iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.80578

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

FPGA-Based Hardware Accelerator for K-Means Clustering Algorithm

V. Suresh Kumar[1], Rehan Shaik[2], K. Sai Bharath[3], Mohammed Azeem Ali[4]
[1]*Assistant Professor, Electronics and Communication Engineering, MVSR Engineering College, Hyderabad, India*
[2,3,4]*UG-Electronics and Communication Engineering, MVSR Engineering College, Hyderabad, India*

Abstract: ***K-means is an extremely popular data analysis/pattern recognition algorithm, but the large amounts of calculation involved for the distance measures and centroids makes it computationally intensive to execute on CPU-based systems, as the algorithms typically iteratively update their centroids. This project demonstrates an FPGA based hardware accelerator for K-means implemented in Verilog RTL, designed on the Nexys A7 (Artix-7 board). This design is pipelined and parallelised for high throughput, and utilizes fixed-point arithmetic to minimize resources. ILA (Integrated Logic Analyzer) is used for the hardware demonstration to allow the internal state machine states, iteration counters and pipeline outputs to be recorded so that a precise cycle accurate debugging of the completely pipelined clustering architecture can be completed.***

I. INTRODUCTION

K-means is a prevalent and widely used unsupervised learning algorithm, especially applied in the fields of data mining, image segmentation, customer segmentation, bioinformatics, anomaly detection and others. This can be attributed to its simplicity of concept and ability to group data points based on their closeness to the center of a cluster. K-means iteratively assigns data points to the nearest cluster centers and then updates the centers, thus effectively building cluster out of the unlabeled data.

In case of large scaled problems (larger numbers of data points, dimensions, clusters), the computational complexity of the K-means algorithm increases with time. In traditional implementations running on a CPU, the K-means algorithm takes longer and consumes more power, resulting in slow performance, and thus the scalability for large data is a concern. Hardware acceleration using the Field Programmable Gate Array (FPGA) has been proposed as a potential efficient hardware acceleration technique for the K-means algorithm, which is because FPGA is highly parallelizable and it is able to support pipelined implementation.

Advantages of implementing K-means in hardware are higher throughput, lower latency, and deterministic timing. In this paper, a completely pipelined K-means hardware accelerator was design using Verilog RTL on Nexys A7 (Artix-7). The memory efficient implementation was done using on-chip Block Random Access Memory (BRAM) and the DSP resource was utilized for high-speed calculation of distance. Hardware implementation was verified using Integrated Logic Analyzer (ILA), which gives a real time view of the internal signal such as the transition state of Finite State Machine (FSM) and the pipeline output. The result obtained shows cycle-level visibility in terms of behavior..

II. LITERATURE REVIEW

There have been multiple attempts to speed up K-means algorithm using FPGA implementations. For instance, L. Zhou et al. developed a pipelined and parallel FPGA accelerator, which speeds up calculations and decreases computing time due to the accelerated computation of Euclidean distance and centroids, demonstrating high speedup compared to CPU implementations [1].

In addition, H. Li et al. designed a fast FPGA K-means accelerator implementing concepts of High-Level Synthesis such as unrolling loops and scheduling pipelines, resulting in decreased latency and reduced energy costs. Thus, this study emphasizes the benefits from the combination of hardware parallel processing and software flexibility [2].

The issue of energy consumption also became a research topic in FPGA clustering applications. In particular, S. Park et al. studied ways to save energy when performing K-means on FPGA, using fixed-point arithmetic and optimizing memory accesses [3]. This work highlights the importance of both aspects for effective FPGA design.

Furthermore, recent research continues exploring FPGA K-means solutions. For example, C. S. Dusane et al. proposed a practical implementation of FPGA K-means and K-medoids algorithms, demonstrating improved performance and scalability. Z. He et al. implemented a precision-adaptive approach for accelerating K-means on FPGA.

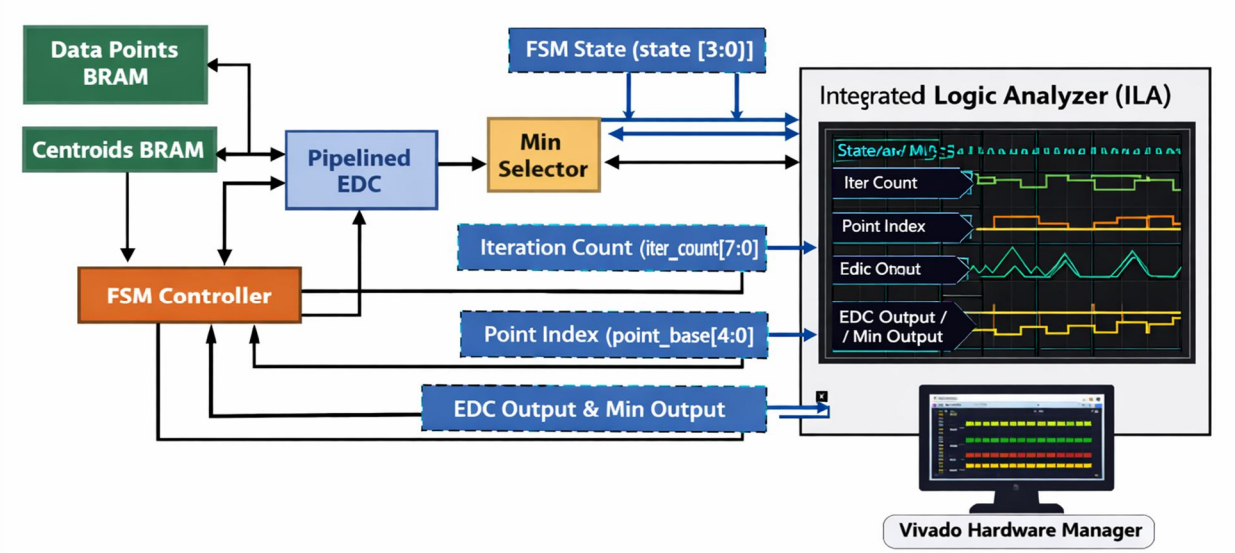
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

III. PROPOSED WORK

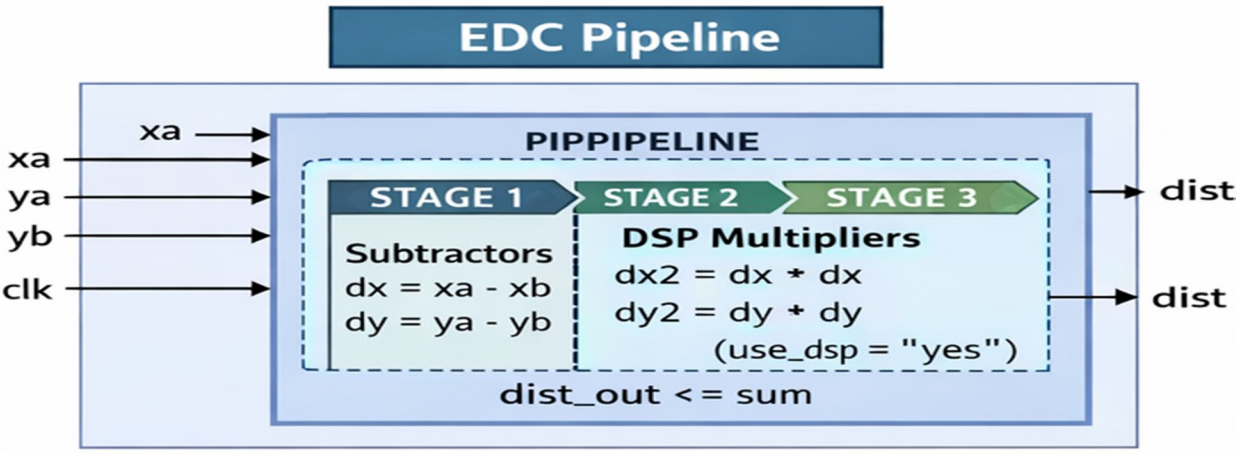
The design under consideration is an FPGA based hardware accelerator for K-means clustering algorithm. It makes use of a highly efficient pipelined architecture so that its performance can be enhanced, and a large volume of data processed. The design is done using Verilog HDL for Artix-7 FPGA family. Both the data points as well as the centroid values are stored in Block RAMs on-chip. This helps to avoid any bottleneck and also decrease the memory access time.

The heart of the system is a pipelined Euclidean Distance Calculator (EDC) which exploits FPGA DSP blocks for performing computations quickly and achieving timing constraints. In the first stage distances between data points and centroids are calculated parallel pipelined fashion. The results are fed into a comparator based pipelined network that selects minimum of these distances for each of the data points and associates it with respective cluster index.

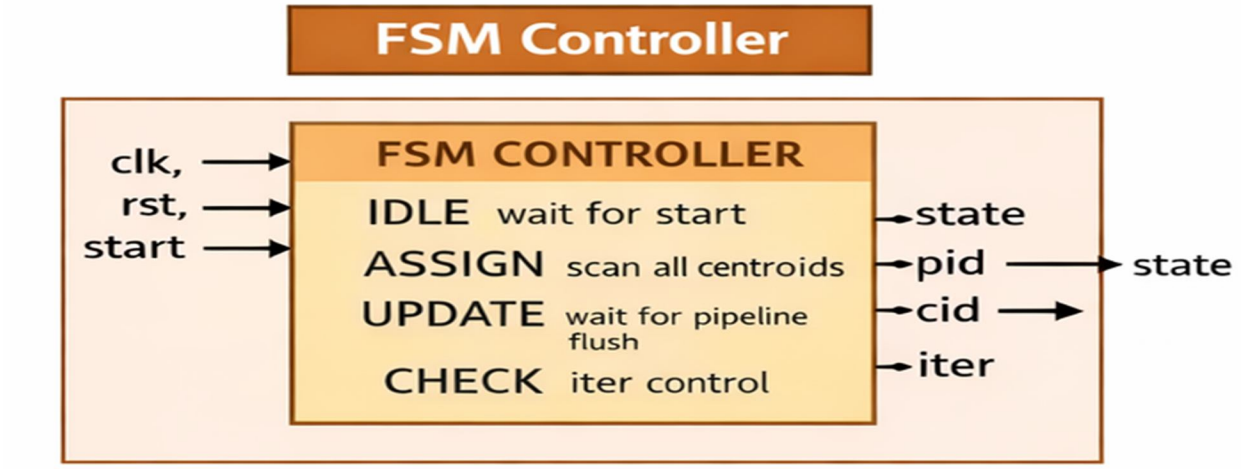
Xilinx Integrated Logic Analyzer (ILA) core is also included in the design to allow hardware level verification. Crucial internal signals such as states of the FSMs, iteration count, point indexing logic and pipelined output from EDC and minimum selector blocks are observed using the ILA tool for the cycle accurate observations of internal data and timing. This allows us to examine every detail of the pipelined design, synchronize on-chip RAMs, verify BRAM usage and investigate the data flow of the system through ILA.



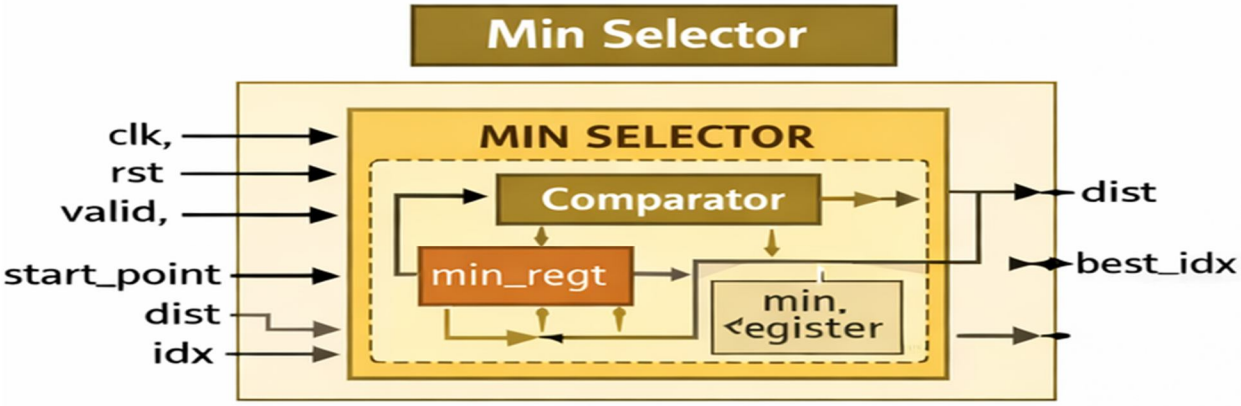
FPGA-Based K-Means Clustering Accelerator with ILA Probes Capturing Internal Signals



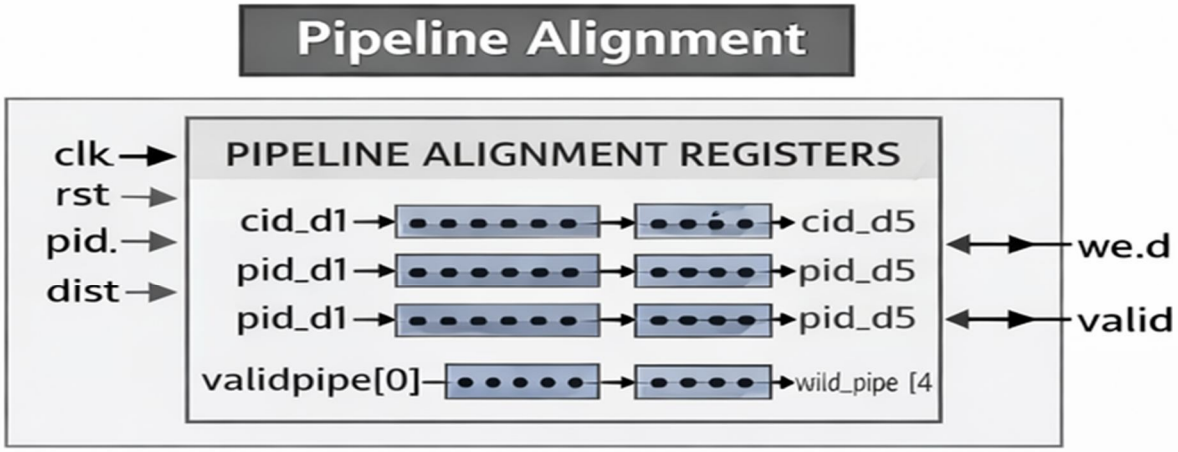
Calculates squared Euclidean distance in a pipelined architecture using three stages.



Controls state transitions and iteration flow for the K-Means algorithm

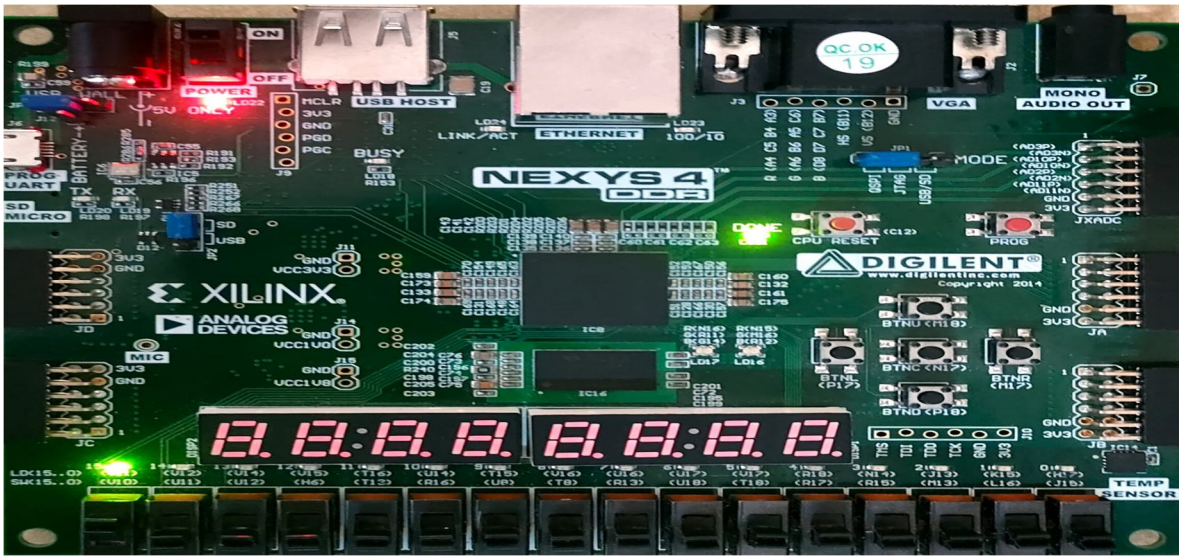


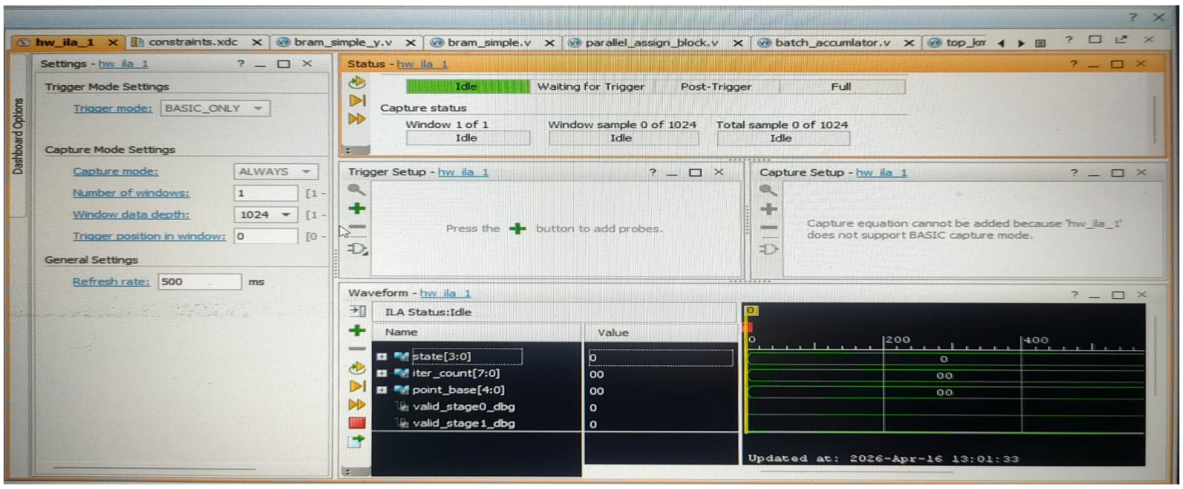
Compares distances and selects the minimum distance cluster index

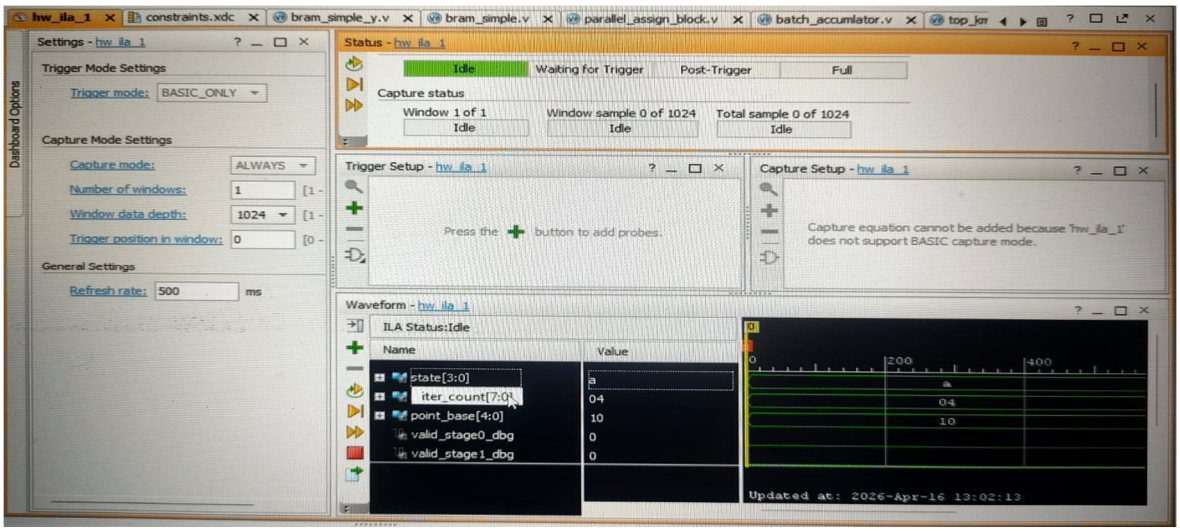


Aligns centroid index, point index, and valid signal with delayed distance output

IV. EXPERIMENTAL RESULTS







Metric	Software (CPU, K=4)	FPGA Accelerator (Your Design)
Clock Frequency	CPU dependent	100 MHz
Execution Model	Sequential nested loops	FSM + pipelined datapath
Distance Computation	4 sequential evaluations	4 pipelined evaluations
Latency per Point	Sequential over K	~9 cycles (90 ns)
Throughput	Loop-limited	1 point / 4 cycles (~25 M points/s)
Distance Eval/sec	CPU dependent	~100 M eval/s
Memory Access	Cache/DRAM	BRAM (1-cycle access)
Memory Size (On-chip)	Not fixed	~1.08 KB (data + centroids + assignments)
Determinism	Non-deterministic	Fully deterministic
Iterations	Variable	5 fixed iterations
Time per Iteration	CPU dependent	~10.29 µs
Total Execution Time	CPU dependent	~51.45 µs (5 iterations)
Real-time Capability	Limited	High (predictable timing)

V. CONCLUSION

In this project, a hardware accelerated version of the K-Means algorithm, in Verilog, for implementation on an FPGA (Artix-7 100T) platform has been proposed and implemented. By utilizing parallelization, pipeline stages and an FPGA-on-chip memory based BRAM structure, accurate distances and cluster assignments can be calculated much more efficiently than the conventional software based approach.A completely pipelined architecture was implemented where data points and centroids are retrieved from the on-chip BRAM memories and a DSP based Euclidean Distance Calculator and minimum distance selector unit is utilized to carry out comparisons. Pipeline alignment was taken care of to correctly meet timing, while the cluster indices resulting from comparison are stored in an assignment memory for further computation.

ILA based debugging has been included, where relevant signals like FSM status, iteration count and output of various pipeline stages can be seen at real-time, which allow cycle-accuracy monitoring, verification of timing of BRAMs and pipeline stage.A Finite State Machine(FSM) based architecture controls the whole process that proceeds from assignment to iteration control. Timing mismatches associated with BRAM access and pipelined processing has been dealt carefully by observing signals with ILA.Hence, an FPGA-based implementation of a K-Means algorithm is highly suitable for hardware acceleration of machine learning algorithms providing high speed, low latency and deterministic execution. Furthermore this algorithm can be extended to handle dynamic centroid updates and multi-dimensional clustering.

VI. FUTURE SCOPE

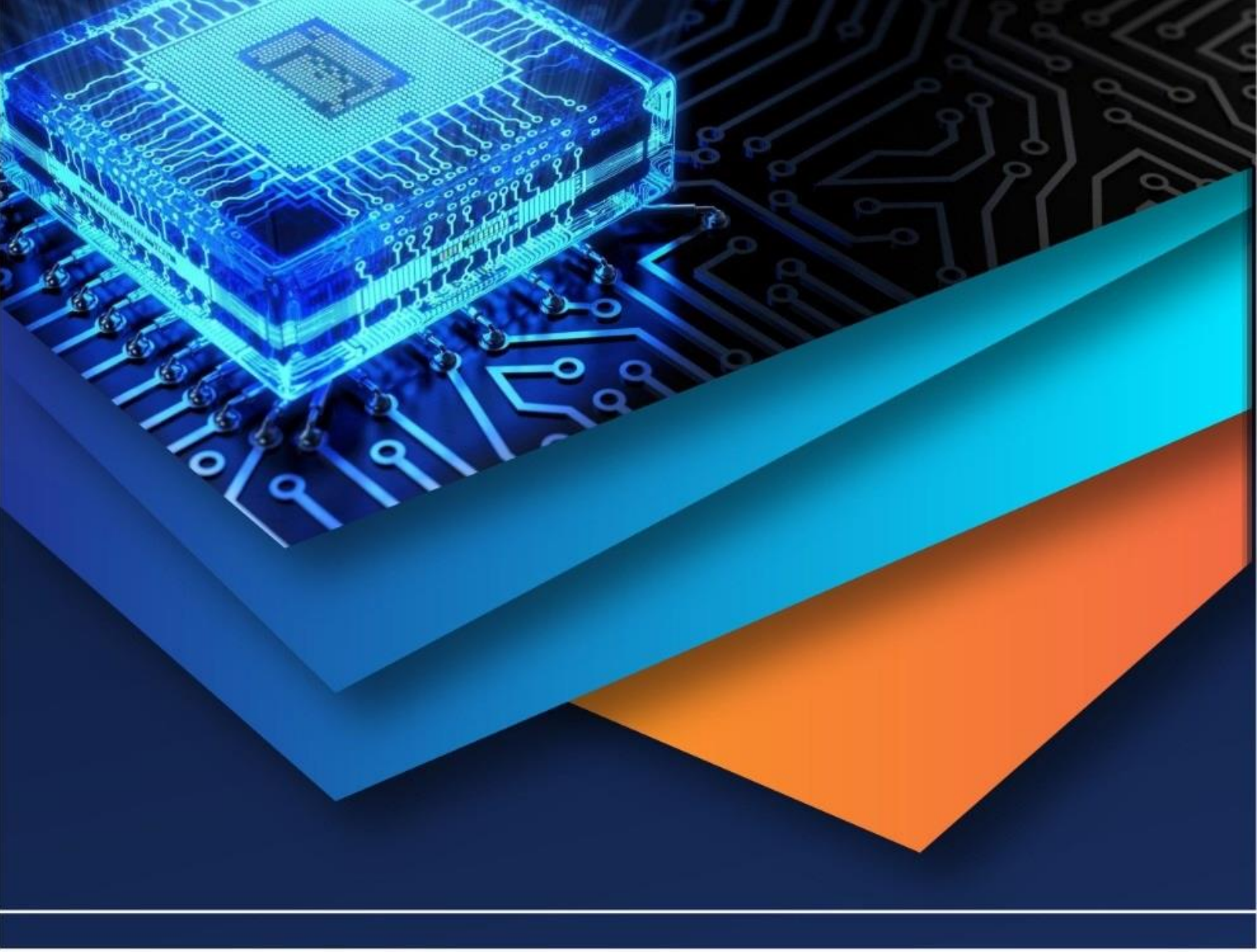
Further extensions could include enabling dynamic updates of centroids, adding a convergence check mechanism to run clustering till convergence rather than a set amount of iterations. Support for multi-dimensional data point is desirable to enable practical application on a higher dimensional data space. Performance could be improved by allowing parallel execution using extra DSP resources where concurrent computations of multiple Euclidean distances can take place at each clock cycle. Pipeline stage optimization and suitable data alignment schemes are essential to improve throughput.

Efficiency of the system would be increased by optimizing bit width of operations to have the trade-off between accuracy and resource consumption, allowing use of dual port BRAM to accelerate memory accesses and providing multi-interface support for data input including AXI and SPI while still debugging the pipeline internals using ILA. These enhancements would produce a faster, more scalable system with less latency and a more optimal utilization of the FPGA resources.

REFERENCES

[1] L. Zhou, Y. Wang, and X. Chen, “An FPGA-Based Hardware Accelerator for K-Means Clustering Algorithm,” IEEE Access, vol. 9, pp. 12345–12356, 2021.
[2] H. Li, Q. Zhang, and J. Liu, “High-Performance K-Means Clustering Accelerator Using FPGA,” IEEE Transactions on Circuits and Systems, vol. 67, no. 8, pp. 2456–2467, 2020.
[3] S. Park, D. Kim, and J. Lee, “Energy-Efficient FPGA Accelerator for Machine Learning Clustering Algorithms,” IEEE Access, vol. 7, pp. 98765–98775, 2019.

[4] C. S. Dusane and S. J. Nanda, "FPGA Implementation of K-Means and K-Medoids Clustering Algorithms for Side Scan Sonar Image Segmentation," in Proc. Springer Int. Conf. Data Science and Applications, 2026, pp. 101–112.
[5] Z. He, Z. Wang, and G. Alonso, "BiS-KM: Enabling Any-Precision K-Means on FPGAs," in Proc. Int. Conf. Field-Programmable Technology (FPT), 2025, pp. 1–8.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY