INTELLIGENT TRANSPORTATION SYSTEMS WITH NETWORK MONITORING AND PASSENGER RECOGNITION


Chew Vee Jeng


Tham Mau-Luen


Chang Yoong Choon

The expansion of Internet of Things (IoT) networks and the evolution of machine learning technologies have resulted in overall smart city frameworks, including intelligent transport systems with network monitoring and passenger recognition.

Recent years have seen the growing popularity of Intelligent Transportation Systems (ITS) owing to the evolution of wireless communications and machine learning. 5G technology has also pushed the boundaries of achievable transmission rates.

According to International Telecommunication Union (ITU) [1], 5G capabilities can reach up to 20 Gbps for peak data rate and machines can learn and subsequently execute human tasks, such as speech translation and object recognition, in real time. For instance, the average time per inference of image classification is less than one second.

Transportation systems equipped with machine learning and networking, fall into the scope of IoT. The role of IoT is to connect all things together and allow the sharing of data. For example, with IoT, transportation authorities will be able to track the GPS coordinates of each vehicle and monitor its movements.

Bus service is likely to be main bottleneck of the entire public transportation system since it serves as the bridge to the first and last-mile connectivity. Excellent bus service quality can enhance commuter experience and subsequently boost overall public transport ridership. This motivates us to consider IoT bus systems with two features: Network monitoring and passenger recognition.

Network monitoring enables bus operators to identify possible network blind spots which can pose a significant challenge to bus fleet management system in terms of capturing accurate real-time bus GPS data. In turn, this may affect user experience as passengers rely on the estimated time arrival calculated from unreliable GPS data.

Apart from that, passenger recognition is also vital for urban mobility planning as gender and age are some of the decisive variables which will influence travel patterns and transit ridership. To this end, we have developed a prototype which monitors the mobile signal strength, measures GPS data, counts the on-board passengers and predicts their age and gender. Collected data is uploaded to backend and visualised on a smart dashboard for further analysis.

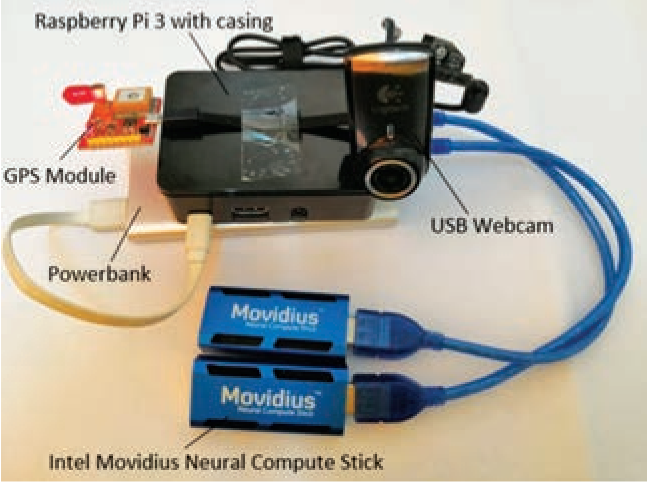

Figure 1: Developed prototype

METHODS

Figure 1 depicts the embedded system, which consists of Raspberry Pi 3 as the host, with Movidius Neural Compute Sticks (NCSs), USB webcam and GPS sensor as peripherals.

Passenger Counting with Gender & Age Recognition: We leveraged on deep learning approach to realise passenger counting with gender and age recognition. We employed 3 pre-trained models – MobileNet-SSD [3], AgeNet and GenderNet [4] – all of which were converted into graph file for Movidius NCS compatibility using Neural Compute Software Development Kit (NCSDK) [5] with Caffe framework [6].

The graph conversion enables inference acceleration of these three deep learning models. MobileNet-SSD can

Figure 2: Test-bed setup

detect person accurately, but it cannot discern between individual passengers. To circumvent this issue, we implemented centroid tracking algorithm as in [7] to track the same person. Once the same person was detected, the passenger count would be incremented. The image was then fed into AgeNet and GenderNet for age and gender analysis. The process was repeated for every single person boarding the bus and the collected data uploaded to a backend server. Figure 2 shows the test-bed setup where we placed the prototype inside a bus.

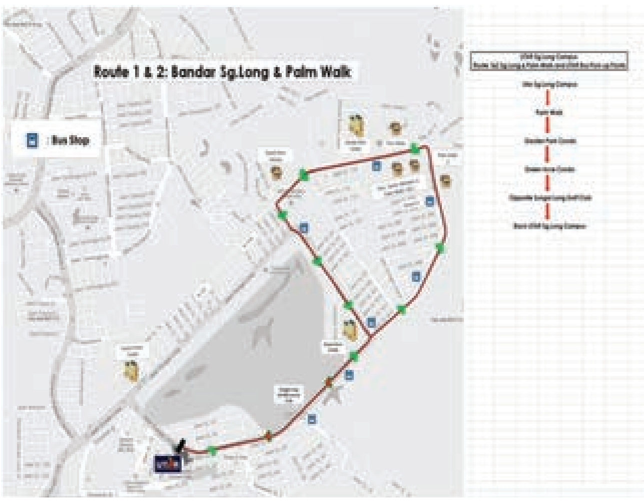

Figure 3: UTAR bus route

Network Monitoring: Download speed, upload speed and ping were measured periodically as the bus travelled across the Bandar Sungai Long area (Figure 3). Corresponding GPS coordinates were recorded in order to discover potential network blind spots. Similar to Section II-A, the collected data was uploaded to the same backend server.

RESULTS AND DISCUSSIONS

For the performance of deep learning models, we adopted the commonly used precision and recall as evaluation metrics. Precision means how much of the output predicted as positive, is correct. Recall indicates that among all of the actual positive results, how many are predicted as positive. Mathematically, precision and recall can be expressed

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive+False Positive}} \tag{1}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive+False Negative}} \tag{2}$$

as respectively. The reason why we leveraged the offline dataset for performance evaluation was that the sample size of the collected data inside bus was too small to obtain statistically meaningful results. The actual test is visualised in Figure 4.

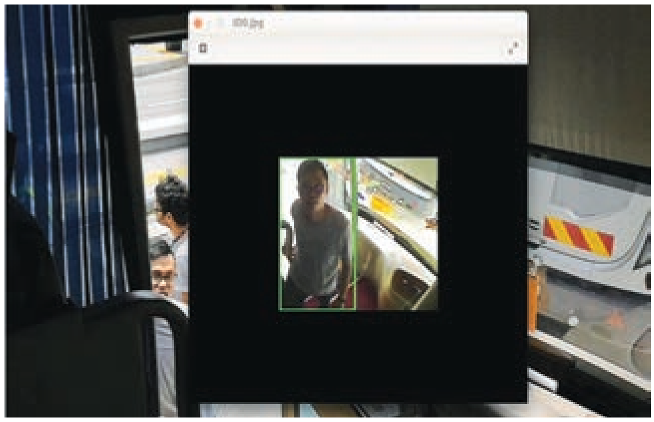

Figure 4: Passenger detection

Passenger Counting: The test images, extracting from INRIA Person Dataset [8], consisted of 1,391 images with 1,129 human (positive) and 262 non-human (negative) images. Figure 5 shows the tradeoff between precision and recall for different thresholds. It can be observed that the solution yields high precision and high recall, with almost all results labelled correctly.

Gender Prediction: The test images, extracting "Looking at People Workshop" [9], consisted of 703 images (471 male, 232 female). Different from Section III-A, we employed confusion matrix as the evaluation metric as gender prediction belonged to the multi-class classifier. Table 1 displays the computed value. The precision-recall pairs (in unit of %) for MALE and FEMALE are (86.28, 78.77) and (63.37, 74.57), respectively.

Table 1: Confusion matrix of gender prediction on test images

n = 703	Predicted: MALE	Predicted: FEMALE
Actual: MALE	371 (True Male)	100 (False Female)
Actual: FEMALE	59 (False Male)	173 (True Female)

Age Prediction: The test images, extracting from LFW datasets [10], consisted of 566 images with 1 (age 8-12), 15 (age 15-20), 154 (age 25-32), 123 (age 38-43), 211 (age 48-53) and 62 (age 60-100) images. Table 2 displays the computed value. The precision-recall pairs (in unit of %) for each age category are (0, 100), (0, 0), (33.49, 50.34), (27.78, 0.04), (36.36, 0.02), and (28.57, 9.67), respectively. It can be observed that the age prediction performs poorer than gender prediction in terms of both precision and recall.

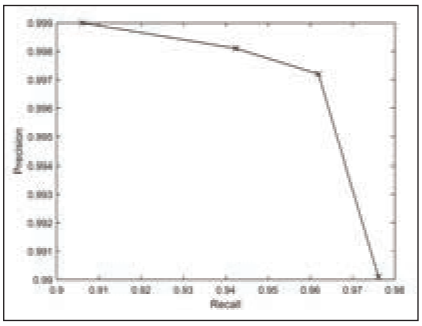

Figure 5: Precision vs. Recall for passenger detection

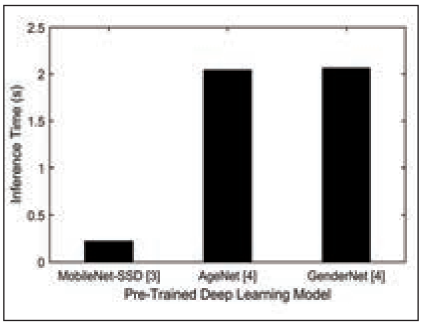

Figure 6: Inference time of 3 deep learning models

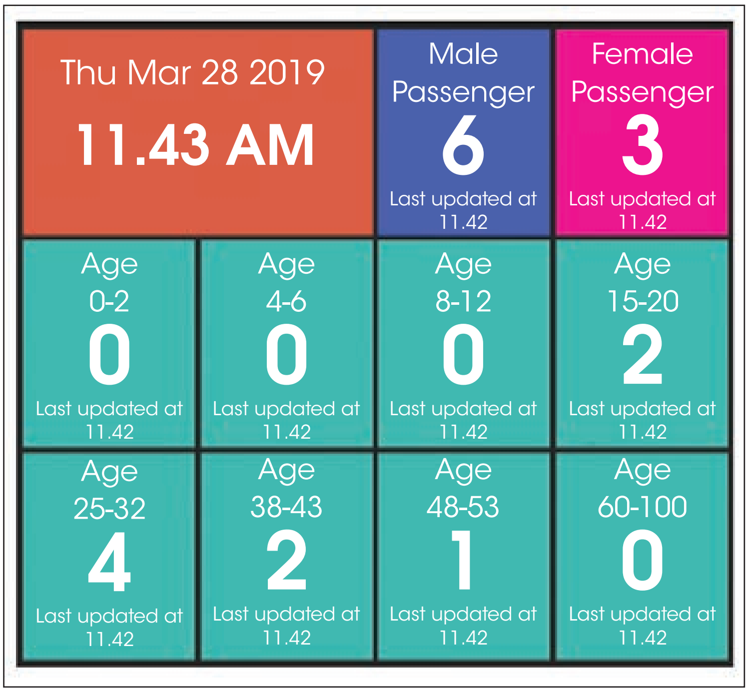

Figure 7: Dashboard on internet browser

One possible reason is that the neural network model is using large-scale non-Asian face dataset during training phase.

Inference Time: Short inference time is another important metric that demonstrates whether the low-power embedded system can be deployed in practical scenarios. Figure 6 shows the inference time needed for Mobilenet-SSD, AgeNet and GenderNet. It can be seen that Mobilenet-SSD performs 9 times faster than the other two models. This is because its neural network structure is lightweight, so it can be comfortably run in real time on resource-constrained devices. Note that we adopted Movidius NCS to amplify the concept of edge computing, where each bus was expected to handle the real-time passenger counting. Nevertheless, some decisions will require peripheral, cloud-based information. For example, real-time seating capacity within each bus can be shared to cloud systems for navigation decisions and route-planning. As a benchmark, Mobilenet-SSD without Movidius NCS is 38 times slower than Mobilenet-SSD with Movidius NCS.

Smart Dashboard: After the local data collection and analysis, the number of passengers as well as their age and gender, were displayed on a dashboard. Operators could access the information through an Internet browser from their mobile devices (Figure 7).

```
Date,Time,Ping (ms),Download (Mbit/s),Upload (Mbit/s)
 + nan,0.0,0.0
04/01/19,12:12,55.068,5.19,3.41
04/01/19,12:13,52.24,8.23,3.32
2019-04-01T04:15:05.000Z + 2019-04-01T04:15:05.000Z,101.805353333,3.046046667
04/01/19,12:15,69.427,5.66,1.55
2019-04-01T04:18:06.000Z + 2019-04-01T04:18:06.000Z,101.805365,3.045983333
04/01/19,12:18,54.68,6.83,3.04
2019-04-01T04:21:06.000Z + 2019-04-01T04:21:06.000Z,101.805201667,3.046153333
04/01/19,12:21,59.009,3.92,2.61
2019-04-01T04:24:05.000Z + 2019-04-01T04:24:05.000Z,101.80503,3.046165
04/01/19,12:24,49.46,9.76,2.30
2019-04-01T04:27:05.000Z + 2019-04-01T04:27:05.000Z,101.805063333,3.046071667
04/01/19,12:28,54.312,6.59,2.58
```

Figure 8: Network and GPS measurements

Network Monitoring: Figure 8 displays the network and GPS measurement results. The average values of ping, download and upload were 55.6 ms, 6.218 Mbps and 3.179 Mbps. Mobile connectivity dropped significantly when the bus was approaching the GPS coordinate of (3.04624, 101.8051). It recorded the values of 143.566 ms, 3.46 Mbps and 0.28 Mbps. One possible reason was that the bus was in the vicinity of high rise building which might have triggered fading.

CONCLUSION

A prototype has been developed to achieve intelligent transport solutions. The IoT solution counts passengers who board the bus and predicts their age and gender. The demographic data is visualised and transmitted to a remote server for future transportation planning. Concurrently, the device is able to perform drive testing automatically. By logging each GPS location, the device can discover network blind spots and inform transportation authorities. Ultimately, the data will offer new insights for the nation's long-term economic, social and environmental sustainability.

In future, we plan to use LoRa for more precise GPS coordinates of a vehicle. Besides that, the accuracy of age classification can be enhanced with more relevant training dataset. ■

REFERENCES

[1] Series, M., 2015. IMT Vision–Framework and overall objectives of the future development of IMT for 2020 and beyond. Recommendation ITU, pp.2083-0.

[2] Pena, D., Forembski, A., Xu, X. and Moloney, D., 2017. Benchmarking of CNNs for low-cost, low-power robotics applications. In RSS 2017 Workshop: New Frontier for Deep Learning in Robotics.

[3] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. and Adam, H., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04863

[4] Levi, G. and Hassner, T., 2015. Age and gender classification using convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 34-42).

[5] Intel. Deploying Your Customized Caffe Models on Intel® Movidius™ Neural Compute Stick. Retrieved April 6, 2019, from https://movidius.github.io/blog/deploying-custom-caffe-models/

[6] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S. and Darrell, T., 2014, November. Caffe: Convolutional architecture for fast feature embedding. In Proceedings of the 22nd ACM international conference on Multimedia (pp. 675-678). ACM.

[7] Adrian Rosebruck, 2018, Simple object tracking with OpenCV [Online]. Available at: https://www.pyimagesearch.com/2018/07/23/simple-object-tracking-with-opencv/ [Accessed: 6 April 2019].

[8] Dalal, N. and Triggs, B., 2005, June. Histograms of oriented gradients for human detection. In international Conference on computer vision & Pattern Recognition (CVPR'05) (Vol. 1, pp. 886-893). IEEE Computer Society.

[9] Rothe, R., Timofte, R. and Van Gool, L., 2015. Dex: Deep expectation of apparent age from a single image. In Proceedings of the IEEE International Conference on Computer Vision Workshops (pp. 10-15).

[10] Huang, G.B., Mattar, M., Berg, T. and Learned-Miller, E., 2008, October. Labeled faces in the wild: A database forstudying face recognition in unconstrained environments. In Workshop on faces in'Real-Life'Images: detection, alignment, and recognition.

Authors' Biodata

***Chew Vee Jeng** is a Business & Integration arch analyst in a global IT and management consulting firm and a student at the Department of Electrical and Electronics Engineering, Universiti Tunku Abdul Rahman.*

***Tham Mau-Luen** is assistant professor in the Department of Electrical and Electronic Engineering, Universiti Tunku Abdul Rahman.*

***Chang Yoong Choon** is associate professor at the Lee Kong Chian Faculty of Engineering & Science, Universiti Tunku Abdul Rahman.*

CONGRATULATIONS

Congratulations to **Tan Sri Dato' Dr Francis Yeoh Sock Ping** for his conferment of the honorary British award of Knight Commander of The Most Excellent Order of the British Empire (KBE) by Her Majesty Queen Elizabeth II for services to UK - Malaysian bilateral relations at Clarence House, London, on 13 October 2019.