

1. Стригіна О. А., Ревицька У. С. Побудова моделей сегментації ринку з використанням кластерного аналізу. Grail of science. 2026. № 61. С. 326–330. URL: <https://shorturl.at/TYNg2>
2. Andrusyk Y. V., Kaganovskyi O. S. Data science methods: analyzing approaches to customer segmentation. Business inform. 2025. Vol. 1, no. 564. P. 497–503. URL: <https://doi.org/10.32983/2222-4459-2025-1-497-503>

**Abstract:** *The paper examines the application of cluster analysis methods for the segmentation of customers in the agricultural products market based on behavioral data. The possibilities of using indicators such as purchase frequency, order volume, demand seasonality, and other characteristics of purchasing behavior for forming homogeneous groups of consumers are analyzed. The feasibility of applying clustering algorithms to identify hidden patterns in customer data is substantiated. It is shown that the results of clustering allow enterprises to improve the effectiveness of marketing and sales strategies, optimize logistics processes, and enhance demand forecasting. An example of the practical implementation of customer clustering using the Python programming language and the k-means algorithm is presented.*

**Keywords:** *customer clustering, behavioral data, consumer segmentation, agricultural products market, data analysis, k-means algorithm, agricultural market, marketing analytics.*

**Науковий керівник:**

**Пархоменко О. Ю.,**

*канд. фіз.-мат.наук, доцент, доцент кафедри економічної кібернетики, комп'ютерних наук та інформаційних технологій*

*Миколаївський національний аграрний університет*

**УДК 004.93:006.7**

**Класифікація відеоконтенту за категоріями**

**Касьяненко Олександр,**

*здобувач вищої освіти спеціальності 122 «Комп'ютерні науки»*

*Миколаївський національний аграрний університет*

*м. Миколаїв, Україна*

**Анотація:** *розглянуто методи автоматичної класифікації відеоконтенту за категоріями з використанням методів машинного навчання та глибоких нейронних мереж. Проаналізовано основні підходи до виділення ознак із відеопослідовностей – просторових, часових та аудіовізуальних. Описано архітектури згорткових і рекурентних мереж, трансформерів, що застосовуються для вирішення задачі категоризації відео. Визначено*

*перспективи розвитку систем класифікації у контексті зростаючих обсягів відеоданих.*

**Ключові слова:** *глибоке навчання, нейронні мережі, обробка відеоданих, комп'ютерний зір, мультимодальний аналіз, категоризація контенту.*

Стрімке зростання обсягів відеоконтенту в мережі Інтернет ставить перед дослідниками та розробниками нові завдання автоматичної обробки та організації мультимедійних даних. Щохвилини на відеохостингові платформи завантажуються сотні годин відео, що унеможлиблює ручну систематизацію та класифікацію такого масиву інформації. Автоматична класифікація відеоконтенту за тематичними категоріями є ключовою задачею у сферах рекомендаційних систем, модерації контенту, пошуку та цифрової бібліотекознавства.

Задача класифікації відео суттєво відрізняється від класифікації статичних зображень завдяки темпоральному виміру – відеопослідовність несе в собі інформацію про рух, зміну сцен та динаміку подій. Крім візуальної складової, відео може містити аудіодоріжку, субтитри та метадані, які є важливими джерелами ознак для навчання класифікатора. Відтак ефективні системи категоризації відео мають враховувати мультимодальну природу таких даних.

Традиційні підходи до класифікації відео ґрунтувалися на ручному виділенні ознак – гістограм спрямованих градієнтів (HOG), оптичного потоку, SIFT-дескрипторів із подальшим застосуванням класичних методів машинного навчання: метода опорних векторів (SVM), випадкових лісів (Random Forest), k-найближчих сусідів (k-NN). Такий підхід вимагав значних зусиль у проектуванні ознакового простору та мав обмежену здатність до узагальнення на нових типах контенту [1].

Сучасні методи класифікації відеоконтенту базуються на глибокому навчанні, яке дозволяє автоматично виділяти ієрархічні ознаки безпосередньо з даних. Серед основних архітектур виділяють [3]:

- двовимірні згорткові нейронні мережі (2D CNN), що обробляють окремі кадри відео як зображення та агрегують передбачення по всьому відеофрагменту;
- тривимірні згорткові мережі (3D CNN / C3D), які одночасно враховують просторові та часові залежності між кадрами;
- гібридні моделі типу Two-Stream, де один потік обробляє RGB-кадри, а другий – поле оптичного потоку між послідовними кадрами;
- рекурентні нейронні мережі (RNN, LSTM), що моделюють часові залежності у послідовності ознак, отриманих від CNN;
- трансформерні архітектури (Vision Transformer, Video Swin Transformer), які застосовують механізм само уваги для глобальної обробки просторово-часових патернів. У таблиці 1 наведемо порівняльну характеристику методів класифікації відеоконтенту.

Таблиця 1 Порівняльна характеристика методів класифікації відеоконтенту

Метод	Переваги	Обмеження
2D CNN	Простота реалізації, висока швидкість обробки	Не враховує часову динаміку відео
3D CNN (C3D)	Враховує просторово-часові ознаки	Висока обчислювальна вартість
Two-Stream	Ефективне виділення руху через оптичний потік	Потребує додаткового розрахунку потоку
CNN + LSTM	Модельовання довгострокових часових залежностей	Складність навчання, затримки в обробці
Video Transformer	Глобальний контекст, висока точність	Значний обсяг параметрів, потребує великих датасетів

Джерело: розроблено автором

Для підвищення якості класифікації активно застосовуються мультимодальні підходи, що поєднують аналіз візуального ряду, аудіодоріжки та текстових метаданих (назва, опис, теги). Наприклад, у задачах розпізнавання навчальних відео поєднання CNN-ознак із транскрипцією мовлення через BERT-подібні моделі дає суттєво вищу точність порівняно з однобічними підходами.

Важливим аспектом побудови системи класифікації є формування якісних навчальних наборів даних. Серед найпоширеніших публічних датасетів для класифікації відео виділяють: UCF-101, Kinetics-400/600/700, ActivityNet, YouTube-8M та HowTo100M. Кожен із них охоплює специфічний діапазон категорій – від побутових дій до спортивних сцен і навчальних матеріалів. Зокрема, YouTube-8M містить понад 8 мільйонів відеофрагментів та 3800 класів, що робить його одним із найбільших публічних відеодатасетів [2,3].

Типовий пайплайн класифікації відеоконтенту включає такі етапи: попередня обробка відео (декодування, семплування кадрів, нормалізація), виділення ознак за допомогою обраної нейромережевої архітектури, агрегація часових ознак, класифікація та після-обробка результатів. Для систем реального часу критичним є збалансування між точністю класифікації та обчислювальною ефективністю, що зумовлює застосування легких архітектур типу MobileNet або EfficientNet із механізмом уваги [3].

Перспективними напрямками розвитку систем класифікації відеоконтенту є комплексне поєднання сучасних підходів штучного інтелекту, спрямованих на підвищення точності, адаптивності та обчислювальної ефективності моделей. Зокрема, особливої актуальності набуває застосування методів само-навчання (self-supervised learning), які дозволяють використовувати великі обсяги нерозмічених відеоданих для попереднього навчання моделей, суттєво знижуючи залежність від дорогих і трудомістких процесів анотації. У цьому контексті ефективними є контрастивні підходи та трансформерні архітектури, що здатні виявляти латентні просторово-часові залежності у відеопотоці [2].

Не менш важливим є розвиток методів кросмодальної передачі знань (cross-modal transfer learning), які забезпечують інтеграцію інформації з різних джерел – зокрема відео, аудіо та текстових метаданих – для формування більш повного семантичного представлення контенту. Це відкриває можливості для побудови мультимодальних моделей, здатних підвищувати якість класифікації в умовах неповних або шумних даних, а також покращувати узагальнювальну здатність систем [1].

Окремим перспективним напрямом є впровадження інтерпретованих моделей (explainable AI), які дозволяють не лише отримувати результат класифікації, але й пояснювати причини прийнятих рішень. Це особливо важливо для критичних застосувань (безпека, медіамоніторинг, державне управління), де необхідна прозорість алгоритмів. Використання механізмів уваги (attention), візуалізація активацій (Grad-CAM) та побудова сурогатних моделей сприяють підвищенню довіри до таких систем [3].

З практичної точки зору вибір моделі класифікації відеоконтенту залежить від обсягу даних, вимог до точності та обчислювальних ресурсів. Для задач із обмеженими ресурсами та потребою у швидкій обробці доцільно використовувати легковагові архітектури, такі як MobileNet або EfficientNet у поєднанні з простими класифікаторами. У випадку середніх за складністю задач ефективними є ансамблеві методи машинного навчання (Random Forest, Gradient Boosting), які можуть працювати на попередньо витягнутих ознаках. Водночас для високоточного аналізу складних відеосцен перевагу слід надавати глибоким нейронним мережам — 3D CNN, CNN+LSTM або трансформерним моделям, які здатні враховувати як просторові, так і часові залежності [2, 3].

Особливо перспективним є використання гібридних підходів, що поєднують різні типи моделей та модальностей даних. Наприклад, комбінація згорткових мереж для обробки відеоряду, рекурентних або трансформерних моделей для аналізу часової динаміки та мовних моделей (BERT-подібних) для текстових описів дозволяє значно підвищити якість класифікації. Такий підхід забезпечує більш повне семантичне розуміння контенту та підвищує стійкість системи до шумів і неповноти даних, що є критично важливим для реальних застосувань.

Таким чином, класифікація відеоконтенту за категоріями є актуальною та динамічно розвиваючою областю досліджень на перетині комп'ютерного зору, обробки природної мови та машинного навчання. Впровадження ефективних систем автоматичної категоризації відкриває широкі можливості для покращення досвіду користувачів медіаплатформ, підвищення ефективності модерації та розвитку персоналізованих рекомендаційних сервісів.

### **Список використаних джерел:**

1. Zisserman A., Simonyan K. Two-Stream Convolutional Networks for Action Recognition in Videos // Advances in Neural Information Processing Systems. 2014. Vol. 27. URL: <https://arxiv.org/abs/1406.2199>.

2. Tran D., Bourdev L., Fergus R., Torresani L., Paluri M. Learning Spatiotemporal Features with 3D Convolutional Networks // Proceedings of the IEEE ICCV. 2015. P. 4489–4497. URL: <https://doi.org/10.1109/ICCV.2015.510>.

3. Abu-El-Hajja S. et al. YouTube-8M: A Large-Scale Video Classification Benchmark. 2016. URL: <https://arxiv.org/abs/1609.08675>.

4. Liu Z. et al. Video Swin Transformer // Proceedings of the IEEE CVPR. 2022. P. 3202–3211. URL: <https://doi.org/10.1109/CVPR52688.2022.00320>.

**Abstract:** *Methods for automatic classification of video content by categories using machine learning methods and deep neural networks are considered. The main approaches to feature extraction from video sequences - spatial, temporal and audiovisual - are analyzed. The architectures of convolutional and recurrent networks, transformers used to solve the problem of video categorization are described. The prospects for the development of classification systems in the context of growing volumes of video data are determined.*

**Keywords:** *deep learning, neural networks, video data processing, computer vision, multimodal analysis, content categorization.*

**Науковий керівник:**

**Жебко О.О.,**

*асистент кафедри економічної кібернетики,  
комп'ютерних наук та інформаційних технологій,  
Миколаївський національний аграрний університет*

**УДК 004.932.2:004.932.72:004.91**

**Інформаційні технології розпізнавання рукописних даних у розбудові цифрової інфраструктури сільського господарства**

**Козуб Іван,**

здобувач вищої освіти спеціальності 122 «Комп'ютерні науки»

Миколаївський національний аграрний університет,

м. Миколаїв, Україна

**Анотація:** *Сучасний етап цифровізації сільського господарства вимагає інтеграції всіх доступних масивів інформації у єдину систему управління. Значна частина історичних та поточних агрономічних даних досі фіксується на паперових носіях. Журнали обліку використання добрив, ветеринарні записи та польові нотатки залишаються поза межами автоматизованого аналізу. Впровадження технологій розпізнавання рукописного тексту є критично важливим кроком для повноцінної розбудови цифрової інфраструктури аграрного сектора.*

**Ключові слова:** *розпізнавання тексту, нейронні мережі, цифровізація, сільське господарство, автоматизація, документообіг.*