Resource allocation for adaptive video streaming with machine learning-based subjective quality of experience

Dr. Abhinav Kumar,
Dept. of EE
KID: 20200207

Our research focuses on utilizing machine learning and resource allocation for improving the quality-of-experience (QoE) for video streaming users over a wireless network. The video streaming in mobile environments is challenging due to the time-varying nature of the wireless channels and is affected by the fluctuating network conditions. Given these factors, it is important that the wireless networks perform careful management and optimal utilization of the available resources such that the video streaming users' demands are met to the best possible extent without degrading their QoE. The Dynamic Adaptive Streaming over HTTP (DASH) standard provides a media delivery framework that allows video users to adapt over different bitrates according to the varying network conditions so that an uninterrupted playback is maintained. However, it often results in a user video quality that keeps varying with time. Moreover, the video playback buffer can run out of the content resulting in rebuffering events. Both time-varying quality and rebuffering affect the perceptual QoE of video users. In this work, we have addressed the problem of understanding and modelling the streaming QoE as perceived subjectively by the users. Streaming QoE modelling is challenging as it is continuous, dynamic, and time-varying in nature.

Hence, the QoE must be evaluated in a continuous manner so that suitable actions can be taken at the network controller such as eNodeB in cellular networks to minimize the QoE degradation of the video users. Various machine learning models like support vector regression QoE (SVR-QoE), C3D time-varying subjective quality (C3D-TVSQ), non-linear state-space QoE (NLSS-QoE), and long short-term memory QoE (LSTM-QoE) based models have been considered.

Given the benefits offered by the QoE-centric methodologies for video streaming in cellular networks, we propose Video Quality-of-experience Aware Resource Allocation (ViQARA), a QoE-centric method for allocating resources to the video users in LTE cellular networks. In ViQARA, the potential of the proposed machine learning-based QoE models are leveraged for resource allocation. A comparison between ViQARA and the conventional throughput based resource allocation indicates that a significant improvement in the average QoE levels, as well as the reduction in the average rebuffering times of the video users, can be achieved with ViQARA.

AI @ Advanced Embedded Systems and IC Design Laboratory

Dr. Amit Acharyya,
Dept. of EE
KID: 20200208

I. Compression Methodology for CNN Targeting Resource--Constrained Edge Devices

In this study we proposed a fragmented Huffman coding methodology, that can be applied to the binary value equivalent of the numeric weights of a neural net model when stored in hardware memory.

Continued...

Huffman compression technique has been revisited by applying it on binary files, from a hardware design perspective, based on multiple bit pattern sequences, to achieve a maximum compression rate of 64 %. This is followed by a compressed hardware memory architecture and the decompression module design which has been synthesized at 500 MHz, using GF 40-nm low power cell library with a nominal voltage of 1.1V achieving a reduction of 62 % dynamic power consumption with a minimal single module decompression time of about 63 microseconds without trading-off accuracy.
[Funding Acknowledgment - DRDO, NPDF (DST, SERB, Govt of India), Redpine Signals Inc.]

II. Development of See-through Armour for Defence

We propose to develop an agile situational awareness prototype model that will employ advanced monocle and imaging technology, providing soldiers with a 360-degree, real-time view outside of their combat vehicles. External optical sensors operating in normal vision or infrared feed imagery to either a helmet-mounted display that synchronizes with head movements and stitches together a complete picture of the battlespace or plugin a tablet display to view 360 degrees to digitally collate, map, and classify various features on the battlefield to track their environment. Better knowledge of surroundings brings increased combat effectiveness and survivability, enabling safer route planning, speedier communication and information sharing, and more rapid targeting. In addition, soldiers can share what they are seeing with other crew members or their commanders to boost response time.

III. Reconfigurable Machine Learning Accelerator

DEEP neural networks (DNNs) are extremely popular and have been adopted to solve problems in a wide variety of fields, including image recognition, semantic segmentation, language translation, and autonomous driving. DNN inference is currently performed on a range of traditional computing systems, including CPUs, field-programmable gate arrays (FPGAs), and GPUs, which provide different

tradeoffs between efficiency, cost, performance, and programmability. CNN algorithms and designs are developing rapidly, however, the low-level operations such as convolution, pooling, activation, etc. remain the same. Hence, we are working for the reconfigurable ML accelerator on which varying shapes and sizes of CNN can be implemented.

The proposed accelerator will be efficient, parameterized and run-time configurable hardware architecture with high-level parallelism and efficient memory designs to support various networks that fit into various FPGA platforms. In other words, we like to design a hardware/software co-designed library to efficiently accelerate an entire CNN on FPGAs. *[Funding acknowledgement - DRDO]*

IV. Intelligent metal corrosion detection based on Acoustic Emission (AE) signals.

Acoustic Emission (AE) signals are sound waves, generated by the rapid release of energy from a localized source within a material when it is stressed. These localized sources of stress wave include different types of defect in metals like cracks, pits, corrosion etc.

Continued…

These stress waves travel through the material and captured by the AE sensors placed over it. The stress wave from the sources mentioned above has some specific unique feature which will be identified and characterized. These real-time characterized AE signals from material (High strength steel) along with simulated AE signals from simulation software like Ansys will be pre-processed and mixed with noise (real and simulated) like different environmental conditions, work as an input to train the neural network. This trained neural network model will be used to predict the type of defect which can classify the defects due to different causes, and ageing of the sample. This customized deep learning model will work efficiently for the specific task as mentioned above. This model will be compact and of less size, so it can work on a mobile platform and require less computation power.

[Funding Acknowledgment – Naval Research Board (NRB) and IMPRINT-II]

V. IoT Based Holistic Prevention and Prediction of Cardiovascular diseases and Assistive Technology

Cardiovascular diseases (CVD) are the leading cause of deaths across the world. Recent years have witnessed a growing interest in developing personalized and non-hospital based care systems to improve cardiac care consequently to reduce morbidity and mortality.

Electrocardiogram (ECG) is the only easy to use diagnosis tool useful for assessment of the probability of cardiac arrhythmias. Clinical CVD diagnosis is generally carried out using the standard 12 lead system.

However, recording all the 12 leads is often difficult, cumbersome and impractical considering the high memory, bandwidth and data transmission time & power in remote health care systems. Moreover, a reduced lead system with 2-3 leads is generally utilized in the telemonitoring application, which is not significant for clinical diagnosis. To bridge this gap, we are researching towards reconstructions methodologies, where a single channel ECG will be recorded from transmission end and all the remaining leads will be reconstructed at the receiver end, thus limiting the bottlenecks mentioned above.

For this reconstruction methodologies, we are using data-driven deep learning approaches using Convolutions neural networks (CNN) and LSTMs (Long short term memory). While convolutions extract the inherent features adaptively, LSTMs are useful to extract the temporal dependency, thus maximizing the performance of the reconstruction. Apart from this ECG to ECG reconstruction, we are also working on Photo plethysmograph (PPG) to ECG reconstruction. PPG is a non-invasive technique which records the blood volume changes happening due to cardiac pumping of the blood using optical techniques and very low-cost technology than for ECG. Moreover, the patient has to wear a minimum of 3 electrodes on the body to get one lead of the ECG where PPG can be recorded from a single sensor placed on the wrist or fingers. With the intuition of PPG correlates with ECG and considering the advantages of PPG over ECG, we started working on PPG to ECG reconstruction methodology. Here also we follow the same data-driven deep learning approaches using CNN and LSTMs.

This research work is partly supported by Department of Science & Technology (DST) under the Internet of Things (IoT) Research of Interdisciplinary Cyber-Physical Systems (ICPS) Programme, GOI, New Delhi, with the Project entitled “IOT Based Holistic Prevention and Prediction of CVD (i-PREACT)”.

Continued…

[Funding Acknowledgment – MEITY-IoT for Smarter Healthcare project; DST-ICPS Project; Xilinx Inc., USA; SERB-ECR and MEITY-SMDP-C2S]

VI. Machine Learning techniques to build a cost-effective framework for fault-tolerant training of RRAM based neural computing systems.

An RRAM-based computing system (RCS) is widely used in neuromorphic computing systems due to its fast computation and low cost. However immature fabrication processes cause a high rate of hard faults. Also, the limited endurance of RRAMs restricts the life of RCS. We are using Machine Learning techniques to build a cost-effective framework for fault-tolerant training of RRAM based neural computing systems.

Physical design flow is an extremely time-consuming process when it comes to optimizing the designs. Due to multiple back and forth within the flow makes it a heavily time-consuming process and increases the turn around time of the final product. We are using ML and AI to reduce the turn around time and the cost. *[Funding Acknowledgment – DRDO ERIPR Project]*

Reforming Video Analytics

Prof. C. Krishna Mohan
Dept. of CSE
KID: 20200209

In recent years, the amount of visual data in the computer vision community is proliferating due to the reducing size and increasing reach of sensors. The understanding and analysis of visual data is hence indispensable to solve various computer vision tasks. We, VIGIL group at IIT Hyderabad, focus on cutting-edge visual understanding tasks includes surveillance video analytics for smart cities, fine-grained action recognition, spontaneous facial expression recognition, scalable and distributed methods for large scale visual computing, remote imagery analysis on satellite and radar data, a semantic description of video activities, autonomous vehicle technology, weather forecasting using a live camera, radar and satellite data, content-driven advertisement insertion, as well as anomaly detection in fine-grained actions. We focus on solving various computer vision tasks by constructing a semantically meaningful representation of videos. Our recent collaborations have included projects with OPPO India to address video blurring and de-duplication of images, as well as with Weathernews, Japan, to address the problem of precipitation now-casting by analyzing weather and road scenes. We have harvested datasets such as SkyEye, IITH-accident database (IITH-AD), and IITH-1 to investigate road user's collision behaviors. SkyEye dataset was introduced to detect collision prone vehicle behavior at intersections and contains 1 hour of continuous aerial footage from 4 major intersections in the city of Ahmedabad in India. IITH-AD and IITH-1 are captured from surveillance videos to investigate road traffic accidents. Our research has resulted in direct application and deployment in real-world applications, as well as publications at top-tier venues of high impact.