

Міністерство освіти та науки України
Національна академія наук України
Координаційна рада НАН України з питань штучного інтелекту
Харківський національний університет радіоелектроніки
Харківський національний університет імені В.Н.Каразіна
Карпатський національний університет імені Василя Стефаника
Північного-Східний координаційний науковий центр з питань штучного інтелекту
Інститут кібернетики імені В.М. Глушкова НАН України
Університет технологій в Лодзі
Університет Павла Йозефа Шафарика в Кошице

СУЧАСНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА СИСТЕМИ ШТУЧНОГО ІНТЕЛЕКТУ MIT&AIS-2026

**Матеріали
2-ї Міжнародної науково-практичної конференції**

Частина 2

**27-29 квітня 2026 р.
Харків - Яремче, Україна**

Харків 2026

ЗАСТОСУВАННЯ АСОЦІАТИВНИХ ПРАВИЛ У МЕТОДАХ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ

Запорожченко А.П.¹, Гороховатський В.О.¹

¹ Харківський національний університет радіоелектроніки, пр. Науки, 14, м. Харків, 61166, Україна

Анотація

У роботі розглянуто застосування апарату асоціативних правил у задачі класифікації зображень. Асоціативні правила синтезовано на даних кластерного подання для множини дескрипторів ключових точок зображення. Проаналізовано особливості формування простору структурних ознак для зображень, обговорюються характеристики та особливості побудови правил. Наведено приклад видобутку правил і результат їх застосування.

Ключові слова: класифікація зображень; дескриптор ключової точки; інтелектуальний аналіз даних; асоціативні правила; підтримка.

У задачах класифікації зображень широкого застосування набули методи розпізнавання, що базуються на представленні зображення як сукупності дескрипторів ключових точок (КТ) у формі багатовимірних векторів [1-3].

Існуючі підходи до редукції даних зазвичай орієнтовані або на відбір ознак, або на відбір окремих прикладів і при цьому часто не враховують взаємозалежності між комбінаціями значень ознак, які можуть бути вилучені з початкової вибірки. У зв'язку з цим актуальним є створення ефективних інтелектуальних правил, здатних виявляти приховані закономірності в даних і використовувати їх для прискорення процесу класифікації із збереженням прийняттого рівня точності.

Формально асоціативне правило визначається як імплікація вигляду $X \Rightarrow Y$, де X і Y є підмножинами множини ознак [2-4]. Основними характеристиками правила є підтримка (*support*) та достовірність (*confidence*), які визначаються наступним чином:

$$\text{supp}(X \Rightarrow Y) = \frac{|\{t \in D: X \cup Y \subseteq t\}|}{|D|} \quad (1)$$

$$\text{conf}(X \Rightarrow Y) = \frac{\text{supp}(X \cup Y)}{\text{supp}(X)}, \quad (2)$$

де D – множина транзакцій (об'єктів), а t – окремий елемент цієї множини. Підтримка вказує, наскільки часто група ознак з'являється разом у транзакціях, а впевненість – в скількох випадках при появі об'єкта X (антецедент) присутній також об'єкт Y (консеквент). Обидва параметри приймають значення від 0 до 1.

У розглянутому методі асоціативні правила пропонуються для використання не на повній множині дескрипторів, а на його кластерному поданні.

Розглянемо базу E описів зображень еталонів розмірністю $N: E = \{E_1, E_2, \dots, E_n\}$. Кожен еталонний опис E_i представляє у задачі розпізнавання окремий клас та має вид скінченної множини дескрипторів ключових точок зображення: $E_i = \{e_v(i)\}_{v=1}^s$, де s – число дескрипторів ключових точок у описі. Приклад зображення з координатами ключових точок (червоні пятки) наведений на рисунку 1.

Кожний дескриптор $e_v(i)$ характеризує деякий окіл ключової точки зображення та є елементом векторного простору R^n скінченної розмірності $n: e_v(i) \in R^n$ з дійсними, цілими або бінарними компонентами. Потужності описів еталонів вважаємо рівноцінними для спрощення аналізу. Застосуємо відображення $E \rightarrow T$ із простору образів (множини дескрипторів) у множину T непересічних кластерів, сформованих за деяким принципом [5, 6].

Кластер – це підмножина опису. Кожний образ E_i еталону тепер трансформується до M його непересічних підмножин $T_k(E_i)$:

$$E_i = T(E_i) = \cup_{k=1}^M T_k(E_i), k = \overline{1, M}, T_k(E_i) \cap T_j(E_i) = \emptyset.$$



Рисунок 1 – Приклад зображення з відміченими ключовими точками

За множиною $T_k(E_i)$ елементів кожного кластеру визначимо його параметр центру $b_{k,i}$, $k = \overline{1, M}$, який є базовою характеристикою побудованої кластерної системи для аналізованих даних. Отримана репрезентація бази еталонів множиною кластерів може бути представлена у вигляді матриці розміром $N \times M$, де кожна клітина відповідає кількості дескрипторів еталону, які в результаті аналізу були віднесені до того чи іншого кластеру M_i

Таблиця 1 – Приклад розбиття числа дескрипторів по кластерам для кожного еталону E_i

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
E_1	45	69	30	116	36	88	36	28	41	11
E_2	30	92	68	46	30	35	18	14	82	85
E_3	37	85	80	34	15	55	13	14	71	96
E_4	12	72	70	67	20	22	23	9	84	121

Дана таблиця може бути основою для видобутку асоціативних правил. Проте перед цим доцільно буде зробити операцію квантування, яка в даному контексті передбачає зменшення діапазону шляхом бакетування всіх можливих значень в невелику кількість груп [3].

Враховуючі, що кількість дескрипторів для кожного еталону є сталою і в нашому прикладі дорівнює $s = 500$, розіб'ємо діапазон $[0; 500]$ на 15 рівних підгруп s_i так, що $s_1 = [0; 33)$, $s_2 = [34; 67)$ і т.д. Тоді таблицю 1 можна представити у квантованому вигляді як табл. 2.

Таблиця 2 – Результат квантування кластерного подання

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
E_1	2	3	1	4	2	3	2	1	2	1
E_2	1	3	3	2	1	2	1	1	3	3
E_3	2	3	3	2	1	2	1	1	3	3
E_4	1	3	3	3	1	1	1	1	3	4

Для побудови ефективної системи правил треба шукати значення, які однозначно вказують на той чи інший клас. Також важливо мати якнайбільше зразків різних класів – таким чином можна сформулювати більш репрезентативні умови. Для даних в таблиці 2, де кожний еталон представлений одним зображенням, правила можуть виглядати наступним чином:

$$\begin{aligned}
 &\text{Якщо } M_4 = 4 \text{ або } M_6 = 3 \Rightarrow E_1 \\
 &\text{Якщо } M_1 = 1 \text{ та } M_4 = 2 \Rightarrow E_2 \\
 &\text{Якщо } M_1 = 2 \text{ та } M_8 = 1 \text{ та } M_9 = 3 \Rightarrow E_3 \\
 &\text{Якщо } M_1 = 1 \text{ та } M_6 = 1 \text{ та } M_{10} = 4 \Rightarrow E_4
 \end{aligned} \tag{3}$$

Кожне зображення-еталон E_i може бути представлено багатьма зразками, $E_{i1}, E_{i2}, \dots, E_{ij}$. У такому випадку ефективне правило-класифікатор R_i має максимізувати підтримку серед N представників одного класу (E_{ij}) та впевненість серед усіх представників еталонів:

$$\text{supp}(R_i) \subset E_{ij}, j = \overline{1, N} \rightarrow \max$$

$$\text{conf}(R_i) \subset E_{ij}, i = \overline{1, k}, j = \overline{1, N} \rightarrow \max$$

Для практичних задач рівень підтримки має складати не менше 0.5 серед усіх представників еталону, а рівень впевненості – не менше 0.75 по всій базі. Застосовуючи (3) до даних таблиці 2, і рівень підтримки, і рівень впевненості кожного правила складає 1.0.

Отримана шляхом навчання система правил (подібна до (3)) може бути використана в якості швидкого класифікатора для нових зображень. Для цього на зображенні як множині дескрипторів проводиться кластеризація і отримуються дані у формі рядку таблиці 1. Після квантування дані трансформуються у таблицю 2. Після цього на основі системи правил приймається рішення про віднесення зображення до того чи іншого класу.

Є ймовірність, що спрацюють кілька правил або не спрацює жодне. У такому випадку систему можна «донавчити» і модифікувати. Розмір та склад систем правил може різнитися і має підбиратися для досягнення бажаного рівня ефективності.

Підсумовуючи, перехід від множини дескрипторів до кластерного подання суттєво скоротив обсяг даних і зробив можливим пошук і видобування асоціативних правил, які дозволяють швидко (за умови попереднього представлення зображення-зразка у вигляді множини кластерів) визначити приналежність до класу або констатувати неможливість однозначної класифікації.

Перелік використаних джерел

1. Гороховатський В. О. Продуктивні моделі аналізу даних у методах розпізнавання зображень : монографія / В. О. Гороховатський, І. С. Творошенко. – Харків : ХНУРЕ, 2026. – 153 с. - DOI : <https://doi.org/10.30837/978-966-659-347-7>.
2. Zayko, T., Oliinyk, A., & Subbotin, S. (2014). TRAINING SAMPLE DIMENSION REDUCTION BASED ON ASSOCIATION RULES. Radio Electronics, Computer Science, Control, (1). <https://doi.org/10.15588/1607-3274-2014-1-15>.
3. Гороховатський В. О. Аналіз багатовимірних даних за описом у формі множини компонент: монографія / В. О. Гороховатський, І. С. Творошенко. – Харків : ХНУРЕ, 2022. – 124 с. – ISBN 978-966-659-379-8.
4. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., and Hudáková M. (2026) Feature space quantization and application of metrics in structural methods of image recognition, IEEE Access, vol. 14, pp. 17812-17824.
5. Gorokhovatskyi, V., Chmutov, Y., Tvoroshenko, I., & Kobylin, O. (2025). Reducing computational costs by compressing the structural description in image classification methods. Advanced Information Systems, 9(1), 5–12.
6. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., and Hudáková M. (2025) Image description compression in classification structural methods, IEEE Access, vol. 13, pp. 43631-43641.

Секція 5 Інтелектуальний аналіз даних, DataMining і BigData-технології Section 5 Data mining, DataMining and BigData technologies.....	189
АНАЛІЗ ТА СЕГМЕНТАЦІЯ ДАНИХ КОРИСТУВАЧІВ ДЛЯ ВИЯВЛЕННЯ ПОВЕДІНКОВИХ ЗАКОНОМІРНОСТЕЙ У СФЕРІ ЕЛЕКТРОННОЇ КОМЕРЦІЇ.....	190
Попова К.А., Кудрявцева М.С.	
ІНТЕЛЕКТУАЛЬНЕ NEURO-FUZZY МОДЕЛЮВАННЯ СТАНІВ ДЕГРАДАЦІЇ АГРОРЕСУРСІВ НА ОСНОВІ ЕНТРОПІЙНИХ ТА ФІЗИКО-ІНФОРМОВАНИХ ГРАФОВИХ АРХІТЕКТУР.....	192
Сільванович К.В., Гриньова О. Є.	
ЗАСТОСУВАННЯ АСОЦІАТИВНИХ ПРАВИЛ У МЕТОДАХ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ	195
Запорожченко А.П., Гороховатський В.О.	
ЕМПІРИЧНИЙ АНАЛІЗ АПАРАТНОГО МАСШТАБУВАННЯ ФРЕЙМВОРКІВ MAPREDUCE ТА АРАСНЕ SPARK ДЛЯ ІТЕРАТИВНИХ АЛГОРИТМІВ ПОКРИТТЯ.....	198
Лученко Я.В., Кобзев В.Г.	