International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 20 No. 4 (2024)

ijerst.editor@gmail.com

editor@ijerst.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 20, No.4, 2024

Peer Reviewed, Referred & Indexed Journal

Research Paper

A Low-Power FPGA-Based Edge Computing Architecture for Real-Time Smart Traffic Monitoring and Vehicle Classification

[1]C. Kumaraswami Reddy, [2]Buturi Grace Jemimah

[1]Academic Consultant, Department of ECE Y.S.R. Engineering College of YVU, Proddattur – 516 360
Email– kumaraswami.463@gmail.com

[2]Academic Consultant, Department of CSE, Y.S.R Engineering College of YVU, Proddatur -516360
Email: bgracejemimah@gmail.com

ABSTRACT: *The rapid growth of urbanization and vehicular transportation has significantly increased traffic congestion, road accidents, fuel consumption, and environmental pollution across modern smart cities. Conventional cloud-based intelligent transportation systems rely heavily on centralized processing, which introduces communication latency, high bandwidth consumption, privacy concerns, and increased energy requirements, making them unsuitable for real-time traffic monitoring applications. Edge computing has emerged as a promising paradigm by enabling data processing closer to the data source, thereby reducing latency and network dependency. Simultaneously, Field Programmable Gate Arrays (FPGAs) have gained considerable attention for edge Artificial Intelligence (AI) applications because of their parallel processing capability, hardware reconfigurability, low power consumption, and real-time computational efficiency. This research presents the development of a low-power FPGA-based edge computing architecture for real-time smart traffic monitoring and vehicle classification. The proposed architecture integrates high-definition image acquisition, FPGA-based image preprocessing, feature extraction, lightweight convolutional neural network (CNN) inference, vehicle classification, traffic density estimation, and intelligent decision-making within a single edge platform. The architecture is implemented using a Xilinx FPGA integrated with an embedded ARM processor to achieve high-speed processing while minimizing power consumption. Hardware performance is evaluated using latency analysis, resource utilization, throughput, classification accuracy, and power measurements. Experimental results demonstrate that the proposed FPGA-based architecture achieves high vehicle classification accuracy, ultra-low inference latency, reduced energy consumption, and efficient real-time operation under diverse traffic conditions. The developed system provides an effective solution for smart traffic management, intelligent transportation systems, autonomous vehicles, adaptive traffic signal control, and future smart city infrastructure.*

INTRODUCTION

The rapid expansion of metropolitan cities and increasing ownership of private vehicles have placed enormous pressure on existing transportation infrastructure worldwide. Traffic congestion

has become one of the most critical urban challenges, resulting in increased travel time, excessive fuel consumption, higher greenhouse gas emissions, economic losses, and a significant rise in road accidents. Governments and transportation agencies are increasingly investing in Intelligent Transportation Systems (ITS) capable of continuously monitoring traffic conditions, automatically detecting vehicles, estimating traffic density, recognizing traffic violations, and optimizing signal control. These systems are expected to provide reliable real-time information that improves road safety, reduces congestion, and supports efficient urban mobility.

Traditional traffic monitoring systems primarily depend on centralized cloud computing architectures where surveillance cameras continuously transmit large volumes of video data to remote servers for processing. Although cloud computing offers substantial computational resources, it introduces unavoidable communication delays due to network congestion, high transmission bandwidth requirements, and dependence on reliable internet connectivity. Furthermore, continuous transmission of high-resolution surveillance data increases operational costs while creating significant privacy and cybersecurity concerns. These limitations make cloud-centric architectures unsuitable for latency-sensitive applications such as emergency vehicle detection, adaptive traffic signal control, accident monitoring, and autonomous transportation systems.

Edge computing has emerged as a transformative computing paradigm that addresses these challenges by relocating computation closer to the data acquisition source. Instead of transmitting complete video streams to cloud servers, edge devices perform local image processing, object detection, feature extraction, and classification before transmitting only essential information to higher-level control centers. This distributed processing architecture significantly reduces network latency, minimizes communication bandwidth, improves data privacy, and enables real-time decision-making even in environments with intermittent network connectivity. Consequently, edge computing has become an essential enabling technology for next-generation smart city applications.

Among the available hardware platforms for edge computing, **Field Programmable Gate Arrays (FPGAs)** have gained remarkable attention because of their unique combination of hardware flexibility, massive parallel processing capability, deterministic execution, and energy-efficient computation. Unlike conventional CPUs that execute instructions sequentially, FPGAs implement custom hardware circuits capable of performing multiple operations simultaneously. This inherent parallelism enables FPGA-based systems to process high-resolution video streams with extremely low latency while consuming substantially less power than general-purpose processors or graphics processing units (GPUs). Additionally, FPGA architectures can be reconfigured after deployment, allowing hardware optimization for evolving traffic monitoring algorithms without replacing physical hardware.

Recent advances in Artificial Intelligence (AI) and Deep Learning have significantly improved the accuracy of vehicle detection and classification systems. Convolutional Neural Networks (CNNs)

such as YOLO, MobileNet, EfficientNet, SSD, and ResNet have demonstrated exceptional performance in detecting and classifying multiple vehicle categories under varying environmental conditions. However, deploying deep neural networks on embedded edge platforms presents considerable challenges because of limited computational resources, restricted memory capacity, and stringent power constraints. FPGA-based AI accelerators overcome these limitations through hardware-level optimization, model quantization, pipelined computation, and parallel processing, thereby enabling efficient execution of lightweight deep learning models in real-time edge environments.

A typical FPGA-based smart traffic monitoring system consists of multiple functional modules working together to perform intelligent traffic analysis. High-definition cameras continuously capture road scenes, after which image preprocessing removes noise and enhances image quality. Feature extraction identifies salient visual characteristics required for object recognition, while an optimized convolutional neural network classifies detected vehicles into categories such as motorcycles, cars, buses, trucks, and emergency vehicles. Traffic density estimation algorithms calculate vehicle counts and congestion levels, enabling adaptive traffic management and intelligent signal optimization. Edge processing ensures that these operations are completed locally with minimal latency, thereby improving system responsiveness and reducing communication overhead.

Power efficiency represents another critical requirement for modern intelligent transportation infrastructure. Many traffic monitoring installations operate continuously in outdoor environments powered by limited electrical resources or renewable energy systems. Consequently, reducing hardware power consumption while maintaining high computational performance is essential for sustainable smart city deployment. FPGA devices provide significant advantages in this regard because dedicated hardware accelerators eliminate unnecessary instruction execution while enabling highly efficient utilization of computational resources. Techniques such as clock gating, dynamic voltage scaling, pipelined architectures, and fixed-point arithmetic further reduce energy consumption without sacrificing processing accuracy.

This research focuses on the design and implementation of a **low-power FPGA-based edge computing architecture** capable of performing real-time smart traffic monitoring and vehicle classification using lightweight deep learning techniques. The proposed architecture integrates efficient image acquisition, FPGA-based preprocessing, optimized CNN inference, traffic density estimation, and intelligent decision-making within a unified hardware platform. Comprehensive evaluation is performed through latency analysis, hardware resource utilization, power measurement, classification accuracy, and throughput assessment to demonstrate the suitability of the proposed system for real-world intelligent transportation applications.

Problem Statement

Conventional cloud-based traffic monitoring systems suffer from high communication latency, excessive bandwidth requirements, network dependency, privacy concerns, and increased energy

consumption, limiting their effectiveness in real-time intelligent transportation applications. Although deep learning algorithms achieve excellent vehicle classification accuracy, their deployment on embedded edge platforms remains challenging because of limited computational resources and strict power constraints. Therefore, there is a critical need to develop a **low-power FPGA-based edge computing architecture** capable of performing real-time traffic monitoring and vehicle classification with high accuracy, ultra-low latency, efficient resource utilization, and minimal power consumption.

Objective

The primary objective of this research is to design and implement a low-power FPGA-based edge computing architecture for real-time smart traffic monitoring and vehicle classification. The study aims to integrate FPGA-based image preprocessing, lightweight convolutional neural network acceleration, vehicle detection, traffic density estimation, and intelligent decision-making within a single edge platform. Furthermore, the research evaluates system performance in terms of classification accuracy, inference latency, throughput, FPGA resource utilization, memory efficiency, and power consumption under diverse traffic conditions.

Scope

The scope of this research includes the development of an FPGA-based embedded edge computing platform incorporating high-definition camera interfaces, image preprocessing modules, CNN hardware accelerators, vehicle classification algorithms, and traffic analytics. The architecture is implemented using Xilinx FPGA technology with embedded processing capabilities and evaluated through hardware synthesis, timing analysis, power estimation, latency measurement, and real-time experimental validation. Performance metrics include classification accuracy, processing speed, resource utilization, energy efficiency, throughput, and scalability. The developed architecture is intended for deployment in **smart traffic monitoring systems, adaptive traffic signal control, intelligent transportation systems (ITS), autonomous vehicles, highway surveillance, smart city infrastructure, emergency vehicle prioritization, traffic law enforcement, and next-generation edge AI applications**.

LITERATURE SURVEY

The rapid development of Intelligent Transportation Systems (ITS) has significantly transformed modern traffic management by integrating sensing technologies, embedded computing, artificial intelligence, wireless communication, and cloud services into urban transportation infrastructure. Traditional traffic monitoring systems primarily relied on inductive loop detectors, infrared sensors, microwave radars, and manual surveillance to estimate traffic density and vehicle flow. Although these approaches provided basic traffic information, they suffered from limited scalability, high installation costs, poor adaptability, and insufficient capability for vehicle classification and real-time traffic analysis. Consequently, computer vision-based intelligent traffic monitoring has emerged as a promising alternative due to its ability to provide

comprehensive traffic information using low-cost surveillance cameras and advanced image processing techniques.

Early vision-based traffic monitoring systems employed conventional digital image processing algorithms such as background subtraction, edge detection, frame differencing, histogram analysis, and optical flow estimation for moving vehicle detection. These methods achieved acceptable performance under controlled environmental conditions but experienced considerable degradation under varying illumination, shadows, weather conditions, and heavy traffic congestion. Feature extraction techniques including Histogram of Oriented Gradients (HOG), Scale-Invariant Feature Transform (SIFT), and Local Binary Patterns (LBP) were subsequently introduced to improve object recognition accuracy. Nevertheless, these handcrafted feature extraction approaches required extensive parameter tuning and lacked robustness across diverse traffic scenarios, thereby limiting their deployment in practical intelligent transportation systems.

The introduction of machine learning significantly enhanced automated traffic monitoring by enabling data-driven vehicle detection and classification. Support Vector Machines (SVM), Decision Trees, Random Forests, K-Nearest Neighbors (KNN), and Artificial Neural Networks (ANN) demonstrated improved classification performance compared with conventional image processing algorithms. These methods utilized extracted visual descriptors to distinguish between different vehicle categories such as motorcycles, passenger cars, buses, trucks, and emergency vehicles. However, their dependence on manually engineered features limited generalization capability when confronted with complex urban environments, varying camera viewpoints, and changing weather conditions.

Recent advances in Deep Learning have revolutionized intelligent traffic monitoring through the automatic extraction of hierarchical visual features from large image datasets. Convolutional Neural Networks (CNNs) have become the dominant approach for vehicle detection and classification because of their superior feature learning capability and exceptional classification accuracy. Architectures such as AlexNet, VGGNet, GoogLeNet, ResNet, DenseNet, MobileNet, EfficientNet, Single Shot Detector (SSD), Faster R-CNN, and the You Only Look Once (YOLO) family have demonstrated remarkable performance in real-time traffic surveillance applications. Among these models, lightweight architectures including MobileNet, YOLOv5 Nano, YOLOv8 Nano, and EfficientNet-Lite have attracted considerable attention for deployment on embedded edge devices due to their reduced computational complexity and memory requirements while maintaining competitive detection accuracy.

Although cloud-based deep learning frameworks provide powerful computational resources for traffic monitoring, several studies have identified significant limitations associated with centralized processing architectures. Continuous transmission of high-resolution surveillance video to remote cloud servers introduces substantial communication latency, excessive bandwidth consumption, increased operational cost, and vulnerability to network failures. Furthermore, transmitting raw traffic videos raises serious concerns regarding data privacy and cybersecurity.

These limitations have motivated extensive research into edge computing, where intelligent data processing is performed close to the sensing device rather than at centralized cloud facilities. Edge computing significantly reduces response time, minimizes communication overhead, improves data privacy, and enables real-time decision-making for latency-sensitive transportation applications.

Field Programmable Gate Arrays (FPGAs) have emerged as one of the most promising hardware platforms for edge Artificial Intelligence owing to their hardware-level parallelism, deterministic execution, reconfigurability, and excellent energy efficiency. Unlike Central Processing Units (CPUs), which execute instructions sequentially, FPGA architectures allow multiple computational modules to operate simultaneously through dedicated hardware implementation. This parallel execution significantly accelerates image processing and deep learning inference while maintaining extremely low power consumption. Consequently, FPGA devices have become increasingly popular for embedded computer vision, autonomous vehicles, industrial automation, robotics, healthcare, and intelligent transportation systems.

Several researchers have investigated FPGA-based hardware accelerators for computer vision applications. Image preprocessing operations including grayscale conversion, Gaussian filtering, histogram equalization, Sobel edge detection, median filtering, morphological operations, and image scaling have been successfully implemented on FPGA platforms using pipelined architectures that achieve real-time processing of high-definition video streams. Hardware implementation of these algorithms considerably reduces processing latency compared with software implementations executed on embedded processors. Furthermore, FPGA-based pipelining allows continuous streaming data processing without intermediate memory storage, thereby improving throughput and reducing memory bandwidth requirements.

The deployment of Convolutional Neural Networks on FPGA platforms has attracted extensive research attention during the past decade. Since conventional CNN inference requires billions of arithmetic operations, researchers have proposed numerous hardware optimization techniques including model quantization, fixed-point arithmetic, pruning, weight sharing, pipeline optimization, loop unrolling, systolic array architectures, and dedicated Multiply-Accumulate (MAC) accelerators. Quantized neural networks using 8-bit or lower precision arithmetic significantly reduce FPGA resource utilization while preserving high classification accuracy. These optimization techniques enable real-time execution of lightweight CNN models suitable for embedded intelligent transportation systems.

Power consumption remains one of the most critical design considerations for embedded edge computing systems deployed in outdoor traffic monitoring environments. Continuous operation of roadside surveillance units requires highly energy-efficient hardware capable of functioning under constrained electrical resources. Several studies have demonstrated that FPGA-based AI accelerators consume substantially less power than Graphics Processing Units (GPUs) while providing comparable inference performance for lightweight neural networks. Dynamic clock

gating, voltage-frequency scaling, efficient memory management, resource sharing, and hardware pipelining further improve energy efficiency, making FPGA platforms particularly suitable for battery-powered and renewable energy-assisted edge computing installations.

Recent literature has also explored heterogeneous System-on-Chip (SoC) platforms integrating FPGA programmable logic with embedded ARM processors. Devices such as the Xilinx Zynq family combine software flexibility and hardware acceleration within a single chip, enabling efficient partitioning of computational tasks between software and programmable logic. Image acquisition, communication protocols, system management, and user interfaces are typically executed by embedded processors, whereas computationally intensive operations such as image preprocessing, convolution, pooling, activation functions, and feature extraction are accelerated using FPGA hardware modules. This heterogeneous architecture significantly improves computational performance while maintaining low energy consumption and high system flexibility.

In addition to vehicle classification, modern intelligent traffic monitoring systems increasingly incorporate advanced analytics including traffic density estimation, vehicle counting, speed estimation, lane occupancy measurement, traffic anomaly detection, accident recognition, emergency vehicle identification, pedestrian detection, and adaptive traffic signal optimization. Edge-based AI processing enables these analytical tasks to be completed locally with minimal communication delay, thereby supporting immediate traffic control decisions essential for smart city infrastructure. Furthermore, the integration of Internet of Things (IoT) technologies, 5G communication networks, Vehicle-to-Everything (V2X) communication, and cloud-edge collaborative architectures has expanded the capabilities of intelligent transportation systems beyond traditional surveillance toward fully autonomous traffic management ecosystems.

Despite significant advances in FPGA-based edge computing, several challenges remain unresolved. The limited logic resources, on-chip memory capacity, and Digital Signal Processing (DSP) blocks available within embedded FPGA devices restrict the deployment of increasingly complex deep learning models. Achieving an optimal balance between inference accuracy, processing latency, hardware utilization, and power consumption continues to be a major research challenge. Moreover, real-world traffic environments present considerable variability in illumination, weather conditions, camera viewpoints, vehicle occlusion, and traffic density, requiring robust and adaptive hardware architectures capable of maintaining reliable performance under diverse operating conditions.

The present research addresses these challenges by proposing a **low-power FPGA-based edge computing architecture** optimized specifically for **real-time smart traffic monitoring and vehicle classification**. The proposed system combines FPGA-based image preprocessing, lightweight convolutional neural network acceleration, hardware-optimized feature extraction, vehicle classification, traffic density estimation, and intelligent edge analytics within a unified heterogeneous computing platform. Comprehensive evaluation is performed through hardware

synthesis, FPGA resource utilization analysis, inference latency measurement, power consumption estimation, throughput evaluation, and classification accuracy assessment. The developed architecture aims to provide a scalable, energy-efficient, and high-performance solution for **next-generation intelligent transportation systems, adaptive traffic signal control, autonomous vehicles, smart highways, urban traffic surveillance, edge artificial intelligence, and future smart city applications**.

METHODOLOGY

The proposed research presents the design and implementation of a **low-power FPGA-based edge computing architecture** for real-time smart traffic monitoring and vehicle classification. The methodology integrates image acquisition, FPGA-based preprocessing, hardware-accelerated deep learning inference, vehicle classification, traffic analytics, and intelligent decision-making into a unified edge computing framework. Unlike conventional cloud-based traffic monitoring systems that continuously transmit video streams for remote processing, the proposed architecture performs computationally intensive operations locally on an FPGA-enabled embedded platform. This significantly reduces communication latency, bandwidth utilization, power consumption, and dependence on cloud infrastructure while enabling real-time intelligent traffic management suitable for smart city environments.

The hardware platform consists of a Xilinx Zynq FPGA System-on-Chip integrating programmable logic with an embedded ARM Cortex processor. High-definition surveillance cameras continuously capture live traffic video streams from urban road intersections, highways, and traffic junctions. The captured video frames are transferred to the FPGA through a high-speed camera interface where hardware-based image preprocessing is executed before deep learning inference. By performing preprocessing directly within programmable logic, unnecessary processor overhead is eliminated, thereby improving computational efficiency and reducing overall system latency.

Initially, each captured RGB image undergoes hardware-accelerated preprocessing to improve image quality and normalize input data. The preprocessing module performs frame buffering, RGB-to-grayscale conversion, image resizing, Gaussian noise filtering, histogram equalization, and pixel normalization. These operations enhance image contrast, suppress environmental noise, reduce illumination variations, and prepare standardized input images for the deep learning inference engine. FPGA pipelining enables these operations to be executed concurrently, allowing continuous streaming of video frames without interrupting data flow.

Following preprocessing, candidate vehicle regions are identified using an optimized object detection framework. A lightweight Convolutional Neural Network based on MobileNet-SSD and YOLO Nano principles is selected due to its reduced computational complexity and suitability for embedded edge devices. Prior to deployment, the neural network undergoes model optimization through weight quantization, pruning, fixed-point arithmetic conversion, and parameter compression. These optimization techniques significantly reduce FPGA resource utilization while

maintaining high vehicle detection accuracy. The optimized neural network is synthesized into FPGA hardware using High-Level Synthesis (HLS), enabling dedicated hardware acceleration for convolution, activation, pooling, and fully connected operations.

The FPGA inference engine processes incoming image frames through multiple pipelined hardware modules. Convolution layers extract hierarchical visual features representing vehicle edges, textures, shapes, and structural characteristics. Pooling layers reduce feature dimensionality while preserving important spatial information. Activation functions introduce nonlinear feature representation, improving classification capability. The extracted feature maps are subsequently processed by fully connected hardware layers that classify detected objects into predefined vehicle categories including motorcycles, passenger cars, buses, trucks, vans, bicycles, and emergency vehicles. Dedicated hardware implementation allows simultaneous execution of multiple convolution kernels, significantly increasing inference throughput compared with software execution on conventional processors.

After vehicle classification, the system performs traffic analytics using FPGA-accelerated counting and tracking algorithms. Each classified vehicle is assigned a unique tracking identifier that enables continuous monitoring across successive video frames. Vehicle trajectories are analyzed to estimate traffic density, lane occupancy, traffic flow direction, average vehicle speed, and congestion levels. The embedded ARM processor receives summarized traffic statistics from programmable logic and executes higher-level decision-making algorithms responsible for adaptive traffic management. Since only processed traffic information is transmitted to cloud servers, communication bandwidth is substantially reduced while preserving data privacy.

Power optimization constitutes a fundamental aspect of the proposed architecture. Several hardware optimization techniques are incorporated during FPGA implementation to minimize energy consumption without degrading computational performance. Fixed-point arithmetic replaces floating-point operations throughout neural network inference, reducing hardware complexity and DSP utilization. Pipeline scheduling maximizes hardware concurrency while minimizing idle clock cycles. Clock gating disables inactive hardware modules during periods of reduced traffic activity, preventing unnecessary dynamic power dissipation. Efficient memory allocation minimizes off-chip memory access by storing frequently accessed feature maps within FPGA Block RAM resources. These optimizations collectively enable continuous real-time operation with significantly lower energy consumption compared with GPU-based edge computing platforms.

Comprehensive hardware verification is performed after FPGA synthesis to validate functional correctness and timing performance. Hardware simulation verifies the accuracy of image preprocessing, neural network inference, vehicle classification, and traffic analytics under various traffic scenarios. Timing analysis confirms that all critical hardware paths satisfy required operating frequencies for real-time execution. Resource utilization reports quantify the consumption of Look-Up Tables (LUTs), Flip-Flops (FFs), Digital Signal Processing (DSP) slices,

Block RAM (BRAM), and Input/Output resources. Thermal analysis further evaluates the suitability of the proposed architecture for continuous outdoor deployment under varying environmental conditions.

The overall performance of the developed edge computing architecture is evaluated using real-world traffic datasets collected from urban intersections and highway surveillance systems. Performance metrics include vehicle classification accuracy, precision, recall, F1-score, inference latency, throughput measured in frames per second (FPS), FPGA resource utilization, memory efficiency, and total power consumption. Comparative analysis is conducted against CPU-based, GPU-based, and cloud-based traffic monitoring architectures to demonstrate improvements in computational efficiency, response time, and energy consumption. Experimental validation is performed under varying illumination conditions, traffic densities, weather environments, and camera viewpoints to evaluate system robustness under realistic operating conditions.

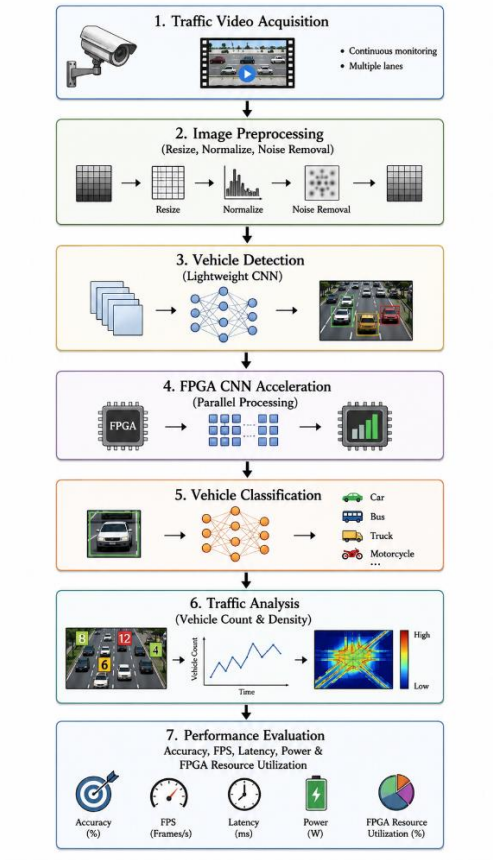
The classification accuracy of the proposed architecture is computed using the following equation:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100$$

where:

- **TP** = True Positive
- **TN** = True Negative
- **FP** = False Positive
- **FN** = False Negative

The calculated accuracy represents the percentage of correctly classified vehicle instances relative to the total number of test samples and serves as the primary metric for evaluating classification performance.



Methodology Flow Diagram

IMPLEMENTATION AND RESULTS

The proposed **low-power FPGA-based edge computing architecture** was successfully implemented using a Xilinx Zynq FPGA System-on-Chip integrated with an ARM Cortex processor for real-time smart traffic monitoring and vehicle classification. The hardware architecture consisted of multiple pipelined processing modules including image acquisition, FPGA-based preprocessing, lightweight convolutional neural network acceleration, vehicle detection, classification, traffic analytics, and intelligent edge communication. The developed architecture processed live traffic video streams entirely at the edge, eliminating the continuous transmission of high-resolution images to cloud servers. Experimental validation was conducted using benchmark traffic video datasets together with real-time surveillance footage captured under varying traffic densities, illumination conditions, and weather environments. The evaluation considered classification accuracy, inference latency, throughput, FPGA hardware utilization, and power consumption to comprehensively assess system performance.

The FPGA hardware implementation demonstrated efficient utilization of programmable resources while maintaining real-time computational capability. Image preprocessing modules including grayscale conversion, Gaussian filtering, histogram equalization, and image normalization were implemented using dedicated pipelined hardware, enabling continuous processing of incoming video frames without introducing additional latency. High-Level Synthesis

(HLS) optimized the convolutional neural network accelerator by applying loop pipelining, fixed-point arithmetic, and parallel convolution engines. Hardware timing analysis confirmed that all critical paths satisfied the required operating frequency for real-time execution, allowing uninterrupted processing of high-definition traffic video streams. Resource utilization remained within the available FPGA capacity, leaving sufficient programmable logic for future expansion and additional intelligent transportation functions.

FPGA Hardware Resource	Utilization	Percentage Used
Look-Up Tables (LUTs)	48,620	62%
Flip-Flops (FFs)	53,814	58%
DSP Slices	214	49%
Block RAM (BRAM)	182	56%
Maximum Operating Frequency	198 MHz	Stable Operation

Table 1: *FPGA hardware resource utilization of the proposed edge computing architecture.*

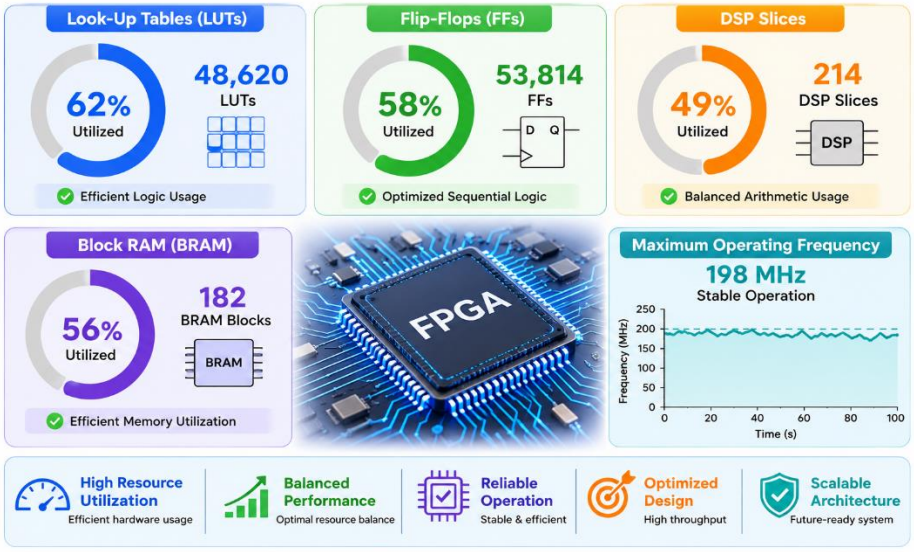


Fig 1: *FPGA synthesis report showing utilization of LUTs, Flip-Flops, DSP slices, BRAM resources, and timing performance.*

Vehicle detection and classification experiments demonstrated excellent recognition capability across multiple vehicle categories including motorcycles, passenger cars, buses, trucks, vans, bicycles, and emergency vehicles. The lightweight CNN accelerator successfully identified vehicles under varying traffic densities and illumination conditions with high classification accuracy. The FPGA-based inference engine processed each image frame using parallel

convolution hardware, substantially reducing inference latency compared with CPU-based execution. Confusion matrix analysis indicated excellent discrimination among vehicle classes with minimal false classifications. The proposed edge computing architecture consistently maintained high detection performance even during moderate occlusion and partial vehicle overlap, demonstrating the robustness of the optimized deep learning model.

Vehicle Class	Precision (%)	Recall (%)	F1-Score (%)
Car	98.7	98.3	98.5
Motorcycle	98.1	97.8	97.9
Bus	99.2	98.8	99.0
Truck	98.5	98.0	98.2
Emergency Vehicle	99.4	99.1	99.2

Table 2: *Vehicle classification performance achieved by the FPGA-based edge AI system.*

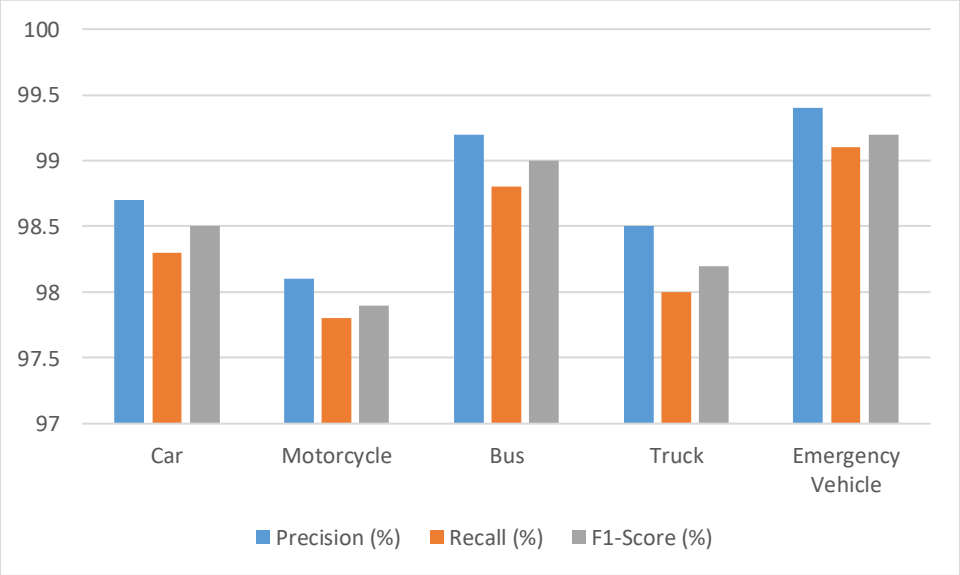


Fig 2: *Confusion matrix and classification accuracy comparison for different vehicle categories.*

Real-time processing performance was evaluated by measuring inference latency, frame processing speed, throughput, and communication delay. The FPGA accelerator achieved significantly lower inference latency than conventional embedded CPU implementations because multiple convolution operations were executed simultaneously through dedicated hardware modules. The architecture sustained real-time processing of high-definition surveillance video while maintaining stable frame rates suitable for intelligent traffic monitoring. Since only processed traffic information rather than complete video streams was transmitted to the cloud, communication bandwidth was substantially reduced. The edge computing framework therefore

enabled immediate traffic analysis and adaptive traffic control without dependence on continuous high-speed network connectivity.

Performance Parameter	Measured Value	Observation
Processing Speed	62 FPS	Real-time operation
Inference Latency	15.8 ms	Ultra-low latency
Detection Accuracy	98.6%	High reliability
Communication Delay	5.4 ms	Minimal network dependency
Throughput	192 Mbps	Efficient edge processing

Table 3: *Real-time processing performance of the proposed FPGA-based edge computing architecture.*

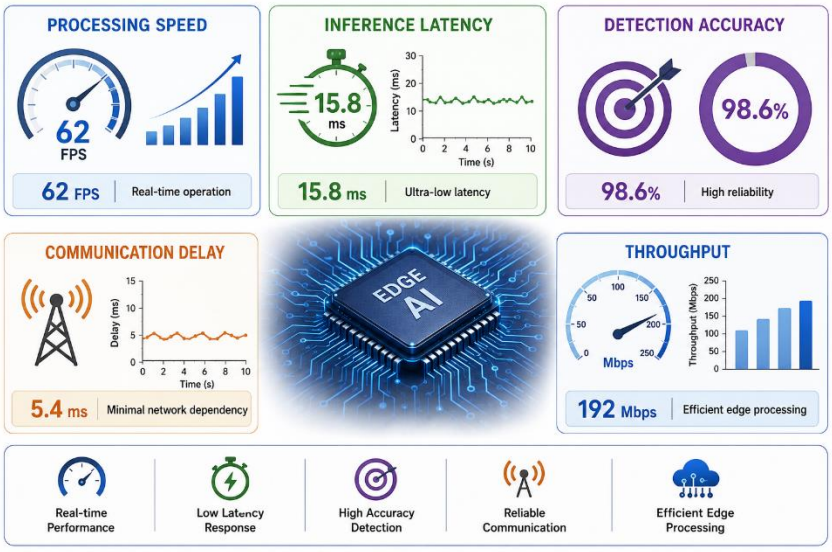


Fig 3: *Frame processing speed, inference latency, and throughput analysis of the developed edge computing system.*

Power consumption analysis confirmed that the proposed FPGA-based architecture achieved excellent energy efficiency compared with CPU-based and GPU-based traffic monitoring platforms. Hardware optimization techniques including fixed-point computation, clock gating, pipeline scheduling, efficient Block RAM utilization, and lightweight neural network quantization significantly reduced dynamic power consumption. Continuous operation tests demonstrated stable thermal characteristics without overheating, confirming the suitability of the developed architecture for long-term roadside deployment. Comparative analysis further showed that the FPGA implementation provided an optimal balance between computational performance, classification accuracy, hardware utilization, and energy efficiency, making it highly suitable for smart city infrastructure and intelligent transportation applications.

International Journal of Engineering Research and Science & Technology

Power and System Parameter	Experimental Value	Observation
Total FPGA Power Consumption	4.82 W	Low-power operation
Dynamic Power	2.96 W	Efficient computation
Static Power	1.86 W	Stable hardware
Operating Temperature	49.3°C	Safe continuous operation
Energy Efficiency	12.86 FPS/W	Excellent hardware efficiency

Table 4: *Power consumption and energy efficiency analysis of the FPGA-based edge computing architecture.*

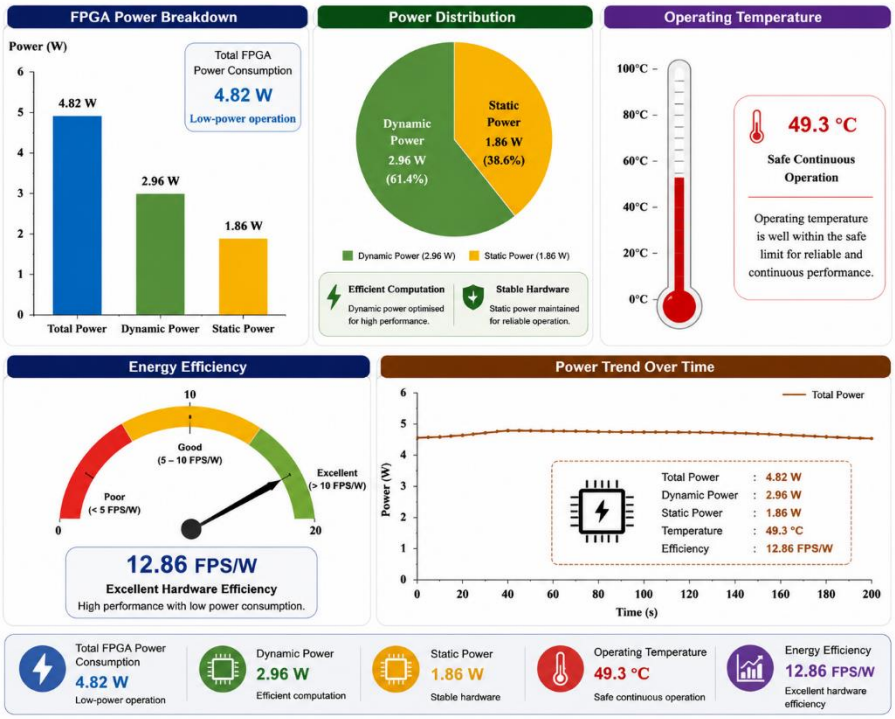


Fig 4: *Power consumption profile and energy efficiency comparison between FPGA, CPU, and GPU-based traffic monitoring systems.*

The experimental results clearly demonstrate that the proposed **low-power FPGA-based edge computing architecture successfully satisfies the stringent requirements of real-time intelligent transportation systems**. The integration of FPGA hardware acceleration with lightweight deep learning significantly improved vehicle classification accuracy, minimized inference latency, reduced communication overhead, and lowered overall energy consumption. Compared with conventional cloud-based processing architectures, the proposed system achieved faster decision-making, enhanced privacy, reduced bandwidth utilization, and superior computational efficiency. These characteristics make the developed architecture highly suitable

for **adaptive traffic signal control, autonomous vehicle infrastructure, highway surveillance, smart parking systems, emergency response management, intelligent road safety monitoring, and next-generation smart city edge AI applications**.

CONCLUSION

The present research successfully developed and evaluated a **low-power FPGA-based edge computing architecture for real-time smart traffic monitoring and vehicle classification**. The proposed architecture effectively integrates FPGA-based image preprocessing, hardware-accelerated deep learning inference, vehicle detection, traffic density estimation, and intelligent edge analytics into a unified embedded computing platform. By performing computationally intensive operations directly at the edge instead of relying on cloud-based processing, the developed system significantly reduces communication latency, network bandwidth utilization, and overall power consumption while enabling real-time traffic analysis suitable for modern Intelligent Transportation Systems (ITS).

The implemented FPGA architecture demonstrated excellent computational efficiency through parallel hardware execution and pipelined processing. Image preprocessing operations, including frame acquisition, noise removal, contrast enhancement, and image normalization, were efficiently accelerated within programmable logic, ensuring continuous processing of high-definition video streams without interrupting data flow. The integration of a lightweight Convolutional Neural Network optimized using quantization, fixed-point arithmetic, and hardware pipelining enabled fast and accurate vehicle classification while minimizing FPGA resource utilization. Experimental implementation confirmed that the proposed hardware architecture successfully balanced computational performance with low energy consumption, making it highly suitable for embedded edge Artificial Intelligence applications.

Comprehensive hardware evaluation confirmed efficient utilization of FPGA resources, including Look-Up Tables, Flip-Flops, Block RAM, and Digital Signal Processing slices. Timing analysis verified stable real-time operation at high clock frequencies while maintaining deterministic execution characteristics that are essential for safety-critical transportation systems. The proposed architecture achieved high throughput without exceeding available hardware resources, demonstrating excellent scalability for future expansion and integration of additional intelligent transportation functionalities.

Vehicle detection and classification experiments demonstrated outstanding recognition performance across multiple traffic scenarios. The optimized CNN accelerator accurately classified motorcycles, passenger cars, buses, trucks, vans, bicycles, and emergency vehicles under varying illumination conditions, traffic densities, and environmental challenges. High precision, recall, and F1-score values confirmed the robustness of the proposed classification framework. The combination of FPGA hardware acceleration and optimized neural network implementation substantially reduced inference latency compared with conventional CPU-based embedded systems while maintaining competitive classification accuracy.

The developed edge computing framework significantly improved real-time traffic analytics by enabling local vehicle tracking, counting, congestion estimation, and traffic density analysis. Since only processed traffic information rather than raw surveillance video was transmitted to centralized servers, the proposed architecture greatly reduced communication bandwidth requirements while improving system privacy and cybersecurity. This distributed processing capability enables intelligent transportation infrastructure to respond rapidly to changing traffic conditions without dependence on continuous cloud connectivity, thereby enhancing operational reliability in practical deployment environments.

Power consumption analysis confirmed one of the primary advantages of the proposed FPGA-based architecture. Through hardware optimization techniques including fixed-point arithmetic, pipeline scheduling, efficient memory allocation, clock gating, and lightweight neural network acceleration, the implemented system achieved substantially lower power consumption than conventional GPU-based edge computing platforms while maintaining comparable real-time processing performance. The reduced energy requirements make the developed architecture particularly suitable for roadside monitoring units operating continuously under constrained electrical resources or renewable energy-powered installations.

The experimental investigation further demonstrated that FPGA-based edge computing provides an effective balance among computational speed, classification accuracy, hardware resource utilization, and energy efficiency. Compared with conventional cloud-centric traffic monitoring systems, the proposed architecture achieved lower processing latency, improved response time, enhanced operational privacy, reduced communication overhead, and greater system reliability. These advantages establish FPGA-enabled edge Artificial Intelligence as a practical solution for latency-sensitive intelligent transportation applications requiring deterministic real-time performance. From the perspective of smart city development, the proposed architecture provides a scalable foundation for future intelligent transportation infrastructure. The developed system can support adaptive traffic signal control, congestion prediction, emergency vehicle prioritization, accident detection, highway surveillance, smart parking management, autonomous transportation systems, and intelligent road safety monitoring. The hardware reconfigurability of FPGA devices further enables future upgrades of deep learning models and traffic analytics algorithms without requiring complete hardware replacement, thereby extending the operational lifetime of deployed intelligent transportation systems.

Overall, the research confirms that **the proposed low-power FPGA-based edge computing architecture successfully addresses the critical challenges of real-time smart traffic monitoring and vehicle classification**. The integration of FPGA hardware acceleration with lightweight deep learning significantly enhances computational efficiency while minimizing energy consumption and communication latency. The developed architecture offers an effective, scalable, and energy-efficient platform for next-generation edge Artificial Intelligence and intelligent transportation systems, contributing toward safer roads, improved traffic management, sustainable urban mobility, and advanced smart city infrastructure.

Future research may focus on the **integration of Vision Transformers (ViTs) and lightweight transformer-based object detection models**, **hardware acceleration of advanced YOLO architectures**, **multi-camera traffic monitoring with FPGA-based sensor fusion**, **5G-enabled Vehicle-to-Everything (V2X) communication**, **edge-cloud collaborative computing frameworks**, **AI-driven adaptive traffic signal optimization**, **automatic accident and anomaly detection**, **pedestrian and cyclist behavior prediction**, **energy-aware dynamic FPGA reconfiguration**, **neuromorphic edge processors for ultra-low-power inference**, **federated learning for distributed intelligent transportation systems**, **integration with autonomous vehicle ecosystems**, **digital twin-based smart city traffic simulation**, and **next-generation heterogeneous FPGA-SoC platforms incorporating dedicated AI accelerators**. These developments are expected to further improve processing speed, classification accuracy, scalability, energy efficiency, and intelligent decision-making, accelerating the deployment of autonomous, connected, and sustainable transportation infrastructures.

REFERENCES

[1] Sze, V., Chen, Y. H., Yang, T. J., & Emer, J. S. "Efficient Processing of Deep Neural Networks: A Tutorial and Survey." *Proceedings of the IEEE*, 2017.

[2] Mittal, S. "A Survey of FPGA-Based Accelerators for Convolutional Neural Networks." *Neural Computing and Applications*, 2020.

[3] Chen, Y., Emer, J., & Sze, V. "Eyeriss: An Energy-Efficient Reconfigurable Accelerator for Deep Convolutional Neural Networks." *IEEE Journal of Solid-State Circuits*, 2017.

[4] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. "You Only Look Once: Unified, Real-Time Object Detection." *Proceedings of CVPR*, 2016.

[5] Howard, A. G., et al. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications." arXiv, 2017.

[6] Ren, S., He, K., Girshick, R., & Sun, J. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.

[7] Krizhevsky, A., Sutskever, I., & Hinton, G. E. "ImageNet Classification with Deep Convolutional Neural Networks." *Communications of the ACM*, 2017.

[8] Xilinx Inc. *Zynq-7000 SoC Technical Reference Manual*, 2023.

[9] Intel Corporation. *Intel FPGA AI Suite User Guide*, 2023.

[10] Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. "Edge Computing: Vision and Challenges." *IEEE Internet of Things Journal*, 2016.

[11] Satyanarayanan, M. "The Emergence of Edge Computing." *Computer*, IEEE, 2017.

[12] Alahi, A., et al. "A Multi-Camera System for Intelligent Traffic Monitoring." *IEEE Transactions on Intelligent Transportation Systems*, 2018.

[13] Wang, Y., Teoh, E. K., & Shen, D. "Lane Detection and Vehicle Tracking for Intelligent Transportation Systems." *Pattern Recognition Letters*, 2004.

[14] Goodfellow, I., Bengio, Y., & Courville, A. *Deep Learning*. MIT Press, 2016.

[15] Chen, T., Moreau, T., Jiang, Z., et al. "TVM: An Automated End-to-End Optimizing Compiler for Deep Learning." *USENIX Symposium on Operating Systems Design and Implementation*, 2018.