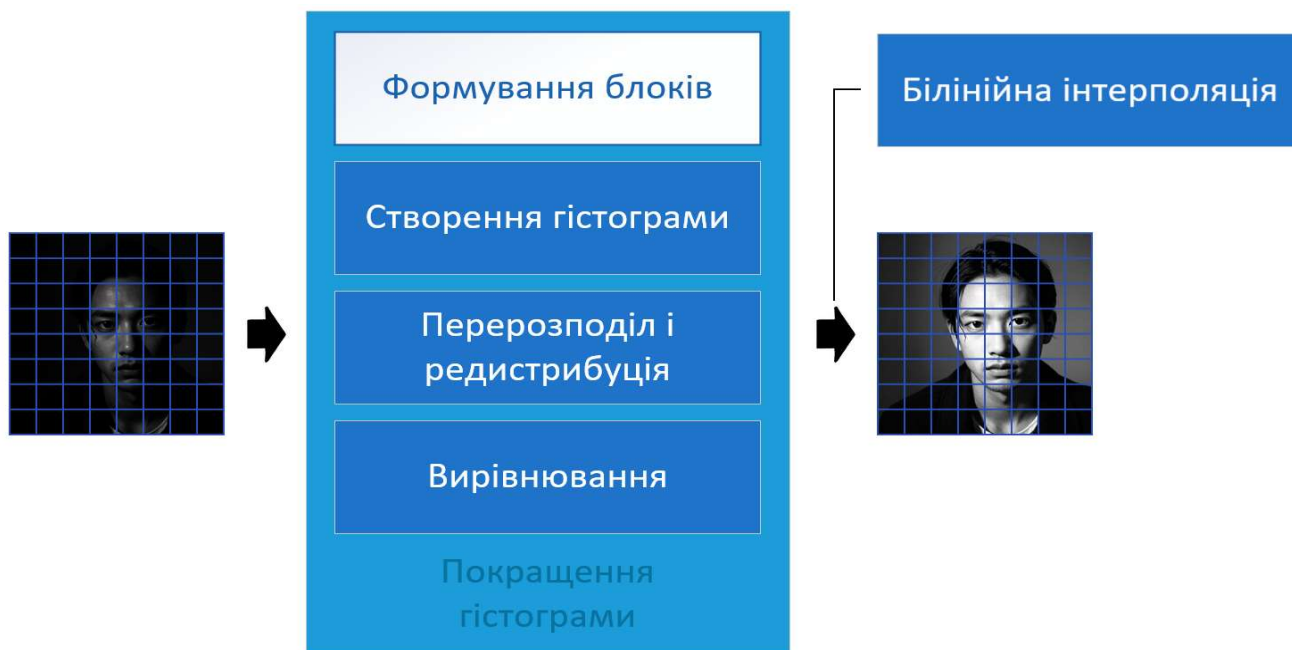


Рисунок 2 – Етапи CLAHE

Таким чином, коефіцієнт α визначає рівень підвищення контрастності. Чим більше α (ближче до 100), тим вищим буде рівень посилення контрасту. Те саме стосується CDF і функції ремапування:

$$CDF(g) = \sum_{l=0}^k pdf(g), \quad (4)$$

$$R(g) = CDF(g) \cdot g_{max}, \quad (5)$$

де $R(g)$ визначає функцію, яка повертає відрізаних частину до кривої гистограми; g_{max} – максимальне значення сірого кольору для пікселів у блоці.

Перевагою CLAHE є те, що він має відносно менші обчислювальні витрати завдяки тому, що впливає на невеликі підблоки зображення, а не на все зображення в цілому. На рис. 3 показано: а) точку фіксації та б) концепцію білінійної інтерполяції.

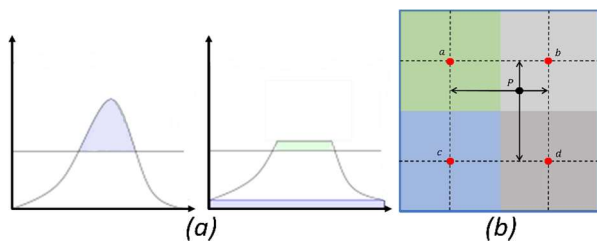


Рисунок 3 – Білінійна інтерполяція: а – точка відсікання та перерозподіл у CLAHE; б – концепція

Хоча методи підвищення контрастності дають прийнятні результати в багатьох випадках, вони все ще стикаються з недоліками, коли вхідне зображення має сильні темні ділянки й глибокі тіні. Це головним чином пов'язано з тим, що β або точка відсікання є фіксованою величиною, і тому вона не може змінюватися залежно від характеру поганого освітлення. Іншими словами, фіксована глобальна точка відсікання може давати задовільні результати для деяких частин бази даних, тоді як можуть бути випадки, що страждають від надмірного посилення контрасту, а також випадки, які викликають артефакти. Крім того, покращення контрастності зображення не є достатнім для того, щоб підвищити освітленість до достатнього рівня. Це особливо помітно у випадках з дуже темними неоднорідними ділянками.

Таким чином, такі моменти обґрунтовують визначення адаптивної точки відсікання та гамма-корекції як запропонованого методу. Запропонована методика використовує адаптивну точку відсікання на основі параметрів кожного блоку. Алгоритм також встановлює гамма-корекцію для g_{max} для покращення динамічного діапазону блоку.

У звичайному CLAHE з фіксованою точкою α і максимальний градієнт (Gm) є параметрами, що визначають точку відсікання. Щоб покращити продуктивність CLAHE, алгоритм розташовує точку кліпу таким чином, щоб мати нижчі значення для

згладжених і гармонізованих ділянок і вищі значення для блоків, що містять більшу варіацію освітленості.

Двома цінними параметрами, які розкривають корисну інформацію про блок, є медіанний рівень сірого та стандартне відхилення. Використання цих параметрів для встановлення адаптивної точки кліпу представлено у вигляді наступного рівняння:

$$\delta = \frac{g_{mod}}{D_b}, \quad (6)$$

$$\beta = \frac{N}{D_b} \left(1 + Q \frac{\delta}{D} + \frac{a}{100} \left(\frac{\sigma}{Avg+c} \right) \right), \quad (7)$$

де g_{mod} – найбільш повторюване сіре значення блоку; δ – динамічний фактор; σ – стандартне відхилення; c – мале значення.

Q та a контролюють динамічний діапазон та ентропійну вагу відповідно.

Нова точка кліпування має більше значення для блоків з більшою кількістю текстурної інформації, оскільки ці блоки мають більший коефіцієнт σ/Ave . На рис. 4 показано приклад нової точки відсікання для різних блоків зображення.

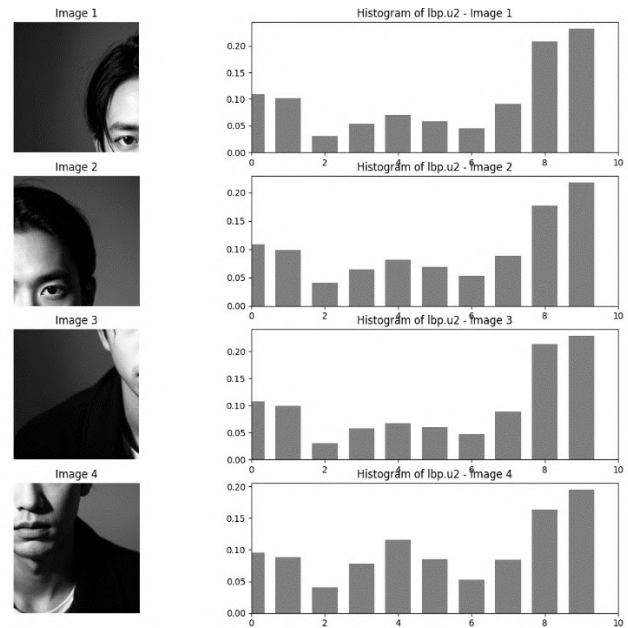


Рисунок 4 – Точка кліпу, керована вмістом

Хоча метод покращує контрастність на основі глобальної та локальної інформації та властивостей зображення, саме лише вирівнювання гистограми не може забезпечити достатнього покращення зображення, особливо в однорідних темних ділянках зображення. Це проявляється на ділянках з екстремально відкинутими тінями або в ситуації, коли в якості джерела освітлення використовується прожектор, оскільки зображення обличчя містить тіні частин обличчя, що блокують світло (в районі носа або очей). Для вирішення цієї проблеми ми пропонуємо гамма-корекцію. Процес гамма-корекції

полягає в технічному збільшенні варіації рівня сірого в затінених ділянках. Це покращує освітлення за рахунок збільшення деталей у темних областях. Функція відображення гамма-корекції визначається наступним чином:

$$R(g) = g_{max} \left(\frac{g}{g_{max}} \right)^\gamma, \quad (8)$$

Це співвідношення між g_{max} , $R(g)$ та α визначає поведінку гамма-корекції. Чим менше значення γ , тим інтенсивніше збільшується значення відповідного пікселя. Проте, через попередньо визначене значення гамма-корекції, ця зміна відбувається у всіх пікселях блоку зображення і може призвести до погіршення контрастності. Незважаючи на те, що значення гамма-корекції діє локально, погіршення контрастності нівелює ефект CLAHE.

Запропонований алгоритм впроваджує гамма-корекцію у введену CLAHE, щоб запобігти цьому. Для цього спочатку потрібно ввести зважену корекцію рівня сірого на основі значення γ .

$$w = \left(\frac{GL_{mod}}{GL_{med}} \right)^{1-\gamma_1}, \quad (9)$$

де GL_{mod} – найбільш повторюваний рівень сірого; GL_{med} – середній рівень сірого зображення.

Проведене дослідження існуючих методів попередньої обробки зображень дало змогу запропонувати власний метод на основі моделі CNN, який складається з ретельно розроблених шарів (рис. 5):

- згорткові шари (Conv1-Conv5) з різними розмірами та кроками ділянок;

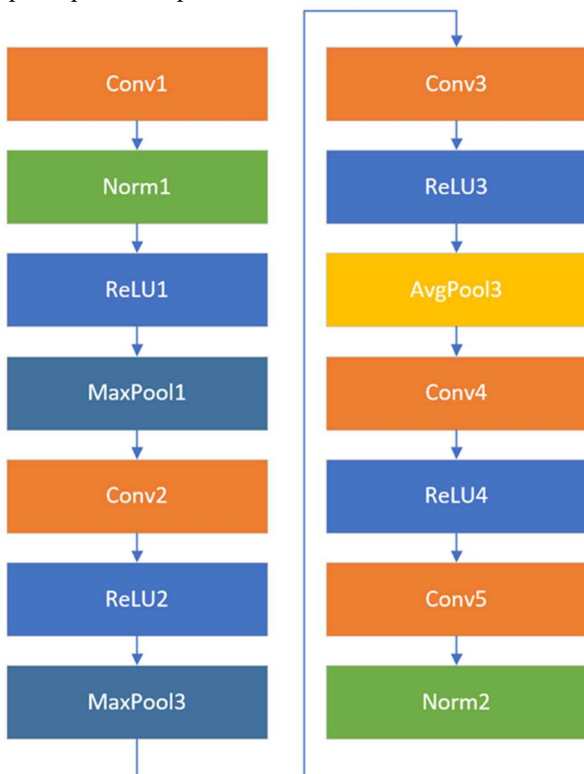


Рисунок 5 – Структура шарів методу попередньої обробки зображень на основі моделі CNN

- шари нормалізації (Norm1-Norm2) для стабілізації навчання;
- функції активації ReLU (ReLU1-ReLU4) для введення нелінійності;
- об'єднання шарів (MaxPool1, MaxPool3, AvgPool3) для зменшення розмірності.

У табл. 1 детально описано конкретні параметри для кожного шару, оптимізація яких здійснена шляхом ретельних експериментів для балансування можливостей вилучення ознак з обчислювальною ефективністю.

Таблиця 1 – Параметри шарів моделі CNN

Тип	Розмір патча	Крок
Conv1	5X5	2
Conv2	3X3	2
Conv3	3X3	1
Conv4	3X3	2
Conv5	3X3	1
Pool1	3X3	1
Pool2	3X3	1
Pool3	8X8	-

Таким чином, для розроблення нового методу попередньої обробки зображень на основі згорткової нейронної мережі використовується унікальна структура, яка містить декілька шарів згортки, активації, об'єднання та нормалізації.

Апробація результатів

Для експериментальної перевірки результатів дослідження використано розширений набір даних Єльського університету [17] для оцінки якості запропонованого методу шляхом порівняння точності різних методів розпізнавання зображень із запропонованим методом. Набір даних включає 21 888 фотографій, зроблених при моноосвітленні. У цьому тесті ми спробуємо оцінити ефективність запропонованого методу порівняно з іншими методами попередньої обробки. Для цього ми застосували різні методи попередньої обробки до однієї автоматизованої системи розпізнавання об'єктів, щоб мати можливість оцінити результати кожного методу належним чином. Ми обрали локальний бінарний шаблон як екстрактор ознак моделі системи розпізнавання обличчя, а також машину опорних векторів (SVM) як класифікатор цієї моделі. Знову ж таки, ми змінюємо розмір зображень до 64 на 64 перед тим, як подавати їх до екстрактора ознак. Ми також використовуємо ті самі 40 та 24 числа як навчальну та тестову множини для цього тесту.

Табл. 2 містить результати такого експерименту.

Таблиця 2 – Результати експерименту з різними методами попередньої обробки

Метод попередньої обробки	Коефіцієнт точності розпізнавання, %
None	81
HE	84
SQI	87
GIC	91
DoG	90
CLANE	91
Розроблений метод	93.74

Висновки

У дослідженні розроблено новий метод попередньої обробки зображень для застосування в автоматизованих системах розпізнавання об'єктів. Запропонований метод базується на адаптивному вирівнюванні гістограми, обмеженому контрастом, та гамма-корекції. У процесі покращення освітлення метод використовує нову адаптивну точку відсікання для вирівнювання гістограми на основі інформації про текстуру блоків зображення. Для вирішення

проблем, таких як артефакти ореолів, недостатнє та надмірне підсилення, запропоновано впровадити в структуру CLANE функцію першої гамма-корекції, яка дає змогу покращити загальну варіацію відтінків сірого на зображенні. Нарешті, оскільки деякі зразки можуть містити темні та світлі ділянки, запропоновано ввести другу функцію гамма-корекції, щоб запобігти негативному впливу цих двох типів освітлення на зображенні. Щоб оцінити успішність розробленого методу, ми провели експериментальні випробування. В якості еталонної моделі було обрано систему розпізнавання облич LBP+SVM. Ефективність запропонованого методу була оцінена шляхом порівняння коефіцієнта точності розпізнавання в системі з різними методами попередньої обробки зображень.

Наукова новизна дослідження полягає в розробленні методу розпізнавання зображень з використанням згорткових нейронних мереж, який значно покращує точність за різних умов освітлення та інших зовнішніх впливів. Практичне значення наукового результату полягає у застосуванні методу для автоматизованого розпізнавання об'єктів у системах різного функціонального призначення. Апробація результатів підтвердила, що такий метод може удосконалити існуючі системи розпізнавання об'єктів або стати основою для розроблення нових автоматизованих додатків, що вимагають високої надійності у розпізнаванні зображень.

Список літератури

1. Zinchenko O. V., Zvenihorodskyy O. S., Kysil T. M. Convolutional neural networks for solving computer vision problems. *Telecommunication and Information Technologies*. 2022. 2, 4–12. <https://doi.org/10.31673/2412-4338.2022.020411>.
2. Su J., Zhang Z., Liu L. Median pixel difference convolutional network for robust face recognition. *British Machine Vision Conference*. 2022. 1–14. <https://doi.org/10.48550/arXiv.2205.15867>.
3. Lakshmi N., Arakeri M. P. Face recognition under illumination based on optimized neural network. *International Journal of Advanced Computer Science and Applications (IJACSA)*. 2022. 13 (9), 131–137. <https://doi.org/10.14569/ijacsa.2022.0130915>.
4. Yildirim I., Belledonne M., Freiwald W. A., Tenenbaum J. B. Efficient inverse graphics in biological face processing. *Science Advances*. 2020. 6 (10), 1–18. <https://doi.org/10.1126/sciadv.aax5979>.
4. Sarkar S. D., Ajitha Shenoy K. B. Face recognition using artificial neural network and feature extraction. *2020 7th International Conference on SPIN*. 2020. 417–422. <https://doi.org/10.1109/SPIN48934.2020.9071378>.
5. Ke C., Ai-min D., Xiao-Hua L., Li-peng Z., Ling W., Xue-mei S. Difference of gaussians (DOG) model. *PLOS ONE*. 2015. 1. <https://doi.org/10.1371/JOURNAL.PONE.0144403.G001>.
6. Zhuowen L., Kejun W., Guofeng Z., Lei Y. Illumination compensation method for face image based on improved gamma correction. *Proceedings of the 32nd Chinese Control Conference*. 2013. 3733–3737.
8. Chang Y., Jung C., Ke P., Song H., Hwang J. Automatic contrast-limited adaptive histogram equalization with dual gamma correction. *IEEE Access*. 2018. 6, 11782–11792. <https://doi.org/10.1109/ACCESS.2018.2797872>.
7. O'Shea K., Nash R. An introduction to convolutional neural networks. *ArXiv*. 2015. abs/1511.08458, 1–11.
8. Гончаренко Т.А. Застосування BIM-технології для створення інформаційної моделі території під забудову. *Управління розвитком складних систем*. 2018. 33, 138–145.
9. Chernyshev D., Dolhopolov S., Honcharenko T., Haman H., Ivanova T., Zinchenko M. Integration of building information modeling and artificial intelligence systems to create a digital twin of the construction site. *2022 IEEE International Conference on Computer Sciences and Information Technologies (CSIT)*. 2022. 36–39. <https://doi.org/10.1109/CSIT56902.2022.10000717>.

10. Mihaylenko V., Honcharenko T., Chupryna K., Andrashko Yu. Modeling of Spatial Data on the Construction Site Based on Multidimensional Information Objects. *International Journal of Engineering and Advanced Technology (IJEAT)*. 2019. 8 (6), 3934–3940.
11. Terentyev O., Tsiutsiura S., Honcharenko T., Lyashchenko T. Multidimensional Space Structure for Adaptable. *International Journal of Recent Technology and Engineering*. 2019. 8 (3), 7753–7758.
12. Berezutskyi I., Honcharenko T., Ryzhakova G., Tykhonova O., Pokolenko V., Sachenko I. Methodological Approach for Choosing Type of IT Projects. Management. *2024 IEEE 4th International Conference on Smart Information Systems and Technologies*. 2024. 14–19. <https://doi.org/10.1109/SIST61555.2024.10629587>.
13. Ryzhakova G., Honcharenko T., Predun K., Petrukha N., Malykhina O., Khomenko O. Using of Fuzzy Logic for Risk Assessment of Construction Enterprise Management System. *2023 IEEE International Conference on Smart Information Systems and Technologies (SIST)*. 2023. 208–213. <https://doi.org/10.1109/SIST58284.2023.10223560>.
14. Гончаренко Т.А. Кластерний метод формування метаданих багатовимірних інформаційних систем для розв'язання задач генерального планування. *Управління розвитком складних систем*. 2020. 42, 93–101. <https://doi.org/10.32347/2412-9933.2020.42.93-101>.
15. Cambridge University. *The ORL face database*. n.d. <https://cam-orl.co.uk/facedatabase.html>.
16. Chernyshev D., Dolhopolov S., Honcharenko T., Haman H., Ivanova T., Zinchenko M. Integration of Building Information Modeling and Artificial Intelligence Systems to Create a Digital Twin of the Construction Site. *2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT)*. 2022. 36–39. <https://doi.org/10.1109/CSIT56902.2022.10000717>.

Стаття надійшла до редколегії 08.06.2025

Dolhopolov Serhii

Postgraduate of the Department of Information Technologies,
<https://orcid.org/0000-0001-9418-0943>

Kyiv National University of Construction and Architecture, Kyiv

Kruk Pavlo

Postgraduate of the Department of Information Technologies,
<https://orcid.org/0000-0002-6786-452X>

Kyiv National University of Construction and Architecture, Kyiv

Zhuk Roman

Postgraduate of the Department of Information Technologies,
<https://orcid.org/0009-0007-1657-6281>

Kyiv National University of Construction and Architecture, Kyiv

Honcharenko Tetyana

DSc (Eng.), Professor, Head of the Department of Information Technologies,
<https://orcid.org/0000-0003-2577-6916>

Kyiv National University of Construction and Architecture, Kyiv

METHODS OF IMAGE PRE-PROCESSING FOR AUTOMATED OBJECT RECOGNITION SYSTEMS

Abstract. This study developed a new method of image pre-processing for use in automated object recognition systems, allowing for increased accuracy, reliability, and speed of recognition under varying lighting conditions and other external influences. The proposed method is based on contrast-limited adaptive histogram equalization (CLAHE) and gamma correction. To enhance illumination, the method employs a novel adaptive cutoff point to equalize the histogram based on the texture information of image blocks. To address issues such as halo artifacts, under-gain, and over-gain, a first gamma correction function is proposed for integration into the CLAHE structure, which enhances the overall variation of gray shades in the image. A second gamma correction function is proposed to mitigate the negative impact of varying illumination on dark and light regions. For the development of a new image pre-processing method based on a convolutional neural network, a unique layer arrangement is proposed, featuring several layers of convolution, activation, pooling, and normalization. Experimental tests have demonstrated the effectiveness of the developed method, which achieved an accuracy coefficient of 93.74%. This was evaluated by comparing its performance with other image pre-processing methods within an automated face recognition system based on Local Binary Patterns (LBP) and the Support Vector Machine (SVM). The results confirmed that the developed method can improve existing object recognition systems or serve as a foundation for creating new automated software applications requiring high reliability in image recognition.

Keywords: convolutional neural networks; CNN; method of image pre-processing; CLAHE; LBP; SVM; automated object recognition systems

References

1. Zinchenko, O. V., Zvenihorodsky, O. S., & Kysil, T. M. (2022). Convolutional neural networks for solving computer vision problems. *Telecommunication and Information Technologies*, 2, 4–12. <https://doi.org/10.31673/2412-4338.2022.020411>.
2. Su, J., Zhang, Z., & Liu, L. (2022). Median pixel difference convolutional network for robust face recognition. *British Machine Vision Conference*, 1–14. <https://doi.org/10.48550/arXiv.2205.15867>.
3. Lakshmi, N., & Arakeri, M. P. (2022). Face recognition under illumination based on optimized neural network. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 13 (9), 131–137. <https://doi.org/10.14569/ijacsa.2022.0130915>.
4. Yildirim, I., Belledonne, M., Freiwald, W. A., & Tenenbaum, J. B. (2020). Efficient inverse graphics in biological face processing. *Science Advances*, 6 (10), 1–18. <https://doi.org/10.1126/sciadv.aax5979>.
5. Sarkar, S. D., & Ajitha Shenoy, K. B. (2020). Face recognition using artificial neural network and feature extraction. In *2020 7th International Conference on SPIN* (pp. 417–422). <https://doi.org/10.1109/SPIN48934.2020.9071378>.
6. Ke, C., Ai-min, D., Xiao-Hua, L., Li-peng, Z., Ling, W., & Xue-mei, S. (2015). Difference of gaussians (DOG) model. *PLOS ONE*, 1. <https://doi.org/10.1371/JOURNAL.PONE.0144403.G001>.
7. Zhuowen, L., Kejun, W., Guofeng, Z., & Lei, Y. (2013). Illumination compensation method for face image based on improved gamma correction. In *Proceedings of the 32nd Chinese Control Conference* (pp. 3733–3737).
8. Chang, Y., Jung, C., Ke, P., Song, H., & Hwang, J. (2018). Automatic contrast-limited adaptive histogram equalization with dual gamma correction. *IEEE Access*, 6, 11782–11792. <https://doi.org/10.1109/ACCESS.2018.2797872>.
9. O'Shea, K., & Nash, R. (2015). *An introduction to convolutional neural networks*. ArXiv. <https://arxiv.org/abs/1511.08458>.
10. Honcharenko, T. A. (2018). Application of BIM-technology for creating an information model of the territory for development. *Management of Complex Systems Development*, 33, 138–145.
11. Chernyshev, D., Dolhopolov, S., Honcharenko, T., Haman, H., Ivanova, T., & Zinchenko, M. (2022). Integration of building information modeling and artificial intelligence systems to create a digital twin of the construction site. In *2022 IEEE International Conference on Computer Sciences and Information Technologies (CSIT)* (pp. 36–39). <https://doi.org/10.1109/CSIT56902.2022.10000717>.
12. Mihaylenko, V., Honcharenko, T., Chupryna, K., & Andrashko, Y. (2019). Modeling of spatial data on the construction site based on multidimensional information objects. *International Journal of Engineering and Advanced Technology (IJEAT)*, 8 (6), 3934–3940.
13. Terentyev, O., Tsiutsiura, S., Honcharenko, T., & Lyashchenko, T. (2019). Multidimensional space structure for adaptable. *International Journal of Recent Technology and Engineering*, 8 (3), 7753–7758.
14. Berezutskyi, I., Honcharenko, T., Ryzhakova, G., Tykhonova, O., Pokolenko, V., & Sachenko, I. (2024). Methodological approach for choosing type of IT projects. Management. In *2024 IEEE 4th International Conference on Smart Information Systems and Technologies* (pp. 14–19). <https://doi.org/10.1109/SIST61555.2024.10629587>.
15. Ryzhakova, G., Honcharenko, T., Predun, K., Petrukha, N., Malykhina, O., & Khomenko, O. (2023). Using of fuzzy logic for risk assessment of construction enterprise management system. In *2023 IEEE International Conference on Smart Information Systems and Technologies (SIST)* (pp. 208–213). <https://doi.org/10.1109/SIST58284.2023.10223560>.
16. Honcharenko, T. A. (2020). Cluster method for forming metadata of multidimensional information systems for solving general planning problems. *Management of Complex Systems Development*, 42, 93–101. <https://doi.org/10.32347/2412-9933.2020.42.93-101>.
17. Cambridge University. (n.d.). *The ORL face database*. <https://cam-orl.co.uk/facedatabase.html>.
18. Chernyshev, D., Dolhopolov, S., Honcharenko, T., Haman, H., Ivanova, T., & Zinchenko, M. (2022). Integration of building information modeling and artificial intelligence systems to create a digital twin of the construction site. In *2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT)* (pp. 36–39). <https://doi.org/10.1109/CSIT56902.2022.10000717>.

Посилання на публікацію

- APA Dolhopolov S., Kruk P., Zhuk R., & Honcharenko T. (2025). Methods of image pre-processing for automated object recognition systems. *Management of Development of Complex Systems*, 62, 155–163. [dx.doi.org/10.32347/2412-9933.2025.62.155-163](https://doi.org/10.32347/2412-9933.2025.62.155-163).
- ДСТУ Долгополов С. Ю., Крук П. М., Жук Р. В. Гончаренко, Т. А. Моделі і методи штучного інтелекту в процесі виконання будівельно-технічної експертизи. *Управління розвитком складних систем*. Київ, 2025. № 62. С. 155 – 163. [dx.doi.org/10.32347/2412-9933.2025.62.155-163](https://doi.org/10.32347/2412-9933.2025.62.155-163).