

RECENT INNOVATIONS IN DEEP LEARNING

¹G.Ramachandran, ² A.Malarvizhi, ³S.Kannan, ⁴ T.Sheela, ⁵ PM Murali,
^{1-2,3-5} Assistant Professor of Electronics and Communication Engineering, Vinayaka Mission's Kirupananda Variyar Engineering College, Vinayaka Mission's Research Foundation (Deemed to be University), Salem 636308, Tamilnadu, India.

⁴Associate Professor of Electronics and Communication Engineering, Vinayaka Mission's Kirupananda Variyar Engineering College, Vinayaka Mission's Research Foundation (Deemed to be University), Salem 636308, Tamilnadu, India.

Abstract:-

Deep learning uses artificial neural networks or computational models that resemble the intricate neurons of the human brain. Together these networks 'learn' from data and recognize patterns that ordinary computing would not be able to do, for example, in areas such as speech recognition and computer vision. Today, it takes on important classification, prediction and decision-making roles that experts say is furthering success in creating a fully functional Artificial Intelligence (AI) system. Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL) are used interchangeably in this paper. The larger field is artificial intelligence. This article is focusing on the Machine Learning piece of AI or more specifically the multi-layered neural networks form of Machine Learning called Deep Learning.

Keywords:- Automation ,Artificial Intelligence ,Robotics, Neural networks, Machine Learning

1.Introduction

The healthcare industry, faced with an overwhelming number of regulatory requirements and the need to maintain accurate and up-to-date patient records has turned to deep learning to provide tech- based assistance. The technology can be implemented to identify individuals who display a higher propensity for certain illnesses by programs that study the vast base of electronic health records. Based on the results pro-active action can be taken particularly when lifestyles are successfully matched with medical conditions.

1.1 Deep learning

The benefits of deep learning are also becoming more accessible to individual consumers and small businesses through open source, customizable software. One platform allows users to build smart applications that pre-empt user behaviour based on data science. Machine learning algorithms that are scalable and utilize existing big data networks can create predictive features such as personalization, recommendations, and content discovery for a number of digital properties. Now a days various applications, such as consumer, industrial, IoT, automotive, medical, drones and surveillance; thanks to ever-growing speeds, shrinking geometries, and ultra-low power semiconductor technologies and System-on-Chip (SoC) devices.

Various sensors including imaging, motion and environmental are attached to these edge devices. These sensors at edge devices generate enormous amounts of data including images, video, speech and other non-imaging data, which need to be transmitted back to cloud. Even if there is abundant and reliable transmission channel bandwidth, the round trip delays in transmitting data to cloud and getting back commands to be executed at the edge device, is prohibitive in most of the real-time, latency-sensitive applications. Further, security and privacy are the biggest concerns in transferring user data from edge devices to the cloud. Hence, there is a huge demand in enabling intelligent decisions in next generation edge devices, either fully autonomous way or semi-autonomous way.

2. Literature survey

The autonomous level of the edge devices depends on applications' criticality, latency sensitivity, security, privacy and available transmission bandwidth. For some applications, such as industrial, semi-autonomous edge devices may serve the purpose of several use cases by reducing volume of data to be transferred to cloud and by minimizing the round trip delays with distributed decision-making. For other mission-critical applications such as autonomous car, fully autonomous edge devices are necessary, with fusion of intelligence derived from various classes of sensors such as imaging, RADAR and LIDAR(which stands for Light Detection and Ranging, is a remote sensing method that uses light in the form of a pulsed laser to measure ranges) etc.

The entire spectrum of expected Machine Learning (ML) inference in edge devices can be categorized threefold- deriving intelligence out of imaging data, non-imaging data and their fusion.

Traditionally, ML and to be precise, Deep Learning (DL), is associated with the necessity of a huge number of powerful CPUs, GPUs, Cloud infrastructure and massive software packages. These DL frameworks need heavy compute requirement and longer inference run times up to several minutes. These compute requirements and run times make it infeasible to apply DL inference for real-time applications in edge devices. Nevertheless, the need of the hour is to bring down DL inference from cloud to next generation edge devices. The off-line machine learning process can still use compute farms or cloud, but the real-time inference shall run on edge-devices, which shall be very tiny, extremely low power consuming with minimal compute and storage needs. Further, additional learning by the machine is also possible during inference phase. Any such additional learning models shall be periodically downloaded into edge devices inference engine.

In the recent past, Artificial Neural Networks (ANN) based DL started to gain momentum and has shown promising results. However, there is no single magical ANN based DL solution for all the

applications. There are several types of neural networks for classification, prediction, clustering and association, out of which convolutional neural networks and recurrent neural networks are mostly deployed. The accuracy of ANN models, coupled with authenticity and volume of training datasets available, dictate the outcome of acceptance and rejection ratios, which are the key parameters for success of DL in general and edge devices in particular. Another challenge with DL today is existence of several frameworks and models. The convergence of some of these are highly desired for near-standard way of deploying DL in edge devices, from ease of usability and interoperability.

3. Proposed methodology

The expectations from edge devices are ever growing - in terms of integrating more and more functionalities, and sensors, but with not much increase in silicon area, power consumption or cost. In this context, bringing DL inference to edge devices should also adhere to these expectations. Any DL solution in the edge device, which needs power hungry and expensive discrete components including CPUs, GPUs and FPGAs (Field Programmable Gate Arrays), is detrimental to the said expectations. Even though such solutions can be a good proof-of-concept model, the power consumption and cost is prohibitive for many real-life applications.

The highly integrated and configurable SoC solutions with embedded DL hardware accelerators are necessary for DL inference needs of edge devices. It is also not practical to develop a single SoC edge device solution, which can cater to a broad range of applications. Different applications need different levels of DL inference complexity. At the same time, developing different DL based SoC edge device solutions for different applications is not practical either, unless the volumes are very high, from the ROI point of view. The more practical way is to develop DL based SoC edge device solutions, which can cater to a range of related applications. This would be a reasonable trade-off between power, performance and cost.



Future work

Medical images are valuable in diagnosing a wide range of health issues, as well as planning treatment and assessing its results. Aggregated with other sources of healthcare and demographic information, imaging data can also lead to novel insights that inspire next-generation treatment breakthroughs. Deep learning and other forms of artificial intelligence (AI) offer exciting potential to improve medical imaging work flows and enhance medical imaging quality. Deep learning is a branch of AI and machine learning in which developers create mathematical and neural network models, and use vast amounts of data to “train” them to perform tasks such as recognizing and classifying medical images. Once a trained model is sufficiently accurate, it can be deployed with other algorithms as an “inference engine,” to evaluate and categorize real-world inputs from digital cameras and sensors.

6. Conclusion,

Machine learning enables businesses and organizations to discover insights previously hidden within their data. Whether exploring oil reserves, improving the safety of automobiles, or mapping genomes, machine learning algorithms are at the heart of innovation and business intelligence. Unleashing the power of machine learning, however, requires certain ingredients: access to large amounts of diverse data and the right skill sets, optimized data platforms, powerful data analysis tools, and a highly scalable and flexible compute and storage infrastructure.

Reference:-

1. David Hanes, Gonzalo Salgueiro, Patrick Grossetete, Robert Barton, Jerome Henry, "IoT Fundamentals: Networking Technologies, Protocols, and Use Cases for the Internet of Things", 1st Edition, Pearson Education (Cisco Press Indian Reprint). (ISBN:978-9386873743)
2. Srinivasa K G, "Internet of Things", CENGAGE Learning India, 2017
3. Vijay Madiseti and Arshdeep Bahga, "Internet of Things (A Hands-on-Approach)", 1st Edition, VPT, 2014. (ISBN:978-8173719547)

4. Raj Kamal, "Internet of Things: Architecture and Design Principles", 1st Edition, McGraw Hill Education, 2017. (ISBN:978-9352605224)

5. Ahmed Abdulkadir Ibrahim, Wang Zhuopeng, College of Electronic Communication and Physics, Shandong University of Science and Technology, China "IOT Patient Health Monitoring System" Int. Journal of Engineering Research and Application "ISSN : 2248-9622, Vol. 8, Issue 1, (Part -III) January 2018, pp.77-80"

6. G. Ramachandran, PM Murali, T. Sheela, A. Malarvizhi, S. Kannan "Future Information And Communication Technology On The Health Care" International Journal of Advanced Research in Basic Engineering Sciences and Technology" Vol.5, Issue.1, January 2019" ISSN (ONLINE):2456-5717" Article DOI 10.20238/IJARBEST.2019.0501001"