

АНАЛІЗ МЕТОДІВ ФОРМУВАННЯ ТЕПЛОВИХ КАРТ ДЛЯ АДАПТИВНОГО ПОКРАЩЕННЯ ЗОБРАЖЕНЬ З БПЛА

Теплові карти нейронних мереж використовуються для візуалізації регіонів, які модель вважає найбільш інформативними. Вони можуть слугувати індикатором інформативності кадру або окремого його фрагмента. Це забезпечує можливість активувати суперрезолюцію (SR) лише там, де її використання виправдане, оскільки вона потребує значних обчислювальних ресурсів.

Адаптивна логіка використання SR передбачає вибіркове застосування покращення лише в тих випадках, коли зображення містить регіони, здатні принести користь для подальшого розпізнавання. Оскільки SR підсилює локальні ознаки, вона є ефективною тоді, коли в кадрі присутні структури, які модель може використати. Для визначення таких регіонів можуть застосовуватися теплові карти, що відображають, які частини зображення модель вважає важливими у своїх внутрішніх ознаках.

Для формування теплових карт широко застосовуються методи Grad-CAM та Eigen-CAM:

1. Grad-CAM формує теплову карту на основі градієнтів виходу моделі за певним класом щодо активацій вибраного шару [1]. Підхід узгоджує середні значення градієнтів з каналами активацій, створюючи карту важливості, яка відображає регіони, що найбільше впливають на прогноз моделі. Оскільки такий механізм напряду пов'язаний із конкретним класом, тепла карта є класово-залежною і показує лише ті ділянки, які модель використовує саме для відповідного передбачення. Якість карти залежить від стабільності градієнтного сигналу. Якщо модель не впевнена у прогнозі або не розпізнає об'єкт, карта стає фрагментованою або слабо вираженою. Крім того, обчислення Grad-CAM потребує виконання зворотного проходу, що підвищує обчислювальну вартість і ускладнює застосування методу в режимах з жорсткими обмеженнями продуктивності. Підхід добре підходить для аналізу класових ознак та пояснення рішень моделі, однак не є ефективним для оцінки загальної інформативності кадру, оскільки повністю залежить від класових прогнозів.

2. Eigen-CAM використовує активації проміжного шару моделі як матрицю ознак та визначає їх найбільш характерний напрямок зміни за допомогою методу головних компонент [2]. Теплова карта формується

на основі першої головної компоненти, що відображає домінуючу структуру активацій без прив'язки до класу. Такий підхід є класово-незалежним і формує карту важливості, що показує інформативність зображення загалом, а не його зв'язок із певним прогнозом. Eigen-CAM залишається стабільним навіть на кадрах низької якості, оскільки не потребує чіткого класового сигналу і не покладається на градієнти. Метод працює на основі прямого проходу моделі, що забезпечує низьку обчислювальну вартість і робить його придатним для використання в режимах реального часу. Оскільки теплова карта відображає загальну інформативність зображення, Eigen-CAM може слугувати надійним індикатором для вирішення, чи варто застосовувати суперрезолюцію, тобто чи містить кадр ознаки, підсилення яких покращує розпізнавання. Відсутність класової прив'язки може бути недоліком у завданнях, де необхідно аналізувати важливість саме конкретного класу, проте в контексті адаптивної суперрезолюції це не є критичним.

Порівняння двох методів показує, що Grad-CAM та Eigen-CAM суттєво відрізняються за принципом формування теплових карт і характером інформації, яку вони відображають. Grad-CAM є класово-залежним і демонструє максимальну ефективність тоді, коли модель має стабільний прогноз, що робить його корисним для аналізу класових ознак, але обмежує у випадках, коли необхідно оцінити загальну інформативність. Eigen-CAM, натомість, формує класово-незалежну карту важливості та зберігає стійкість на кадрах низької якості.

З огляду на ці властивості, Eigen-CAM є більш практичним для адаптивної суперрезолюції, де потрібно визначити, чи має кадр структури, підсилення яких може покращити результативність детекції. Застосування такого підходу забезпечує більш точне визначення інформативності та дає змогу раціональніше використовувати суперрезолюцію, що загалом підвищує ефективність детекції та якість обробки зображень з БПЛА.

Список використаних джерел:

1. Selvaraju R. R., Cogswell M., Das A., Vedantam R., Parikh D., Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization // Proceedings of the IEEE International Conference on Computer Vision (ICCV 2017). – Венеція, 2017. – С. 618–626.
2. Muhammad M. B., Yeasin M. Eigen-CAM: Class Activation Map using Principal Components // Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN). – Глазго, 2020. – С. 1–7.