

МЕТОДИ СЕМАНТИЧНОГО АНАЛІЗУ ВІДЕОДАНИХ

Макаров Д.С.

e-mail: dmytro.makarov@nure.ua

Науковий керівник – к.т.н., доц. Кобилін О.А.

Харківський національний університет радіоелектроніки, каф. ІНФ
м. Харків, Україна

Semantic analysis of video data involves extracting meaningful information and understanding the context and objects within the video content. With the rapid growth of video data in recent years, there has been a significant interest in developing techniques for automatic interpretation [1]. This report explores various methods for semantic analysis, including object recognition, activity recognition, scene understanding, and deep learning-based approaches. The use of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for analyzing temporal and spatial data is discussed, alongside challenges such as data complexity, scalability, and real-time processing [1].

У зв'язку з експоненційним зростанням обсягів відеоданих у різних сферах діяльності, зокрема в галузі безпеки, медицини та автоматизації, виникає необхідність у застосуванні методів семантичного аналізу відео для автоматичної обробки та інтерпретації даних [2]. Постійне збільшення кількості відеоінформації робить неможливим її ручний аналіз, тому ефективні алгоритми обробки є ключовими для забезпечення швидкого і точного розпізнавання важливих подій. Семантичний аналіз дозволяє витягати змістовну інформацію з відеофайлів, включаючи виявлення об'єктів, розпізнавання дій, ідентифікацію сцен та інші важливі аспекти.

Методи семантичного аналізу відео спрямовані на автоматизацію процесу інтерпретації відеозаписів, що включає аналіз вмісту сцени, об'єктів і подій у динамічному середовищі. Основні підходи включають застосування алгоритмів для виявлення і класифікації об'єктів у відеопотоці. Для цього використовуються глибокі нейронні мережі, такі як CNN, YOLO та SSD, що дозволяють ефективно детектувати та розпізнавати об'єкти в кожному кадрі [3]. Окрім того, просторовий аналіз допомагає побудувати карти об'єктів, визначити їхній контекст і взаємодію в межах сцени.

Розпізнавання дій та активностей у відео передбачає не лише ідентифікацію об'єктів, а й аналіз їхньої взаємодії у часовому контексті. Для цього використовуються методи, що дозволяють обробляти послідовність кадрів. Зокрема, рекурентні нейронні мережі, такі як LSTM (Long Short-Term Memory), ефективно моделюють тимчасові залежності у відеоданих [2].

Розуміння контексту сцени передбачає семантичний аналіз, що поєднує просторові та тимчасові дані для формування цілісного уявлення про події. Важливим аспектом є контекстуальне розпізнавання, яке допомагає визна-

чити значущі об'єкти в конкретній ситуації, наприклад, виявлення небезпеки на камерах спостереження. Також ефективним підходом є аналіз взаємодії об'єктів, що базується на комбінуванні різних типів даних, таких як звук, текстові підписи та зображення.

Глибокі нейронні мережі є ключовими інструментами семантичного аналізу відео, забезпечуючи ефективне розпізнавання об'єктів і їх взаємодій. Конволюційні нейронні мережі (CNN) дозволяють виявляти та класифікувати об'єкти у відео, тоді як гібридні моделі поєднують їх із рекурентними нейронними мережами для врахування тимчасових залежностей. Використання моделей трансформерів забезпечує обробку послідовностей із збереженням глобального контексту, що особливо важливо для аналізу складних сцен [1].

Методи семантичного аналізу відеоданих досягли значного прогресу, проте залишаються певні виклики. Масштабованість та обробка великих обсягів інформації потребують значних обчислювальних ресурсів. Забезпечення аналізу в реальному часі залишається складним завданням, особливо для застосувань, де важлива миттєва реакція, як-от у сфері автономного транспорту. Труднощі виникають також при інтерпретації контексту сцени, коли об'єкти та їхні дії є неоднозначними.

Подальший розвиток семантичного аналізу відео зосереджується на підвищенні точності та адаптивності моделей. Інтеграція мультимодальних даних, таких як відео, аудіо та текст, сприяє кращому розумінню ситуацій. Оптимізація алгоритмів для швидкої обробки дозволяє застосовувати їх у критичних сферах, включно з безпекою та автономним транспортом. Удосконалення методів навчання без нагляду допомагає зменшити залежність від розмічених даних [3].

Методи семантичного аналізу відео широко застосовуються у різних галузях. У сфері безпеки вони допомагають виявляти загрози та аналізувати поведінку в реальному часі. У медицині ці технології використовуються для моніторингу пацієнтів, аналізу хірургічних операцій і діагностики захворювань. Автономний транспорт потребує цих методів для розпізнавання об'єктів на дорозі та навігації в змінних умовах. У розважальній індустрії аналіз відео автоматизує обробку контенту, а у віртуальній реальності забезпечує прозору адаптацію середовищ під користувача [4].

Семантичний аналіз відеоданих є ключовим напрямом сучасних досліджень у галузі комп'ютерного зору та штучного інтелекту. Незважаючи на значні досягнення, технології потребують подальшого вдосконалення для забезпечення точності, швидкості обробки та адаптації до різних типів відеоконтенту. Подальший розвиток алгоритмів включає інтеграцію мультимодальних джерел даних, удосконалення методів навчання без нагляду та розробку більш ефективних механізмів інтерпретації відеосцен [2]. Це до-

зволить підвищити якість аналізу в умовах великих обсягів даних та розширити можливості для реального часу, що є критичним для застосувань в безпеці, медицині та автономному транспорті.

Список використаних джерел:

1. Kobylin, O. A., Gorokhovatskyi, V. O., Tvoroshenko, I. S., & Peredrii, O. O. (2020). The application of non-parametric statistics methods in image classifiers based on structural description components. *Telecommunications and Radio Engineering*, 79(10).
2. Коваленко, О. І., & Мартиненко, В. В. (2021). Аналіз методів розпізнавання об'єктів в відеопотоці на основі глибоких нейронних мереж. *Інформаційні технології і засоби навчання*, 71(2), 46-58.
3. Мельник П. В., Шестопапов С. В. Методи семантичного аналізу в системах відеоспостереження // *Вісник Київського національного університету технологій та дизайну*. 2018. №2. С. 45-51.
4. Xie, L., Zhang, S., & Zhang, C. (2020). Semantic Video Analysis: Current Challenges and Future Directions. *Journal of Visual Communication and Image Representation*, 71, 50-65.