**INTRODUCTION :**

IoT devices has led to an exponential increase in the amount of data generated, especially from image and video sources. Traditional cloud-based processing faces challenges in terms of latency, bandwidth limitations, and privacy concerns. Edge AI, which allows AI algorithms to run on IoT devices at the edge of the network, offers a solution by enabling real-time analysis and decision-making. This paper investigates the integration of AI with IoT for real-time image and video analysis, highlighting its importance, key benefits, and challenges.

* IoT devices have grown significantly,from 13% in 2014 to 25% today, with 43 billion connected devices in 2023.
* Real-time data processing is crucial, especially in remote surveillance systems using edge devices.
* Cloud computing struggles with transmission power, bandwidth, and latency issues, necessitating edge-based solutions.

**Objective :**

Propose an architecture that allows small businesses to process real-time video on IoT devices using edge AI technologies.

The architecture is:

- Easy to implement

- Cost-effective

- Processes video in real-time

- Enhances flexibility and security using

accelerators such as Google Coral and Intel Neural Compute Stick 2.

Real-time image and video analysis is crucial for many IoT applications. This section discusses:

* **Computer Vision Techniques**: Image classification, object detection, and video streaming analysis using edge AI.
* **Machine Learning Models for Image and Video**: Description of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) used for visual data processing.
* **Optimization Techniques**: Methods like model quantization, pruning, and edge-specific optimizations to reduce computational overhead and improve efficiency.

**Need for Edge Computing :**

* Cloud-based video surveillance systems require high data transfer rates (2.3 Mbps for a single H.265 1080p camera at 15 FPS).
* Cloud systems introduce latency of hundreds of milliseconds, which is critical for safety applications (e.g., pedestrian detection).
* Privacy risks exist with cloud-based video transmission, leading to privacy and legal challenges (e.g., GDPR compliance).

Several industries are leveraging edge AI for real-time image and video analysis. Some key applications include:

1. **Smart Surveillance**: Use of edge AI in security cameras to detect intrusions, identify faces, and track movement in real-time without relying on cloud processing.
2. **Autonomous Vehicles**: Real-time object detection and decision-making for self-driving cars to navigate safely.
3. **Healthcare Monitoring**: Real-time image analysis in wearable devices for monitoring patient health.
4. **Industrial Automation**: Detection of defects, monitoring of production lines, and safety alerts in manufacturing environments using edge AI-powered cameras.
5. **Drones and Robotics**: Real-time vision-based navigation and object tracking for autonomous drones and robots.

**WebRTC and RTSP Overview** **:**

* **WebRTC**: Enables real-time peer-to-peer video streaming with low latency, written in JavaScript and supported by most browsers.

Web RTC has always claimed to be a revolutionizing way that helps users communicate, both in the consumer and enterprise world.

Web RTC, as it is called, can support video streaming in real-time context with a P2P connection, making it one of the fastest solutions with zero latency and having a high-quality image.

• **RTSP**: Protocol for real-time streaming of media, commonly used in IP cameras and IoT devices.

RTSP is a protocol that provides essential functions to control the video’s flow and is combined with the Real-Time Transport Protocol (RTCP) to distribute the media flow.

RTSP is commonly used in camera devices (IP cameras or security cameras) to transfer media to a server or another client

• Both are key protocols for real-time video analysis in IoT systems.

**Proposed Architecture :**

* Interior Transfer Component: Captures and processes video frames usin AI accelerators (Google Coral, Intel NCS2).

• Exterior Transfer Component: Manages transfer of processed data to browsers via WebRTC.

• Process Data Component: Handles user requests like object detection, recognition, and blurring.

**Google Coral and Intel Compute Stick 2** :

* **Google Coral:** Edge TPU (Tensor Processing Unit) for AI tasks, supporting Tensor Flow Lite models, achieving 400 FPS (frames per second) for image classification.

Developed by Google, Coral is a complete toolkit for developers to build products with local Artificial Intelligence (AI). It consists of:

* Dev Board: a single-board computer with a removable systemon-module featuring the Edge Tensor Processing Unit (TPU).
* USB Accelerator: an Edge TPU USB compatible with Windows 10, macOS, and Debian Linux (including Raspberry Pi) host computer.
* Mini PCIe Accelerator, M.2 Accelerator A+E Key, M.2 Accelerator B+M Key: PCIe devices that enable easy integration of the Edge TPU into existing systems; support Debian Linux host computer.

• **Intel Compute Stick 2:** Powered by Intel Movidius VPU (vision processing unit), optimized

for real-time image analysis using OpenVINO toolkit. OpenVINO is a cross-platform deep learning toolkit developed by Intel. The name stands for “Open Visual Inference and Neural Network Optimization.” OpenVINO focuses on optimizing neural network inference with a write-once, deploy-anywhere approach for Intel hardware platforms.

**Methodology :-**

* Data Preparation: Collect and curate datasets.

• Data Augmentation: Apply transformations like flipping and brightness adjustments.

• Transfer Learning: Fine-tuning the MobileNetV2 architecture for object detection

• Training and Fine-tuning: Use pre-trained models and refine for IoT device deployment.

**Conclusion and Future Work** **:**

* Conclusion: Architecture is efficient for real-time IoT video and image analysis, achieving low latency and high performance.

• Future Work: Improve resolution, cross-platform compatibility, and performance with additional accelerators.