

Parallel Computing: Offloading Matrix Multiplication to GPU

Authors : Bharath R., Tharun Sai N., Bhuvan G.

Abstract : This project focuses on developing a Parallel Computing method aimed at optimizing matrix multiplication through GPU acceleration. Addressing algorithmic challenges, GPU programming intricacies, and integration issues, the project aims to enhance efficiency and scalability. The methodology involves algorithm design, GPU programming, and optimization techniques. Future plans include advanced optimizations, extended functionality, and integration with high-level frameworks. User engagement is emphasized through user-friendly interfaces, open-source collaboration, and continuous refinement based on feedback. The project's impact extends to significantly improving matrix multiplication performance in scientific computing and machine learning applications.

Keywords : matrix multiplication, parallel processing, cuda, performance boost, neural networks

Conference Title : ICPDS 2024 : International Conference on Parallel and Distributed Systems

Conference Location : Paris, France

Conference Dates : May 16-17, 2024