Aggregated Residual Convolutional Neural Network for Multi-label Pixelwise Classification of Geospatial Features

Timothy Forbes[♦], Yuanlie He[*], Sudhir Mudur[♦], Charalambos Poullis[♦]

[♦]Concordia University
[*]Guangdong University of Technology

Object classification is one of the many holy grails in computer vision and as such has resulted in a very large number of algorithms being proposed already. Specifically in recent years there has been considerable progress in this area primarily due to the increased efficiency and accessibility of deep learning techniques. In fact, for single-label object classification [i.e. only one object present in the image] the state-of-the-art techniques employ deep neural networks and are reporting very close to human-like performance.

There are specialized applications in which single-label object-level classification will not suffice; for example in cases where the image contains multiple intertwined objects of different labels. In this paper, we address the complex problem of multi-label pixelwise classification. We present our distinct solution based on an aggregated residual convolutional neural network (ARCNN) for performing multi-label pixelwise classification and its application to large-scale urban reconstruction. A supervised learning approach is followed for training a 10 block ARCNN using both multispectral data from satellite images and depth. An empirical study has been conducted to determine some of the hyperparameters of the ARCNN for improved performance. The trained network yields six pixelwise classifications for new images. A softmax classifier is used to map the six pixelwise classifications into a unique label per pixel. Lastly, we refine boundary pixel labels using graph-cuts for maximum a-posteriori (MAP) estimation with Markov Random Field (MRF) priors. The results, pixels with object type and depth, simplify reconstruction of different objects in large-scale urban areas. The proposed approach has been extensively tested and the results are reported.