Survey Paper on Real Time Vehicle Detection on Edge for Smart Traffic Management

Shashikala J[1], Sagar H[2], S Tejas Narayan[3], Sanjith R[4], Rohit B C[5]

Department of Electronics and Communication, BMS Institute of Technology and Management
Avalahalli Yelahanka Bengaluru

***Abstract*— This paper describes the various technologies which are employed in Real Time Vehicle Detection using an edge device Intel Movidius Neural Computing Stick. Vehicle detection is a very important task which enables the smart traffic management here moving from cloud centric to edge centric approach greatly minimizes the bandwidth and latency issues. The MobileNet-SSD an ImageNet classification model is a light weight architecture built on a Caffe framework ensures better performance finally implementing the model on low cost high performance oriented Raspberry Pi will greatly minimizes the cost and easy deployment of model thus Real Time Vehicle Detection is achieved.**

***Keywords*— Neural Computing Stick (NCS), MobileNet-SSD, MYRIAD architecture, SHAVE VLIW vector processor, LEON Microprocessor, Raspberry PI.**

I. Introduction

A Real Time Vehicle Detection involves finding whether there is vehicle present or not secondly which type of vehicle is present and how many vehicles are present. Basically, vehicle presence needs to be detected after detecting a vehicle it has to be classified. Classification is the main part which means what type of vehicle it is (car, bus, bike, etc.). Smart Traffic Management aims at the avoidance of traffic on roads especially on highways because where manual governance and management is difficult. Implementation of traffic surveillance camera finds primary application in traffic monitoring through which management will become easier. If those cameras are enabled by the modern technology enables the Smart Traffic Management. Internet of Things (IoT) concept demo where Artificial Intelligence (AI) inferencing happens on the device and in near real time is the key feature employed at camera enables it to detect vehicles in real time.

As a human it is easy for him to detect and classify the vehicle by simply seeing it. But while thinking from surveillance camera and it is not that much easy. This is the main objective addressed and implemented in this project. Usually vehicle detection techniques are dependent on internet and they utilise more power for running. Here, in this project implementation difficulties have been resolved the vehicle detection technology will be independent of the internet to and uses light weight framework. which can be installed in the edge device which involves a computational device in this case Intel Movidius Neural Computing Stick which is interfaced to Raspberry Pi and Surveillance camera can be connected to this edge device.

The Edge device (Raspberry Pi, Intel Movidius and a camera) requires no internet connection for vehicle detection utilizes internet only for connection to the cloud for real time data transfer. Instead of sending the images or video to the cloud for further analysis, the detection happens on the device and only the results (i.e. pure text) are sent to the cloud for i.e. statistical and visualization purposes. Such approach brings numerous benefits like: solves latency and communication bandwidth related concerns.

In the first stage classification is carried out using MobileNet-SSD framework and then uploaded a pre-trained model onto the device i.e., Intel Movidius Neural Computing Stick. Which intern mounted to Raspberry pi board together acts has an edge device. The video feed is given to the edge device through a digital camera in Full-HD, convert that to around 448x448 before it lands in the neural network. Average performance here is of ~5-6fps as it requires a lot of computing power to analyse frames in near-real-time mode on a relatively constrained hardware architecture (many objects being detected). Later stage involves connecting edge device to cloud and it requires a analytics script which is to be developed and addressed to statistical and visualization purposes. Results from the edge device is sent to the cloud and Exploratory analysis is carried out. Finally, Real Time Vehicle Detection is achieved which enables the surveillance camera becoming a part of edge device this helps in Smart Traffic Management.

II. Literature Survey

[1] For efficient performance of traffic data and traffic control, the traffic data of multiple vehicle types is required. With the data collected the type and speed of every vehicle is determined, which is used to arrive at the flows and mean speeds of different vehicle types. In the current paper a colour image-based adaptive background subtraction is proposed to obtain more accurate vehicle objects, and also a series of processes like shadow removal and setting road detection regions are used to improve the system robustness. The results obtained from the classification system for these test videos were compared to the ground truth table, to conclude that the accuracy of vehicle counting was 97.4%. The errors of vehicle classification without data fusion was 18.8% and the errors of vehicle classification using data fusion was 8.3%. It was concluded that data fusion can be used to reduce the errors caused due to vehicular occlusion.

[2] Convolutional neural networks have become ubiquitous in computer vision ever since AlexNet popularized deep convolutional neural networks. The extensive experiments on resource and accuracy trade-offs and showed strong performance compared to other popular models on ImageNet classification A class of efficient models called MobileNets for mobile and embedded vision applications.

The MobileNet model is based on depthwise separable convolutions which is a form of factorized convolutions which factorize a standard convolution into a depthwise convolution and a 1×1 convolution called a pointwise convolution.

ImageNet classifications involves lot of classification models but they compromise with resource and accuracy trade-offs. Hence an efficient model is required for ImageNet classification. MobileNets is the efficient model for ImageNet classification. It finds major applications in mobile and embedded vision applications.

[3] Video Analytics plays a key role in the development of smart cities and is used in applications such as crowd computing, activity detection, event classification, traffic counting. Previously used systems incorporated a cloud centric approach, that presented a number of problems such as available bandwidth, real-time responsiveness and personal data privacy issues. To solve these issues video analytics is now done on the EDGE where raw video feeds can be pre-processed at the point of capture while integration and deeper analytics is performed in the cloud.
Smarter cities use IOT technologies to provide a sustainable environment to accommodate the needs of the increasing urban population of tomorrow and preserve natural resources. For the purpose of crowd monitoring it is better to use an edge centric approach than a cloud centric approach. Using edge centric approach in a video-based crowd monitoring application captured frames can be transformed into meaning information (number of people in the scene, main activity) at the device level, close to point-of-capture, enhancing response-time and improving data privacy by pushing only low-bandwidth, anonymous data (numbers) to the cloud for analysis and aggregation.
The early stages smart city applications were using cloud-centric approach, but now a new paradigm has emerged which balances edge and cloud computing. Here video analytics at the edge has been explored using Dell Edge Gateway 5000. The basic image processing which are not CPU-bound is performing equally good on a low and middle range CPU device. They can process more than one frame per second where most CCTVs only capture one image per second which make them suitable for near real-time crowd monitoring. But, with the algorithm complexity increasing, the size of the frame highly impacts the performance.

[4] Testing the performance of state of the art YOLO system and its quantized version on Raspberry Pi device and found that their speed is not eligible for real time use. Proposed a new object detection system with region proposal based on temporal information and reaches nearly 20X speed up and 15X less storage compared to YOLO. Implement a detector version of MobileNet, namely Mobile-Det, by combining MobileNet classifier and Single Shot MultiBox Detector (SSD) frame work SSD framework uses multiple feature layers as classifiers, where each feature map is evaluated by a set of different (aspect ratio) default boxes at each location in a convolutional manner, and each classifier predicts class scores and shape offset relative to the boxes.
The benefit of using SSD framework is evident a unified model which is able to train end-to-end; which do not rely on the reference frame and hence the temporal information, expanding the application scenarios; it is also more accurate in theory. In comparison to the obtained results the YoloV2 occupies 60.5MB and has a detection speed of 0.458fps, whereas the Mobile-Det occupied a space of 27.5 MB and showed a detection speed of 0.712fps.

[5] Vehicle detection means finding or analysing of the vehicles whether there is vehicle present or not. After analysing vehicle classification is another part which means what type of vehicle it is (car, bus, bike, etc.). Objective is to detect and classify vehicles with increased accuracy, speed and reduced processing time. Propose a new device named Intel Movidius Neural Compute Stick which had convolutional Neural network of Caffe framework for vehicle detection to satisfy all the above requirements. This model will be very effectively used in technologies like self-driving cars. Vehicle detection and classification is one of the important parts in autonomous cars. This involves 3 stages: Training: The main goal is to train a model by giving hundreds of thousands of samples of different features and varieties. Inference: Comparing original images with sample images in order to recognize the vehicles and classify them accurately. Deployment: Implementation in daily life for real time applications like traffic management.

[6] Raspberry Pi is a multi-purpose low-cost Advanced Reduced-Instruction-Set-Computer (ARM) processor-based miniature device that can be used as a standalone machine. Raspbian is the provided operating system (OS), but there are various other ARM-Linux OS variants that can be used. This miniature device can be used for robotics, arcade machines, and temperature probing devices. It can also be used for MATLAB applications, and comes in a variety of models with different interfaces for different requirements.

[7] Real-time detection and recognition of objects is a very important task in image processing and computer vision. Here, an embedded powerful technique for real-time object detection and recognition running at high Frames-per-Second (FPS) on an embedded platform like Movidius Neural Compute Stick. This is done by applying a deep learning for computer vision. It includes a single shot detector algorithm with MobileNet architecture trained with caffe framework used with Raspberry Pi 3. The movidius stick is used in correlation with the Raspberry Pi 3 to achieve high FPS. Certain enhancements like default boxes, multi scale features and depth wise separable convolution.
Recent object detection and recognition is mostly established on utilize of convolution neural network. The three main methods for deep learning based object detection include Single Shot Detector (SSD), Faster Region CNN (F-RCNN) and You Only Look Once (YOLO). Here, caffe deep learning library and OpenCV's deep neural network library to perform object detection and recognition. SSD algorithm with MobileNet architecture to obtain a pre-trained model that was trained and tested with real time video with Raspberry Pi 3 and neural compute stick. The Movidius stick is an embedded machine intelligence platform designed for low power devices to achieve high frame rates. To use the NCS for any application Movidius SDK is installed and then Movidius graph files are generated.
Embedded object detection and recognition system use deep learning techniques to detect and recognize objects. The method can identify and recognize objects in a real-time video. Raspberry Pi 3 is used to execute the systems as it is a low-cost embedded platform with exceptional specifications. NCS benefits to build efficient system for embedded vision applications and 3.5 frames per second can be processed using Movidius NCS whereas the system obtains 0.5 frames per second without using Movidius NCS.

[8] A new generation of ultra low power processors have emerged that are aimed primarily at signal processing in mobile computing. The demand for computing power in mobile systems is driven largely by ever more complex applications. This leads to algorithms developed for scientific computing being used increasingly in signal conditioning and emerging fields such as Computer Vision. Here it describes the design and implementation of dense matrix multiplication on the Movidius Myriad architecture and evaluates its performance and energy efficiency, achieving a performance of 8.11GFLOPS at 180Mhz. Considering Myriad's published power dissipation of 350mW, this results in an outstanding

performance/Watt ratio of 23.17GFLOPS/W for a key computational kernel.
Scientific computing techniques have important applications outside the traditional scientific establishment. Linear algebra is used in game physics, Computer Vision and other applications once confined to large scale computing, now increasingly found on mobile systems. And as computing moves onto mobile devices, the need for energy-efficient high-performance platforms will only increase In this field, the Myriad architecture is unique, mainly because it was designed from scratch with the goal of providing specifically that; we believe that Myriad is representative of a new class of ultra-low power processors.

[9] Data level parallelism (DLP) in application is not fixed and varies largely due to the computational characteristics of applications. Processors today use single width SIMD hardware to exploit DLP, but single width SIMD may not be suitable for applications with variable DLP. To counteract this, VLIW processors with multiple vector widths are used to serve applications with varying DLP. One such example of VLIW processor is SHAVE that provides hardware support for native 32 bit and 128 bit wide vector operations. The main aim here is to improve the performance of compiler generated SIMD code by reducing the number of overhead operations and by improving the SIMD hardware architecture. SIMD is adopted in many modern processor architectures targeting embedded, desktop, graphics and supercomputing domains. Application parallelism to be exploited appears in the form of instruction level parallelism. To exploit both the ILP and DLP, very-long instruction-word (VLIW) architecture is one of the most widely used types of processor architecture in many mobile and embedded platforms.

[10] LEON is a 32-bit microprocessor which implements a RISC architecture conforming to the SPARC v8 definition. It is a open source core written in VHDL and can be implemented both on FPGAs and ASICs. The microprocessor used big-endian byte ordering, has 32-bit internal registers ,72 different instructions in 3 different instruction formats and 3 addressing modes. It implements signed and unsigned operations, and has a 7-stage instruction pipeline. SPARC v8 processor defines three main units, integer unit, floating-point unit and a custom co-processor, each one with its own 32-bit internal registers. The LEON uses the AMBA-2.0 AHB bus to connect the main processor with high-speed controller like cache and memory ones and other optional units like the on chip RAM or PCI or Ethernet interfaces. The external memory access for LEON is provided by a programmable memory controller with interfaces to PROM, SRAM, SSRAM, DDR & SDRAM chips, providing also memory mapped I/O operation.

TABLE I
OVERVIEW TABLE

Author	Year	Technology	Description
Shuguang Li, Hongkai Yu, Jingru Zhang, Kaixin Yang, Ran Bin,	2018	Vehicle Detection	A colour image-based adaptive background subtraction is proposed to obtain more accurate vehicle objects
Andrew Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko , Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam,	2017	MobileNet Image classification model	It finds major applications in mobile and embedded vision applications.
Camille Ballas, Mark Marsden, Dian Zhang, Noel E. O'Connor, Suzanne Little	2018	Video Analytics on edge	Edge technology have better performance than cloud technology
Yundong Zhang, Haomin Peng, Pan Hu	2017	Benefit of using SSD framework	SSD framework improves accuracy when used with MobileNet image classification model.
K.Banumathi, B.Bharathi, S.Daisy Deve Priya, P. Gogulalakshmi	2018	Vehicle detection and classification	NCS helps in detect and classify vehicles with increased accuracy, speed and reduced processing time.
Ar Kar Kyaw, Hong Phat Truong, Justin Joseph	2018	Raspberry Pi	This miniature device can be used for robotics, arcade machines, and temperature probing devices.
Nashwan Adnan OTHMAN, Ilhan AYDIN	2018	Movidius Neural Compute Stick	NCS benefits to build efficient system for embedded vision applications
Mircea Horea Ionică, David Gregg	2015	Myriad architecture	Myriad is representative of a new class of ultra-low power processors
Erkan Diken, Martin J. O'Riordan, Roel Jordans, Lech Jozwiak	2015	SHAVE VLIW Processor	Most widely used types of processor architecture in many mobile and embedded platforms.
A.Muñoz, E.Ostúa, M.J.Bellido, A.Millán, J.Juan, D.Guerrero	2008	LEON microprocessor	Connect the main processor with high-speed controller like cache and memory ones

III. CONCLUSIONS

The Movidius NCS can propel the Raspberry Pi to a 6.88x speedup over the standard CPU object detection. That is 6-7 fps is achieved. Thus, achieving the increase in average performance and accuracy of vehicle detection through light weight architecture. Usage of low power utilization edge device overcomes cloud dependency and internet connection for vehicle detection. An energy efficient, more secure, cost efficient and user friendly model which does Real Time Vehicle Detection On Edge For Smart Traffic Management can be used.

REFERENCES

[1] Shuguang Li, Hongkai Yu, Jingru Zhang, Kaixin Yang, Ran Bin, "Video-based traffic data collection system for multiple vehicle types", IET Intelligent Transport System, 2015.

[2] Andrew Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications", Published at Cornell University, 2017.

[3] Camille Ballas, Mark Marsden, Dian Zhang, Noel E. O'Connor, Suzanne Little, "Performance of video processing at the edge for crowd- monitoring applications", IEEE 4th World Forum on Internet of Things (WF-IoT), 2018.

[4] Yundong Zhang, Haomin Peng, Pan Hu, "Towards Real-time Detection and Camera Triggering", Published at Standford University as CS341 Final Report, 2017.

[5] K. Banumathi, B. Bharathi, S.Daisy Deve Priya, P. Gogulalakshmi, "Real Time Vehicle Detection using Movidius Neural Compute Stick", RTICCT-2018 Conference Proceedings, 2018.

[6] Ar Kar Kyaw, Hong Phat Truong, Justin Joseph, "Low-Cost Computing Using Raspberry Pi 2 Model B". Published at Journal of Computers Volume 13, 2018.

[7] Nashwan Adnan OTHMAN, Ilhan AYDIN, "A New Deep Learning Application Based on Movidius NCS for Embedded Object Detection and Recognition", 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2018.

[8] Mircea Horea Ionică, David Gregg, "An analysis of the suitability of the Movidius Myriad architecture for scientific computing", IEEE Micro, Issue 1, Special Issue on Mobile Systems, Jan-Feb 2015.

[9] Erkan Diken, Martin J. O'Riordan, Roel Jordans, Lech Jozwiak, Henk Corporaal, David Moloney, "Mixed-Length SIMD Code Generation for VLIW Architectures with Multiple Native Vector-Widths", IEEE 26th International Conference on Application-specific Systems, Architectures and Processors (ASAP), 2015.

[10] A.Muñoz, E.Ostúa, M.J.Bellido, A.Millán, J.Juan, D.Guerrero, "Building a SoC for industrial applications based on LEON microprocessor and a GNU/Linux distribution", IEEE International Symposium on Industrial Electronics 2008.