

Abstract

This dissertation focuses on the problem of tiny object detection in high-resolution images, with applications in real-time systems such as autonomous robots and UAVs. Existing methods often require significant computational resources, are too slow or achieve poor detection quality, when objects are very small. To address these limitations, a modular framework (SegTrackDetect) is developed. It combines a lightweight local detector, that operates on selected full resolution areas of the input image, with three supporting modules: ROI Estimation based on semantic segmentation, ROI Prediction using object tracking, and Global Filtering with two novel algorithms to reduce redundant partial detections. Together, these components allow selective high-resolution inference on relevant image regions, improving speed, detection quality, and computational efficiency.

The first component, the ROI Estimation Module, uses semantic segmentation to guide the placement of detection windows. By aligning window selection with meaningful image content, this approach reduces redundant computation while retaining important contextual information, offering a more efficient alternative to uniform sliding-window techniques. The second component, the ROI Prediction Module, integrates object tracking to recover regions that may be missed by the Estimator because of its low input resolution. By exploiting temporal continuity, it ensures that small and difficult-to-detect objects remain visible to the system, while at the same time lowering the computational cost of the segmentation branch. The third component, the Global Filtering Module, addresses one of the core limitations of window-based methods: partial detections in overlapping regions of the detection windows. It introduces two novel algorithms – Overlapping Box Suppression (OBS), which favors complete detections, and Overlapping Box Merging (OBM), which combines multiple partial detections into a full one. Together, they provide a strong alternative to standard Non-Maximum Suppression and significantly improve quality by lowering both false positive and false negative detections.

Through the integration of these three modules, SegTrackDetect enables selective full-resolution inference on relevant subregions of high-resolution images. This modular design allows balancing detection quality, computational efficiency, outperforming traditional sliding-window and several state-of-the-art approaches while remaining lightweight enough for real-time deployment. The dissertation demonstrates that segmentation, tracking, and advanced filtering strategies can be combined into customizable framework for efficient tiny object detection and provides an open-source implementation for real-world applications.

Streszczenie

Niniejsza rozprawa dotyczy problemu detekcji bardzo małych obiektów na obrazach wysokiej rozdzielczości, ze szczególnym uwzględnieniem zastosowań wymagających działania w czasie rzeczywistym na robotach autonomicznych czy dronach. Istniejące metody często wymagają znacznych zasobów obliczeniowych, działają zbyt wolno lub charakteryzują się niską jakością detekcji, zwłaszcza gdy obiekty są bardzo małe. Aby przezwyciężyć te ograniczenia opracowano wielomodułowy system detekcji SegTrackDetect. System ten łączy lekki lokalny detektor, działający na wybranych obszarach obrazu w pełnej rozdzielczości, z trzema modułami wspierającymi: Estymacją ROI opartą na segmentacji semantycznej, Predykcją ROI wykorzystującą śledzenie obiektów oraz Filtrowaniem Globalnym, które obejmuje dwa nowe algorytmy redukujące problem częściowych detekcji. Połączenie tych modułów umożliwia selektywną analizę istotnych obszarów obrazu w wysokiej rozdzielczości, poprawiając czas działania, jakość detekcji i efektywność obliczeniową.

Pierwszy moduł, Estymacja ROI, wykorzystuje segmentację semantyczną do wyznaczania położenia okien detekcji. Dopasowując obszary do istotnych fragmentów obrazu, zmniejsza on nadmiarowe obliczenia, jednocześnie zachowując kontekst, co stanowi bardziej efektywną alternatywę dla prostych metod wykorzystujących okno przesuwne. Drugi moduł, Predykcja ROI, integruje śledzenie obiektów w celu przywrócenia regionów pominiętych przez estymator z powodu jego niskiej rozdzielczości wejściowej. Przez wykorzystanie ciągłości danych wejściowych, zapewnia on widoczność bardzo małych i trudnych do wykrycia obiektów, przy jednoczesnym obniżeniu kosztów obliczeniowych segmentacji. Trzeci moduł, Filtrowanie Globalne, rozwiązuje jedno z głównych ograniczeń metod opartych na oknach: częściowe detekcje we wspólnych obszarach okien detekcji. Wprowadza on dwa nowe algorytmy: Overlapping Box Suppression (OBS), który zachowuje pełne detekcje, oraz Overlapping Box Merging (OBM), który łączy fragmentaryczne wyniki w jedną spójną detekcję. Razem stanowią one alternatywę dla klasycznego Non-Maximum Suppression, poprawiając jakość poprzez ograniczenie zarówno fałszywie pozytywnych częściowych detekcji, jak i pominiętych obiektów.

Integracja tych trzech modułów sprawia, że SegTrackDetect umożliwia selektywną analizę w pełnej rozdzielczości tylko na istotnych podobszarach obrazu. Modułarna konstrukcja pozwala zachować równowagę między jakością detekcji a efektywnością obliczeniową, przewyższając tradycyjne metody sliding-window oraz szereg najnowocześniejszych metod z literatury, a jednocześnie pozostając wystarczająco lekką do zastosowań w czasie rzeczywistym. Wykazano, że poprzez połączenie segmentacji, śledzenia obiektów oraz zaawansowanych strategii filtrowania uzyskano elastyczną i wydajną metodę detekcji bardzo małych obiektów, która została publicznie udostępniona umożliwiając zastosowanie w rzeczywistych problemach zarówno naukowych jak i wdrożeniowych.