**Fatema Tuz Zohra**

**A Review on Real-Time Human Detection as an Edge Service Enabled by a Lightweight CNN**

In this paper, the contributors have mainly focused on Edge Computing using Depthwise Separable Convolutional Network to introduce Lightweight Convolutional Neural Network (L-CNN) to do human object detection in real time [1]. The authors have carried out smart surveillance and have used Raspberry Pi 3 board. The L-CNN algorithm can process from 1.79 to 2.06 frames per second (FPS). The reason for choosing human detection instead of example cars was that, humans walk much slowly and it was possible to cover that in such less memory space. Only humans who appear newly in the frame are detected as the ones who have already been detected are being tracked by a Tracker algorithm and the objects who walk out are automatically deleted. Based on this idea an FPS of around 2 is justified even more. This paper also compared between the results of L-CNN and other algorithms using a Raspberry PI 3 model B device as the selected edge computing node.

Smart Surveillance as an Edge Service is considered to be able to minimize communication delays and message overhead. It is also believed that decision can me made almost instantly in real time [2], [3]. Results with L-CNN has been compared with powerful SSD GoogleNet and two algorithms have been used to draw comparison between them. Algorithms used are Haar-Cascade and HOG + SVM, results with L-CNN shows that it is quite promising and can be used for real time smart surveillance. The writers researched GoogleNet [4] and Microsoft ResNet [5] which can classify up to thousand objects with human but they found that these were not able to find any deep learning network that was specifically designed for human detection [1].

Experiment was conducted as aforementioned with Raspberry PI 3 Model B with ARMv7 1.2 GHz processor and 1 GB of RAM is used as an edge device. GoogleNet and L-CNN have the same training database and memory profiler application was used for recording memory utilization. HAAR- Cascade produced the fastest algorithm (1.82 FPS, 76.9% CPU, 111.6 MB memory, 26.3 FPR) and L-CNN did the second fastest (1.79 FPS, 75.7% CPU, 122.5 MB memory, 6.6 FPR), which was very close to the best in terms of FPS. L-CNN did much better in terms of false positive rate. Different CNN architectures have been compared with L-CNN and it was concluded that a lot more space was needed and it is not possible to do with that in an edge device. Overall L-CNN provided valid and commendable results with edge devices, L-CNN could also handle more complex situations where human objects were not completely in the frame which shows just how promising it can be for the fiuture of edge computing.

**References:**

[1] Seyed Yahya Nikouei, Yu Chen, Sejun Song, Ronghua Xu, Baek-Young Choi, Timothy R. Faughnan, “Real-Time Human Detection as an Edge Service Enabled by a Lightweight CNN”, 2018 IEEE International Conference on Edge Computing (EDGE). IEEE, 2018, pp. 125-129.

[2] N. Chen, Y. Chen, E. Blasch, H. Ling, Y. You, and X. Ye, “Enabling smart urban surveillance at the edge,” in 2017 IEEE International Conference on Smart Cloud (SmartCloud). IEEE, 2017, pp. 109–119.

[3] N. Chen, Y. Chen, S. Song, C.-T. Huang, and X. Ye, “Smart urban surveillance using fog computing,” in Edge Computing (SEC), IEEE/ACM Symposium on. IEEE, 2016, pp. 95–96.

[4] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.

[5] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.