

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
МІНІСТЕРСТВО ВИЩОЇ ОСВІТИ, НАУКИ ТА ІННОВАЦІЙ
РЕСПУБЛІКИ УЗБЕКІСТАН
МІНІСТЕРСТВО НАУКИ ТА ОСВІТИ АЗЕРБАЙДЖАНСЬКОЇ
РЕСПУБЛІКИ
НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
"ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ"**

**ПРОБЛЕМИ
ІНФОРМАТИКИ ТА МОДЕЛЮВАННЯ
(ПІМ-2025)**

**ТЕЗИ ДВАДЦЯТЬ П'ЯТОЇ МІЖНАРОДНОЇ
НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ
(25 – 28 вересня 2025 року)**

Харків

2025

УДК 004.9

Проблеми інформатики та моделювання (ПІМ-2025). Тези двадцять п'ятої міжнародної науково-технічної конференції. – Харків: НТУ "ХПІ", 2025. – 138 с.

ОРГАНІЗАТОРИ КОНФЕРЕНЦІЇ:

- Міністерство освіти і науки України;
- Національна Академія наук України;
- Національний технічний університет "ХПІ", Харків;
- Національний університет "Одеська політехніка", Одеса;
- Інститут проблем моделювання в енергетиці імені Г.Є. Пухова НАНУ, Київ;
- Харківський національний університет радіоелектроніки, Харків;
- Донбаська державна машинобудівна академія, Краматорськ;
- Ташкентський інститут інженерів іригації і механізації сільського господарства, Ташкент, Узбекистан;
- Азербайджанський державний університет нафти і промисловості, Баку, Азербайджан;
- Грузинський технічний університет, Тбілісі, Грузія.

Електронна адреса сайту конференції: <https://web.kpi.kharkov.ua/pim/>



О.А. Горносталь, В.В. Челақ, С.Ю. Гавриленко Використання метрик програмної складності для виявлення шкідливого коду засобами машинного навчання.....	36
В.О. Гороховатський, І.С. Творошенко Оцінювання значущості ознак для підвищення продуктивності структурних методів класифікації зображень	38
М. Hulevych An intelligent method for C++ test case synthesis based on a q-learning agent	44
В.В. Єсінов Розробка системи підтримки прийняття рішень на основі нейронних мереж для діагностики серцево-судинних захворювань	45
О.Ю. Заковоротний, О.С. Євтушенко Вплив коливань рухомого складу на проходження криволінійної ділянки шляху	46
О.Ю. Заковоротний, П.Е. Решетнікова Моделі, методи та програмні компоненти для комп'ютерної системи управління рухом поїзда	47
О.Ю. Заковоротний, Ю.В. Юрченко Дослідження можливості реалізації алгоритму за рівнянням беллмана в дискретній моделі за допомогою мови PL/SQL (ORACLE)	50
М.Й. Запововський, М.В. Мезенцев, М.В. Оліфір Розробка комп'ютерних компонентів для системи керування електроприводом змінного струму дизель-поїзда	52
Y.S. Zarichnyi WEB service recommendation based on hybrid search	53
А.О. Каргін, Р.С. Кузьменко Моделювання ймовірносно-нечіткого актуатора в моделі м'якого управління автономними безлюдними системами	54
О.В. Касілов, К.С. Дмитрук Модель динамічного балансування автономних мереж	55
В.Д. Кассов, Р.А. Бородай, М.К. Шаповалов, Д.М. Григоренко Управління температурно-силовими параметрами процесу механообробки з застосуванням мехатроніки	56
О.А. Кобилін, В.О. Гороховатський, А.П. Запорожченко Нейромережа Хемінга для класифікації зображень за множиною дескрипторів	57

НАУКОВЕ ВИДАННЯ

**ТЕЗИ ДВАДЦЯТЬ П'ЯТОЇ МІЖНАРОДНОЇ
НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ
"ПРОБЛЕМИ ІНФОРМАТИКИ ТА МОДЕЛЮВАННЯ"
(ПІМ-2025)**

*Відповідальний за випуск лауреат Національної премії України
імені Бориса Патона, д.т.н., проф. О.Ю. Заковоротний*

Науковий редактор д.т.н. Леонов С.Ю.
Технічний редактор к.т.н. Мезенцев М.В.

Підп. до друку 03.09.2025 р. Формат 60x84 1/16. Папір Сору Рарер.
Гарнітура Таймс. Умов. друк. арк. 7,5.
Облік. вид. арк. 8,2. Наклад 120 прим.
Ціна договірна

Видавничий центр НТУ «ХП»
Свідоцтво про державну реєстрацію ДК № 5478 від 21.08.2017 р.
61002, Харків, вул. Кирпичова, 2.

Виготовлено у ТОВ ВПП "Контраст".
Україна, 61166, м. Харків, пр. Науки, 40, оф. 221.
Св-во: ДК №1778 від 05.05.2004

ОЦІНЮВАННЯ ЗНАЧУЩОСТІ ОЗНАК ДЛЯ ПІДВИЩЕННЯ ПРОДУКТИВНОСТІ СТРУКТУРНИХ МЕТОДІВ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ

*д-р техн. наук, проф. В.О. Гороховатський, канд. техн. наук, доц.
І.С. Творошенко, ХНУРЕ, м. Харків*

У задачах класифікації зображень, що є зараз досить актуальними для комп'ютерного зору, систему ознак часто формують як множину багатовимірних векторів, що у повній мірі відображають просторові властивості візуального об'єкту для здійснення результативного аналізу [1, 2]. Наприклад, опис візуального об'єкта подається як скінченна множина дескрипторів ключових точок (КТ) зображення [3]. Дескриптор – це багатоконпонентний числовий вектор, що відображає характеристики деякої околиці КТ на зображенні і формується спеціальними фільтрами – детекторами [1].

Фактори, які характеризують важливість (вагомість, значущість, значимість) ознак: розкид значень всередині опису і в межах бази еталонних описів, ступінь стійкості до геометричних перетворень і дії завад, унікальність та цінність присутності на зображенні. Конкретний рівень параметру вагомості, як і результативність класифікації загалом, значною мірою визначається базою еталонних зображень, які треба класифікувати. Введення параметру значущості в процесі класифікаційного аналізу дає змогу перейти від однорідного впливу ознак до врахування їх відносної зваженості, що напряму впливає на характеристики класифікації: продуктивність та результативність [3].

Значущість апріорно можна оцінити для наявної множини компонентів еталонних описів, її використання націлено на більшу адаптацію класифікації до аналізованих даних внаслідок розширення обсягу інформації [2]. На ефективність впровадження параметрів значущості впливає також форма подання простору ознак і застосований спосіб класифікації. Однією із оцінок класифікаційної значущості є параметр інформативності ознаки. Інформативність ознаки є тим вищою, чим краще ця ознака розділяє екземпляри навчальної вибірки на класи. На підставі впровадження вагових коефіцієнтів та параметру інформативності, з'являється можливість здійснення більш глибокого аналізу даних, що покращує показники точності розпізнавання.

Розглянемо множину $E = \{E_i\}_{i=1}^N$ описів бази еталонів $E = \{\{e_v(i)\}_{v=1}^s\}_{i=1}^N$, де N – число класів, $i = 1, \dots, N$ – номер класу, s – потужність множини компонентів окремого опису, $v = 1, \dots, s$ – номер компонента всередині опису класу E_i .

Модель даних для бази класифікації E можна подати у формі матриці розміром $N \times s$, де кожний рядок містить опис окремого еталону (класу) у складі s компонентів – дескрипторів КТ [4].

Опис $Z = \{z_v\}_{v=1}^s$ розпізнаваного об'єкта подається моделлю, аналогічно опису окремого еталону E_i . У новітніх детекторах КТ використовують бінарні подання для дескрипторів КТ [1]. Будемо вважати $z_v \in B^n$, $e_v \in B^n$, $E_i \subset B^n$, $Z \subset B^n$, де B^n – простір векторів розмірності n з бінарними компонентами.

Формально розпізнаваний клас у такій постановці [1, 4] має вид нескінченної множини зображень, отриманої із зображення конкретного еталону застосуванням до нього всіляких наборів геометричних перетворень зсуву, масштабу, повороту.

Для кожного елемента $e_v(i) \in E$ бази еталонів на підготовчому етапі аналізу даних визначимо величину параметра значущості $\lambda_{i,v}$ у формі числа.

Для E отримаємо матрицю $\Lambda = \{\{\lambda_{i,v}\}_{v=1}^s\}_{i=1}^N$ чисел

$$\Lambda = \begin{Bmatrix} \lambda_{1,1}, \dots, \lambda_{1,s} \\ \lambda_{2,1}, \dots, \lambda_{2,s} \\ \dots \\ \lambda_{N,1}, \dots, \lambda_{N,s} \end{Bmatrix}. \quad (1)$$

Матриці Λ та E мають ідентичну структуру. Але компонентами E є бінарні вектори, а компоненти Λ – це числа, що характеризують значущість цих векторів для класифікації. У рядках матриці E містяться значення дескрипторів КТ для описів окремих класів, а Λ містить кількісний параметр значимості для кожного із дескрипторів.

Зазначимо, що параметр $\lambda_{i,v}$ матриці Λ , відображає вагомість кожної ознаки $e_v(i)$ всередині окремого класу E_i чи в рамках усієї наявної множини E класів. Способи [5] визначення елементів $\lambda_{i,v}$ вказують на їх пряму залежність від складу E , так як класифікаційна вагомість компонентів залежить від набору розпізнаваних класів.

Поставимо задачу розроблення процедури $R: E \rightarrow E^*$ для цілеспрямованого скорочення (редукції) складу опису бази E шляхом відбору найбільш значимих за критерієм матриці Λ елементів, причому $\text{card } E^* \ll \text{card } E$, $E^* = \{E_i^*\}_{i=1}^N$.

Потужність $s^* = \text{card } E_i^*$ опису кожного із еталонів E_i при цьому рівноцінно скорочується: $s \rightarrow s^*$, $s^* \ll s$. Формування новоствореної бази E^* на основі редукції R призначено для здійснення класифікації зі значно меншими обчислювальними затратами із забезпеченням необхідного рівня результативності.

Друга нагальна задача полягає у створенні класифікатора F , який для довільного опису $Z = \{z_v\}_{v=1}^s$ об'єкту у формі множини дескрипторів КТ прийме рішення $F[Z] \rightarrow [1, 2, \dots, N]$ про належність аналізованого опису Z до одного із N класів із врахуванням значень матриці Λ .

Класифікація $F[Z]$ здійснюється на основі таких засад: описів Z об'єкту та бази $\{E_i\}_{i=1}^N$, а також визначеної значущості у формі матриці Λ для множини еталонних компонентів. Повна формальна параметризація класифікатора – $F[Z, E, \Lambda] \rightarrow [1, 2, \dots, N]$. Класифікатор F може використовувати редуковану базу E^* замість E .

Така вимога потребує модифікації для моделі оцінювання релевантності описів Z та E_i^* з нерівнозначним числом елементів.

Застосуємо для обчислення матриці Λ метричний критерій інформативності [2]. Для довільного вектора $z \in E$ у системі класів як складового елемента $z \in E_k$ фіксованого еталонного опису E_k з номером k введемо поняття інформативності $I(z, E)$ у складі бази E

$$I(z, E) = \rho_m(z, \bar{E}_k) - \rho_m(z, E_k), \quad (2)$$

де $\rho_m(z, \bar{E}_k) = \min_{v, i \neq k} \rho(z, e_v(i))$ – мінімальна відстань від z до елемента бази, що не належить класу E_k , $\rho_m(z, E_k) = \min_{v, i=k} \rho(z, e_v(i))$ – відстань від z до найближчого елемента із класу E_k (за виключенням відстані $\rho(z, z) = 0$ самого до себе, $z \in E_k$).

При впровадженні нормованих відстаней із величиною $0 \leq \rho \leq 1$ оцінювані значення критерію (2) лежать в інтервалі $-1 \leq I \leq 1$.

У методі Relief використовується один із варіантів моделі (2), де інформативність оцінюється як різниця квадратів відстаней, а відбір аналізованих ознак здійснюється випадковим чином [6]. Як бачимо, критерій (2) має властивість оцінки даних.

Використання моделі (2) для визначення індивідуальної інформативності I для $z \in E_k$ засновано на припущенні про те, що

класифікаційна значущість ознаки є тим вищою, чим краще вона розділяє екземпляри навчальної вибірки на класи. Виходячи з цього, чим далі від екземпляра класу знаходиться найближчий до нього екземпляр іншого класу, тим вищою є індивідуальна інформативність. У той же час, чим далі від екземпляра знаходиться найближчий до нього елемент цього ж класу, тим нижчою є індивідуальна інформативність. Тобто впроваджується принцип: «ближче до своїх, подалі від інших».

Таким чином, ознаки (2) з високими значеннями I індивідуальної інформативності будемо вважати значущими, тобто суттєво інформативними стосовно результативної класифікації, ознаки з низьким I вважаються малозначущими, тобто є претендентами на виключення із системи ознак.

Критерій (2) є більш дієвим за інші [7], так як він уже частково відображає ступінь метричного розрізнення аналізованих компонентів даних, що є підставою для класифікації зображень за сформованими описами, та характеризується числом, що спрощує оброблення.

Результативність використання параметру інформативності (2) для розпізнаваних даних необхідно перевіряти експериментально, що дасть можливість оцінити прикладну точність класифікації з використанням інформативності.

З використанням програмних засобів бібліотеки OpenCV детектором BRISK сформовано опис із множини дескрипторів КТ для аналізованих зображень. У експерименті для трьох різних зображень певного казкового персонажу (рис. 1) з числом $s = 500$ дескрипторів BRISK у описах шляхом відбору 100 дескрипторів із найвищим показником інформативності (2) прогнозовано досягнуто приблизно п'ятикратний вигравш у часі класифікації.



Рис. 1. Зображення в експерименті редукції за критерієм (2)

Точність класифікації для повного і редукованого описів при цьому складає $pr = 1,0$, тобто усі зображення класифікуються вірно.

Важливим показником для методів голосування є відношення максимуму числа голосів $h_{\max 1}$ чи агрегованої значущості, за якими приймається класифікаційне рішення, до найближчого локального максимуму $h_{\max 2}$ класу – конкурента

$$\Delta = h_{\max 2} / h_{\max 1} . \quad (3)$$

Значення Δ характеризує ступінь впевненість рішення. Чим менше Δ , тим надійніше прийняте рішення про клас об'єкту.

Показник впевненості класифікації (3) для редукованого опису у $s = 100$ дескрипторів склав значення 0,29, у той час як для повного опису у 500 дескрипторів цей показник для аналізованих зображень (див. рис. 1) дорівнює 0,59, що значно гірше.

Отриманий експериментальний результат говорить про те, що навіть для достатньо схожих зображень за рахунок редукації за коефіцієнтом інформативності досягнуто підвищення надійності в аспекті розрізнення зображень. Редукація опису у п'ять разів за критерієм інформативності дає не тільки істотний вигравш у часі оброблення, але й забезпечує зростання показника впевненості при збереженні точності класифікаційного рішення. Відібрана підмножина найбільш інформативних дескрипторів забезпечує результативність класифікації з істотно нижчими обчислювальними затратами.

Задля узагальнення результатів дослідження проведено програмне моделювання класифікатора для тих же тестових даних із використанням засобу подвійної перевірки відповідності дескрипторів за моделлю Cross-Checking. Для 500 дескрипторів точність 1,0 не змінилась, але для її забезпечення треба знизити поріг δ_h для числа голосів до значення 220. Це необхідно для правильної класифікації другого еталону, найбільш схожого з іншими. Коефіцієнт впевненості Δ для зображень із бази покращився до рівня 0,08, а час t_c обчислень майже не зріс у порівнянні із підходом без подвійної перевірки.

Для складу із 50 дескрипторів досягнута максимальна точність 1,0 при порозі δ_h у 25 голосів, $\Delta \leq 0,1$ для бази, час класифікації приблизно у 10 разів менше, ніж для повного складу опису. Для 25 дескрипторів досягнута точність 1,0 при порозі δ_h у 15 голосів, $\Delta \leq 0,5$ для бази, час класифікації у 2 рази менше, ніж для складу опису у 50 дескрипторів.

Варіант для 25 дескрипторів із використанням накопичення значущості val_k підтверджує класифікацію із застосуванням лише δ_h .

Важливий результат отримано для невеликого числа дескрипторів у 10 елементів. При виборі порогу $\delta_h = 8$, як у першому методі, точність падає до рівня $pr = 0,84$. Але вже при виборі порогу $\delta_h = 6$ точність значно покращується до $pr = 0,96$ (але при суттєвому збільшенні $\Delta \leq 0,85$). Покращенням результату при невеликих описах є прийняття рішення із перевіркою за параметром δ_h . Класифікація за накопиченою значущістю val_k підтверджує висновок за підрахунком голосів. Важливим є вибір порогу δ_h , ніж використання показника Δ .

Основний результат дослідження полягає у встановленні високої продуктивності і результативності класифікаторів за редукованим складом опису еталонів. Застосування редукції для трансформації множини дескрипторів опису зображень дає можливість прискорити оброблення без втрати точності класифікації. Швидкодія оброблення збільшується пропорційно ступеню редукції даних і для проведеного експерименту покращена у 20 разів. Доцільним є використання 10% найбільш інформативних дескрипторів, що забезпечує приріст швидкодії в 10 разів при збереженні точності. Якщо ж швидкодія є ключовим критерієм, то допустимо використовувати навіть 5% від оригінальної кількості дескрипторів, що забезпечує приріст швидкодії майже у 20 разів, але при незначному зниженні точності до 0,95.

Результати дослідження отримані в рамках міжнародного дослідницького проекту ЄС "Горизонт Європа" – "INITIATE" (грант № 101136775).

Список літератури: 1. *Gorokhovatskyi V.* Application a committee of Kohonen neural networks to training of image classifier based on description of descriptors set / *V. Gorokhovatskyi, I. Tvoroshenko, O. Yakovleva, M. Hudáková, O. Gorokhovatskyi* // IEEE Access. – 2024. – Vol. 12. – P. 73376-73385. 2. *Gorokhovatskyi V.* Search for visual objects by request in the form of a cluster representation for the structural image description / *V. Gorokhovatskyi, I. Tvoroshenko, O. Kobylín, N. Vlasenko* // Advances in Electrical and Electronic Engineering. – 2023. – Vol. 21. – No. 1. – P. 19-27. 3. *Ahmad M.A.* The research of image classification methods based on the introducing cluster representation parameters for the structural description / *M.A. Ahmad, V. Gorokhovatskyi, I. Tvoroshenko, N. Vlasenko, S.K. Mustafa* // International Journal of Engineering Trends and Technology. – 2021. – Vol. 69. – No. 10. – P. 186-192. 4. *Gorokhovatskyi V.* Distance matrix for a set of structural description components as a tool for image classifier creating / *V. Gorokhovatskyi, O. Peredrii, I. Tvoroshenko, T. Markov* // Advanced Information Systems. – 2023. – Vol. 7. – No. 1. – P. 5-13. 5. *Oliinyk A.* The system of criteria for feature informativeness estimation in pattern recognition / *A. Oliinyk, S. Subbotin, V. Lovkin, O. Blagodariov, T. Zaiko* // Radio Electronics, Computer Science, Control. – 2017. – Vol. 43. – No. 4. – P. 85-96. 6. *Daradkeh Y.I.* Improving the effectiveness of image classification structural methods by compressing the description according to the information content criterion / *Y.I. Daradkeh, V. Gorokhovatskyi, I. Tvoroshenko, M. Zeghid* // Computers, Materials and Continua. – 2024. – Vol. 80. – No. 2. – P. 3085-3106. 7. *Gorokhovatskyi V.* Image description compression in classification structural methods / *V. Gorokhovatskyi, I. Tvoroshenko, O. Yakovleva, M. Hudáková* // IEEE Access. – 2025. – Vol. 13. – P. 43631-43641.