3. Arthur D., Vassilvitskii S. k-means++: the advantages of careful seeding. Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA ’07). Philadelphia : SIAM, 2007. P. 1027–1035. URL: https://doi.org/10.1145/1283383.1283494 (дата звернення: 08.10.2026).

4. Rousseeuw P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics. 1987. Vol. 20. P. 53–65. URL: https://doi.org/10.1016/0377-0427(87)90125-7 (дата звернення: 08.10.2026).

5. Davies D. L., Bouldin D. W. A cluster separation measure. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1979. Vol. PAMI-1, No. 2. P. 224–227. URL: https://doi.org/10.1109/TPAMI.1979.4766909 (дата звернення: 08.10.2026).

6. Clustering. Scikit-learn documentation. URL: https://scikit-learn.org/stable/modules/clustering.html (дата звернення: 08.10.2026).

7. Tibshirani R., Walther G., Hastie T. Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society: Series B. 2001. Vol. 63, No. 2. P. 411–423. URL: https://doi.org/10.1111/1467-9868.00293 (дата звернення: 08.10.2026).

8. Pedregosa F., Varoquaux G., Gramfort A. et al. Scikit-learn: machine learning in Python. Journal of Machine Learning Research. 2011. Vol. 12. P. 2825–2830. URL: https://jmlr.org/papers/v12/pedregosa11a.html (дата звернення: 08.10.2026).

УДК 004.932:004.85

ПОРІВНЯННЯ КОМПАКТНОЇ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ ТА LBP-КЛАСИФІКАТОРА ЗА ОБСЯГОМ НАВЧАЛЬНОЇ ВИБІРКИ ПРИ КЛАСИФІКАЦІЇ ТЕКСТУР МАТЕРІАЛІВ

Владислав Макеєв, Олександр Жебко

Миколаївський національний аграрний університет, м. Миколаїв, Україна

Автоматичне розпізнавання матеріалів за зображеннями їхніх текстур застосовують у контролі якості, сортуванні сировини та агропромисловому виробництві. Глибокі згорткові нейронні мережі (CNN) стали основним інструментом такої класифікації [2, 3], однак вони потребують підбору навчальних даних, тоді як класичні дескриптори, як-от локальні бінарні шаблони (LBP) [1], мають вбудовану інваріантність до яскравості та повороту. На практиці розмічених зразків часто мало, тому важливо знати, з якого обсягу вибірки компактна CNN починає відповідати простому базовому методу.

Текстура задається статистикою локальних змін яскравості, тому її опис залежить від масштабу й умов освітлення. Дескриптор LBP порівнює яскравість кожного пікселя з яскравістю сусідів на колі заданого радіуса та кодує результат бінарним числом; гістограма таких кодів у фрагменті слугує вектором ознак [1]. Згорткові мережі натомість навчаються витягувати ознаки безпосередньо з пікселів, а їхні перші шари можна розглядати як банк фільтрів, параметри якого підбираються під задачу [2]. Для оцінювання таких підходів зазвичай використовують великі колекції, зокрема набір DTD із 5640 зображень у 47 категоріях [3]; водночас у прикладних задачах власних розмічених зображень зазвичай значно менше.

Мета дослідження – кількісно порівняти точність компактної CNN і класифікатора LBP+SVM залежно від кількості навчальних фрагментів текстури та перевірити стійкість обох моделей до адитивного шуму.

Використано чотири реальні напівтонові зображення 512×512 пікселів із відкритої бібліотеки scikit-image [4]: цегляна кладка, трава, гравій і поверхня Місяця. Фрагменти розміром 64×64 пікселі стандартизували (нульове середнє, одиничне стандартне відхилення) і квантували до 8 біт. Щоб уникнути витоку даних, навчальні фрагменти випадково добирали зі стовпців 0–319 кожного зображення, а тестові – зі стовпців 352–511 з кроком 16 пікселів (відступ 32 пікселі), що дало 812 тестових фрагментів (по 203 на клас). Кількість навчальних фрагментів на клас n змінювали від 2 до 256; для кожного n виконано 10 незалежних запусків з різними вибірками та ініціалізаціями. CNN

містить три блоки «згортка 3×3 – пакетна нормалізація – ReLU – max-pooling» (16, 32, 64 фільтри), глобальне усереднення та лінійний шар (23 780 параметрів), реалізована в PyTorch [5]; навчання: Adam, швидкість 10^{-3}, 40 епох, пакет 32, без аугментації. Базовий метод – гістограми рівномірних інваріантних до повороту LBP [1] для (P, R) = (8, 1), (16, 2), (24, 3) (54 ознаки, корінь із частот) і SVM із RBF-ядром (C = 10). Точність оцінювали за формулою:

$$Acc = \frac{1}{N}\sum_{i=1}^{N} 1(\hat{y}_i = y_i) \quad (1)$$

де N – кількість тестових фрагментів; y_i та $\hat{y}_i$ – істинний і передбачений клас; $1(\cdot)$ – індикатор події. Код LBP для центрального пікселя з яскравістю g_c та P сусідів g_p на радіусі R визначається співвідношенням [1]:

$$LBP(P,R) = \sum_{p=0}^{P-1} s(g_p - g_c)\cdot 2^p\,,\ s(x) = 1 \text{ при } x \geq 0, s(x) = 0 \text{ при } x < 0 \quad (2)$$

Додатково тестові фрагменти зашумлювали гауссовим шумом із σ = 0,5 (у стандартизованих одиницях); моделі навчали лише на чистих даних.

Порівняння з LBP обрано тому, що цей дескриптор є усталеним базовим методом класифікації текстур [1], а згорткові шари мережі можна розглядати як навчані банки фільтрів [2]. Компактну архітектуру використано навмисно: за малих вибірок велика мережа переозначується, а дослідження вимагає лише порівняння за зростання обсягу даних. Розділення навчальних і тестових фрагментів за просторовими ділянками виключає ситуацію, коли однакові пікселі потрапляють в обидві вибірки.

Узагальнені результати наведено в табл. 1 і на рис. 1. Для LBP+SVM точність на чистих даних практично дорівнює 100 % уже за n = 2. CNN за n = 2 показує 86.2 ± 14.9 %, за n = 16 – 97.8 ± 5.2 %, а починаючи з n = 32 усі 10 запусків дають 100 %. Великий розкид за малих n свідчить, що мережа має «вивчити» фільтри з небагатьох прикладів, тоді як LBP задає ознаки апріорі. За наявності шуму обидві моделі втрачають працездатність: точність CNN становить 29–35 %, а LBP+SVM – близько 25 %, тобто рівень випадкового вгадування для чотирьох класів. Причина: шум змінює знаки малих різниць яскравості, на яких побудовано LBP, а мережа не бачила такого викривлення під час навчання.

Динаміка точності CNN має два характерні етапи. За n від 2 до 8 середня точність змінюється мало (86,2; 89,1 та 89,3 %), а стандартне відхилення залишається великим (12,5–14,9 в. п.): результат істотно залежить від того, які саме фрагменти потрапили до навчальної вибірки та з яких початкових ваг стартувало навчання. Помітніше зростання спостерігається між n = 8 і n = 16, де середня точність досягає 97,8 %, а розкид зменшується до 5,2 в. п.; починаючи з n = 32 усі запуски збігаються.

Результати на зашумлених даних слід трактувати обережно. Для чотирьох збалансованих класів рівень випадкового вгадування становить 25 %, і саме на цьому рівні перебуває LBP+SVM (25,0–25,1 %); це узгоджується з тим, що класифікатор відносить усі фрагменти до одного класу, проте підтвердження потребує матриці помилок, яку в роботі не наводили. Точність CNN (29,0–34,9 %) лише на кілька відсоткових пунктів перевищує випадковий рівень і не зростає з обсягом навчальної вибірки: додаткові чисті приклади не компенсують відсутності шуму в навчальних даних.

Таблиця 1. Точність класифікації на тестових фрагментах, % (середнє ± СВ за 10 запусків)

n	CNN, чисті дані	LBP+SVM, чисті дані	CNN, шум σ=0,5	LBP+SVM, шум σ=0,5
2	86.2 ± 14.9	100.0 ± 0.0	31.0	25.0
4	89.1 ± 12.5	100.0 ± 0.0	29.5	25.0
8	89.3 ± 13.1	100.0 ± 0.0	34.9	25.1
16	97.8 ± 5.2	100.0 ± 0.0	32.3	25.0
32	100.0 ± 0.0	100.0 ± 0.0	30.6	25.0
64	100.0 ± 0.0	100.0 ± 0.0	30.5	25.0
128	100.0 ± 0.0	100.0 ± 0.0	29.0	25.0

n	CNN, чисті дані	LBP+SVM, чисті дані	CNN, шум σ=0,5	LBP+SVM, шум σ=0,5
256	100.0 ± 0.0	100.0 ± 0.0	29.0	25.0

Для задач із малою кількістю розмічених зображень простий базовий метод LBP+SVM слід перевіряти першим; перехід до CNN виправданий, коли є щонайменше близько 30 фрагментів на клас, а також для складніших класів і за потреби навчання з аугментацією, зокрема шумовою.
Практичне значення отриманих результатів полягає в порядку дій під час розроблення систем візуального контролю. Спочатку варто побудувати простий еталон на основі LBP і SVM, який має небагато параметрів і не потребує великого навчального набору; лише після цього, якщо точність недостатня або класи складніші, переходити до CNN. Оцінювати моделі слід і на зашумлених зображеннях, оскільки в реальних умовах зйомки спотворення неминучі.

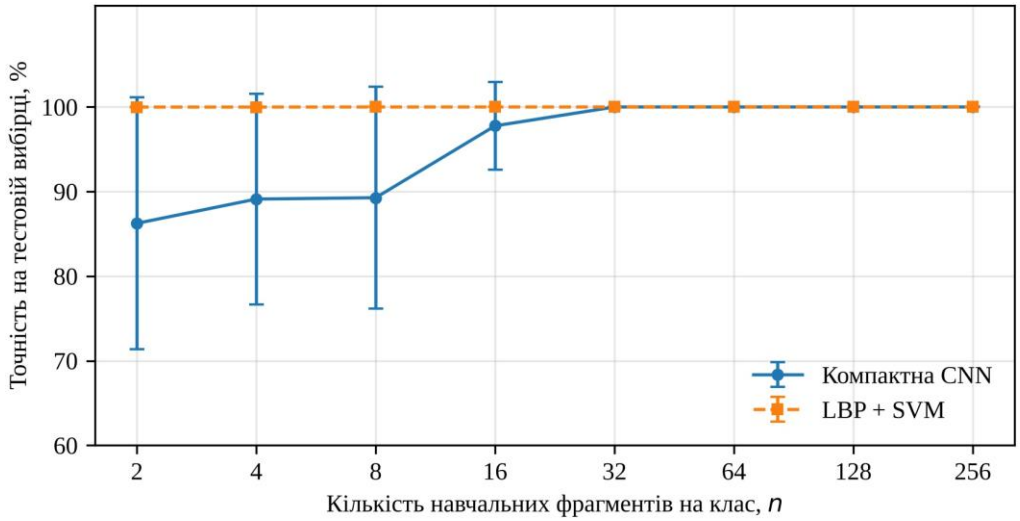


Рис. 1. Залежність точності від кількості навчальних фрагментів на клас (стовпчики похибок – СВ)

Використано лише чотири зображення (по одному на клас), а тестові фрагменти узято з іншої ділянки того самого зображення; тому оцінка не відображає узагальнення між різними зразками матеріалу, і вихід LBP на рівень 100 % зумовлений легкістю задачі. Тестові фрагменти перекриваються, тож вони не є повністю незалежними. Результати не можна переносити на великі набори на кшталт DTD [3].

Подальші дослідження доцільно спрямувати на розділення даних за різними зразками матеріалу, збільшення кількості класів (наприклад, на наборі DTD [3]), шумову аугментацію під час навчання CNN та вимірювання часу навчання й виведення.

Висновки. 1. На чотирьох класах текстур LBP+SVM досягає майже 100 % точності вже за 2 фрагментів на клас, а компактна CNN (23 780 параметрів) – лише від 32 фрагментів на клас. 2. Обидві моделі, навчені на чистих даних, нестійкі до гауссового шуму σ = 0,5. 3. Для коректної оцінки переваг CNN необхідні міжзразкове розділення даних і складніші набори.

Список використаних джерел

1. Ojala T., Pietikäinen M., Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002. Vol. 24, No. 7. P. 971–987. DOI: https://doi.org/10.1109/TPAMI.2002.1017623.

2. Andrearczyk V., Whelan P. F. Using filter banks in convolutional neural networks for texture classification. Pattern Recognition Letters. 2016. Vol. 84. P. 63–69. DOI: https://doi.org/10.1016/j.patrec.2016.08.016.

3. Cimpoi M., Maji S., Kokkinos I., Mohamed S., Vedaldi A. Describing textures in the wild. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2014. P. 3606–3613.

4. van der Walt S., Schönberger J. L., Nunez-Iglesias J. et al. scikit-image: image processing in Python. PeerJ. 2014. Vol. 2. Art. e453. DOI: https://doi.org/10.7717/peerj.453.

5. Paszke A., Gross S., Massa F. et al. PyTorch: an imperative style, high-performance deep learning library. Advances in Neural Information Processing Systems 32 (NeurIPS 2019). 2019. P. 8024–8035.

УДК 519.237.8

КЛАСТЕРИЗАЦІЯ КРАЇН ЗА РІВНЕМ ЕКОНОМІЧНОГО РОЗВИТКУ ТА ОЧІКУВАНОЮ ТРИВАЛІСТЮ ЖИТТЯ МЕТОДОМ K-СЕРЕДНІХ

Олександр Хлівний

Миколаївський національний аграрний університет, м. Миколаїв, Україна

Науковий керівник: О. О. Жебко,
асистент Миколаївський національний аграрний університет

Соціально-економічний розвиток держав є багатовимірним явищем, яке неможливо достовірно охарактеризувати одним рейтингом або окремим абсолютним показником. Обсяг національного виробництва свідчить про економічні можливості, однак сам по собі не описує стан здоров’я населення, тривалість життя чи доступність суспільних благ. Країни зі схожими показниками виробництва на одну особу можуть відрізнятися за соціальними результатами, а держави з різним економічним потенціалом – демонструвати близькі демографічні характеристики. Тому актуальним є використання методів багатовимірного аналізу, які дозволяють виявляти групи подібних країн без попереднього призначення їм нормативних категорій.

Кластерний аналіз групує спостереження за близькістю обраних характеристик і, на відміну від ранжування лише за ВВП, враховує кілька ознак одночасно. Оскільки результат залежить від вибору показників, нормування та складу вибірки, важливо оцінювати не лише склад груп, а й якість розбиття. Теоретичну основу алгоритму k-середніх закладено Дж. Мак-Квіном, а силуетний критерій запропоновано П. Руссеєм [5; 6].

Метою дослідження є визначення можливості виокремити інтерпретовані групи країн за поєднанням економічного та соціально-демографічного показників на невеликій міжрегіональній вибірці за 2023 рік. Дослідницьке питання сформульовано так: чи дозволяє стандартизована двовимірна модель k-середніх розрізнити три групи країн із помітно відмінними профілями ВВП на одну особу та очікуваної тривалості життя? Об’єкт дослідження – відібрані країни світу; предмет – структура їх подібності у просторі двох кількісних ознак. Робота має характер відтворюваного навчально-аналітичного обчислювального експерименту, а не репрезентативної класифікації всіх держав.

Інформаційну основу становлять два відкриті масиви даних за календарний 2023 рік. Економічна змінна – ВВП на одну особу за паритетом купівельної спроможності (ПКС) у постійних міжнародних доларах 2021 року, код Світового банку NY.GDP.PCAP.PP.KD [1]. Урахування ПКС зменшує спотворення порівняння, пов’язане з відмінностями рівнів цін між країнами. Соціальна змінна – очікувана тривалість життя при народженні для обох статей, років, за оцінками Відділу народонаселення Департаменту ООН з економічних і соціальних питань у виданні World Population Prospects 2024 [2]. На рівні країни вона відображає узагальнену поточну структуру смертності, а не гарантовану тривалість життя конкретної людини.

Значення обох показників звірено з відкритими табличними виданнями за 2023 рік, що відтворюють дані відповідних першоджерел [3; 4]. Для зіставності використовуються 12 країн, відібраних цілеспрямовано з різних географічних регіонів і діапазонів доходу: Норвегія, США, Німеччина, Японія, Польща, Туреччина, Китай, Бразилія, Україна, Південно-Африканська