

ПОРІВНЯЛЬНИЙ АНАЛІЗ АРХІТЕКТУР YOLO ДЛЯ ВИЯВЛЕННЯ ОСЕРЕДКІВ ЗАЙМАННЯ НА ВІДЕОПОТОЦІ

UDC 004.932.2:004.421.3:614.84

T. Semtsiv

COMPARATIVE ANALYSIS OF YOLO ARCHITECTURES FOR DETECTION OF ENGAGEMENT CENTERS IN VIDEO STREAM

Проблема оперативного виявлення осередків займання у відкритому міському середовищі є важливим завданням для забезпечення техногенної безпеки, оскільки традиційні системи протипожежної сигналізації часто виявляються непридатними для моніторингу просторих територій. Перехід до комп'ютеризованих систем відеонагляду, підсилених методами глибокого навчання, дозволяє в реальному часі виявляти пожежу. Для служб екстреного реагування критичною є вимога щодо досягнення оптимального балансу між надійністю ідентифікації (mAP) та показником швидкості опрацювання (FPS). Через це актуальними є моделі, які можуть працювати в реальному режимі часу на звичайних Edge-пристроях, і при цьому не втрачати точності. Такими моделями є YOLO, які дозволяють опрацьовувати відеопотік та виявляти в ньому задані об'єкти [1].

Серед версій YOLO варто виокремити YOLOv11, YOLOv12 та YOLOv13, оскільки саме вони пропонують помітні зміни в тому, як модель збирає та інтерпретує інформацію з кадру [1].

YOLOv11 загалом зосереджена на тому, щоб зробити детекцію стабільнішою та швидшою без зайвих ускладнень у структурі. У YOLOv12 з'явилися механізми уваги, які допомагають краще розпізнавати контекст сцени – це важливо тоді, коли дим або слабе полум'я зливаються з фоном. У YOLOv13 використано HyperACE – підхід, що дає змогу моделі «бачити» зв'язки між ознаками ширше, ніж у попередніх версіях, тому вона впевненіше знаходить навіть малопомітні сигнали займання.

Щоб оцінити реальні можливості моделей, застосовувався датасет із ознаками пожежі, а головними метриками були mAP, FPS і FLOPs (див. табл. 1).

Таблиця 1. Порівняння архітектурних характеристик версій моделей YOLO

Метрика	YOLOv11-n	YOLOv12-n	YOLOv13-n
Detection Head	Розділена	Розділена+увага	HyperACE
FLOPs	Дуже низька	Низька	Середня
mAP	Швидкодія	Робота з акцентом	Підвищена точність
FPS	Висока	Висока	Середньо-висока

З огляду на результати, YOLOv11 найкраще підходить для систем, де вирішальними є швидкість і мінімальні затримки. YOLOv12 краще справляється зі складними сценами, де важливо враховувати контекст. YOLOv13 показує найвищу точність у випадках, коли ознаки займання ледве помітні. Такий аналіз може бути корисним при виборі моделі для реальних протипожежних відеосистем.

Література

1. Ultralytics. Models Documentation [Електронний ресурс]. – Режим доступу: <https://docs.ultralytics.com/models/>.