

Tecniche di apprendimento profondo per riconoscimento di oggetti in video termici

Candidato: Federico Schipani federico.schipani@stud.unifi.it

Relatore: Marco Bertini marco.bertini@unifi.it

Abstract

Il *Machine Learning*, detto anche Apprendimento Automatico, nasce come diramazione dell'intelligenza artificiale e trova campi di applicazione sempre maggiori nel mondo contemporaneo. Il trend attuale prevede l'utilizzo estensivo di Reti Neurali artificiali, ovvero dei modelli matematici che per molti versi cercano di ricalcare la funzionalità di un cervello umano allo scopo di sostituirlo per quei compiti che possono essere definiti *ripetitivi*.

Scopo di questo lavoro di tesi è stato migliorare la rilevazione di oggetti all'interno di immagini termiche. Per realizzare ciò si è partiti dalla rete neurale RetinaNet a cui sono state applicate alcune tecniche per ottenere prestazioni sempre maggiori. I dataset a nostra disposizione sono il [KAIST Multispectral Pedestrian Dataset \(KAIST MPD\)](#) ed il dataset di FLIR, su cui sono state considerate solamente le annotazioni riguardanti oggetti come pedoni, ciclisti e veicoli. Dopo un iniziale fase di analisi sono iniziati i primi esperimenti per valutare le performance di detection dove è stato evidenziato come sia possibile ottenere buoni risultati attraverso l'utilizzo di dataset ampi. Successivamente l'interesse è virato sul miglioramento della detection in un sottoinsieme del dataset [KAIST MPD](#) annotato con le automobili.

Nella fase successiva è stata fissata una base di partenza data dall'addestramento di RetinaNet sul dataset di FLIR. Tramite una operazione detta di *fine tuning* è stato raggiunto un incremento complessivo della [Mean Average Precision \(mAP\)](#) su [KAIST MPD](#) del 25%. Sono state usate tecniche di *data augmentation* come *AutoAugment* e *RandAugment* con lo scopo di aumentare la [mAP](#) ottenendo risultati che migliorano di una percentuale che varia dal 2% al 4%.

Sono state utilizzate delle [Generative Adversarial Network \(GAN\)](#) addestrate per generare un dataset termico sintetico partendo dalle immagini RGB di [KAIST MPD](#). La serie di esperimenti effettuati tramite l'utilizzo di questo dataset ha evidenziato come un dataset artificiale può aiutare in quelle situazioni critiche in cui si ha una totale mancanza di dati reali su cui effettuare una fase di addestramento.

Infine sono stati condotti esperimenti su video termici girati sul circuito di test di [Rete Ferroviaria Italiana \(RFI\)](#) di San Donato a Bologna valutando le performance di *Transfer Learning*. Non essendo annotati non è stato possibile svolgere un training accurato o calcolare delle metriche; nonostante ciò l'osservazione dei risultati raggiunti mostra come questi siano più che accettabili, grazie anche allo sviluppo di un algoritmo per migliorare la qualità delle rilevazioni basato sull'indice di Jaccard.

Deep Learning Techniques for Object Detection in Thermal Videos

Candidate: Federico Schipani federico.schipani@stud.unifi.it

Supervisor: Marco Bertini marco.bertini@unifi.it

Abstract

Machine Learning was born as a branch of Artificial Intelligence and nowadays it has several increasing fields of applications . The current trend is to heavily rely on Artificial Neural Networks. An Artificial Neural Network is a mathematical model that tries to mimic some functionality of the human brain to replace it in repetitive tasks.

The objective of this master thesis is to enhance object detections inside thermal images. To do so, we started from the neural network RetinaNet, in which we applied some refinements to obtain better performances. The used datasets are the [KAIST Multispectral Pedestrian Dataset \(KAIST MPD\)](#) and FLIR, where we only considered the annotations about objects like people, cyclists and cars. After an initial phase of analysis, we started the first experiment to assess the detection performances, where has been highlighted the possibility to obtain satisfying results through the use of a wide dataset. We moved then our focus on the improvement of the vehicles detection in the [KAIST MPD](#) dataset that has been annotated with cars.

Afterwards, it has been fixed a baseline coming from the training of RetinaNet on the FLIR dataset. Through a fine tuning operation, it has been reached a 25% increment of the [Mean Average Precision \(mAP\)](#) on [KAIST MPD](#). Data augmentation techniques, such as AutoAugment and RandAument, have been applied to further improve the [mAP](#). A small improvement has been registered for the data augmentation, increasing the overall [mAP](#) by 2 – 4%. Many [Generative Adversarial Network \(GAN\)](#) have been used, trained to generate a synthetic thermic dataset, beginning from RGB images of [KAIST MPD](#). Those series of experiments exposed how such kind of dataset helps in critical situations where real data are missing for the training phase.

At the end, experiments have been performed on thermic videos, shot on the test circuit of [Rete Ferroviaria Italiana \(RFI\)](#) in San Donato in Bologna to evaluate the Transfer Learning performances. Those videos aren't annotated so has not been possible to have an accurate and specific training or to calculate any kind of metric. However, the achieved results are positive, thanks also to the development of an algorithm that improves the quality of the detections based on Jaccard index.