

УДК 004.932:004.8

*Лазаренко М.В., здобувач
Рудніченко М.Д., к.т.н., доцент,
Шибасєва Н.О., к.т.н., доцент,
Національний університет «Одеська політехніка»*

АНАЛІЗ ІЄРАРХІЧНИХ АГЕНТНИХ МОДЕЛЕЙ ОБРОБКИ МУЛЬТИМОДАЛЬНИХ ДАНИХ В ПРОЦЕСАХ РОЗПІЗНАВАННЯ ОБРАЗІВ

Вступ. Ієрархічні агентні моделі пропонують архітектурне рішення, що дозволяє розподілити завдання обробки між рівнями агентів із різною спеціалізованою функціональністю, поєднуючи локальні та глобальні стратегії розпізнавання та інтерпретації образів. Актуальність наряду обумовлена складністю сучасних даних, що надходять із різномірних джерел, і потребою синтезувати семантичну інформацію на різних рівнях абстракції. Ієрархічні агентні структури здатні ефективно інтегрувати фрактальні, часові, контекстні і семантичні ознаки, розподіляючи обчислювальні ресурси та забезпечуючи адаптацію процесу обробки залежно від складності вхідних сигналів [1].

Основна частина. Поява мультимодальних даних зумовила потребу розвитку моделей, здатних не лише обробляти окремі потоки інформації, але й інтегрувати зміст з різних сенсорних доменів. Ієрархічні агентні моделі передбачають організацію обчислювальних одиниць (агентів) у багатоетапні структури, де кожен агент або група агентів спеціалізуються на окремих аспектах мультисенсорної обробки. На нижчому рівні такої ієрархії агенти виконують первинну фільтрацію та локальні трансформації даних.

Вищі рівні агентних ієрархій виконують інтеграцію локально оброблених результатів для формування глобальних представлень. Ці рівні включають більш складні когнітивні механізми для семантичної інтерпретації сцен, встановлення причинно-наслідкових зв'язків, виділення об'єктних категорій та контекстних відношень між ними. Комунікація між агентами здійснюється через стандартизовані інтерфейси обміну інформацією, що дозволяє агентам адаптувати свої алгоритмічні стратегії залежно від отриманих контекстних сигнатур [2].

Ієрархічний підхід дозволяє вирішувати ключові проблеми, характерні для обробки мультимодальних даних: узгодження неоднорідних представлень, усунення суперечливих сигналів, інтеграцію інформації різної дискретності, а також адаптацію до динамічних змін контексту. Наприклад, у складних задачах розпізнавання сцен агенти нижчого рівня можуть фокусуватися на

локальних просторових ознаках, тоді як агенти вищого рівня інтегрують ці ознаки для побудови моделі об'єктних взаємозв'язків.

Однією з архітектурних моделей, що знайшла застосування в цій сфері, є багаторівневий агентний граф, де кожен вузол графу представляє окремих агента з певною спеціалізацією. Такі графи можуть бути побудовані за принципом розширеної глибокої ієрархії, в якій інформація просувається вгору по рівнях для агрегування та семантичної інтерпретації, а також назад для уточнення локальних оцінок із урахуванням глобального контексту.

Деякі підходи передбачають використання гібридних структур, де ієрархічні агентні модулі поєднують класичні алгоритми машинного навчання (нейромережі, класифікатори) зі структурованими комунікаційними правилами, що дозволяють оптимізувати процес розпізнавання шляхом узгодження локальних і глобальних оцінок [3].

Серед сучасних розробок у цій сфері – моделі, що поєднують агентні механізми з глибокими гібридними архітектурами для інтеграції візуальних та текстових даних, що суттєво покращує якість семантичного узгодження розпізнаваних об'єктів у мультисенсорних системах. Серед основних тенденцій у розвитку ієрархічних агентних моделей – активне використання механізмів attention-подібних структур, що дозволяють агентам зосереджувати свою увагу на релевантних частинах вхідних даних залежно від поточного контексту.

Висновки. Ієрархічні агентні моделі становлять перспективний напрям у сфері обробки мультимодальних даних завдяки своїй здатності розподіляти завдання між рівнями агентів, інтегрувати неоднорідні сигнали та забезпечувати адаптивність при розпізнаванні складних образів. Аналіз сучасних архітектур свідчить про необхідність подальшого розвитку механізмів комунікації між агентами, інтеграції гібридних методів навчання та адаптації до динамічних контекстів.

Список використаних джерел:

1. Khan S., Hayat M., Bennamoun M., Sohel F. Multimodal Deep Learning for Robust RGB-D Object Recognition. *IEEE Transactions on Image Processing*. 2023. Vol. 32. P. 3175–3187.
2. Guo Y., Wang Y., Wu H., Lin Y. Hierarchical Agent-Based Framework for Multimodal Perception in Autonomous Systems. *Journal of Artificial Intelligence Research*. 2024. Vol. 71. P. 255–284.
3. Zhang Z., Zheng L., Yang Y. Graph Attention Networks for Multimodal Fusion in Object Recognition. *Pattern Recognition*. 2025. Vol. 127. P. 108538–108538.