

Binarized Neural Networks (BNNs) can significantly reduce the inference latency and energy consumption in resource-constrained devices due to their pure-logi