

TABLE OF CONTENTS

TABLE OF CONTENTS	1
LIST OF TABLES	5
LIST OF FIGURES AND ALGORITHMS	6
ABSTRACT	8
CHAPTER I INTRODUCTION	9
1.1 RESEARCH BACKGROUND	9
1.2 RESEARCH PROBLEM	12
1.3 RESEARCH SCOPE	12
1.4 RESEARCH OBJECTIVES	13
1.5 RESEARCH BENEFITS	13
CHAPTER II LITERATURE REVIEW	14
CHAPTER III THEORETICAL BASIS	19
3.1. TENSORS & COMPUTATIONAL GRAPH	19
3.2. CONVOLUTIONAL NEURAL NETWORK (CONVNET/CNN)	20
3.2.1. <i>Layers and Operations in Convolutional Neural Network</i>	20
3.2.2. <i>LeNet Model</i>	25
3.2.3. <i>ResNet Model</i>	25
3.3. NUMERICAL DATA TYPES	26
3.4. QUANTIZATION	27
3.4.1. <i>Linear Quantization</i>	27
3.4.2. <i>Quantization Granularity</i>	29
3.4.3. <i>Quantization-aware Training (QAT)</i>	29
3.5. PRUNING	30
3.5.1. <i>Structural Pruning</i>	30
3.5.2. <i>Pruning Criterion</i>	32
3.5.3. <i>Iterative Pruning</i>	35
3.6. FAULT AND FAULT INJECTION	36
3.6.1. <i>Evaluation Metrics</i>	38
CHAPTER IV RESEARCH METHODOLOGY	40
4.1. RESEARCH DESCRIPTION	40
4.2. BASE MODELS	41
4.2.1. <i>Dataset</i>	41
4.2.2. <i>Model Preparation</i>	43
4.2.3. <i>Training Hyperparameters</i>	45
4.2.4. <i>Quantization</i>	45
4.3. GROUP IMPORTANCE SCORING	46
4.4. ITERATIVE PRUNING & FINE-TUNING	47
4.5. EVALUATION	48
CHAPTER V IMPLEMENTATION	53

5.1.	SOFTWARE SPECIFICATION.....	53
5.2.	FULL-PRECISION AND QUANTIZED MODELS IMPLEMENTATION.....	53
5.2.1.	<i>Implementation of LeNet-5 Architecture.....</i>	54
5.2.2.	<i>Implementation of ResNet-18 Architecture.....</i>	54
5.2.3.	<i>Implementation of ResNet-50 Architecture.....</i>	55
5.3.	IMPLEMENTATION OF FULL-PRECISION TRAINING AND QUANTIZATION-AWARE TRAINING (QAT)	55
5.4.	IMPLEMENTATION OF GROUP IMPORTANCE SCORING	56
5.5.	IMPLEMENTATION OF PRUNING AND FINE-TUNING LOOPS	62
5.6.	IMPLEMENTATION OF EVALUATION TOOLS.....	66
CHAPTER VI RESULT AND DISCUSSION		71
6.1.	TRAINING IN FULL-PRECISION AND WITH QUANTIZATION-AWARE TRAINING.....	72
6.2.	MAGNITUDE-BASED AND ACTIVATION-BASED PRUNING	73
6.2.1.	Precision Impact on Pruning.....	74
6.2.2.	Scoring Method Comparative Analysis.....	75
6.3.	FAULT INJECTION EXPERIMENTS.....	77
6.3.1.	Number of Fault Injections to Perform.....	78
6.3.2.	Fault Injection Results on LeNet-5, ResNet18, and ResNet50.....	78
A.	Layer Positioning.....	82
B.	Bit Position Vulnerability.....	85
6.4.	FAILURE IN TIME (FIT) RATE	84
CHAPTER VII CONCLUSION AND SUGGESTION		86
APPENDIX.....		87
REFERENCES		87101

LIST OF TABLES

Table 4.1: Dataset Splitting.....	35
Table 4.2: LeNet-5 Architecture.....	35
Table 4.3: ResNet-18 Architecture.....	36
Table 4.4: ResNet-50 Architecture.....	36
Table 4.5: Training Hyperparameters.....	37
Table 4.6: QAT Hyperparameters.....	38
Table 4.7: Aggregation and Normalization Formulas.....	39
Table 4.8: Pruning Hyperparameters.....	40
Table 4.9: Possible Injection Sites for Quantized Models.....	43
Table 6.1 LeNet5 Training QAT Result.....	81
Table 6.2 ResNet Training QAT Result.....	81
Table 6.3: Pruning Performance Result.....	82
Table 6.4: Statistical Fault Injection.....	85
Table 6.5: Software Fault Injection Results.....	86

LIST OF FIGURES

Figure 3.1: Tensors and Simple Computational Graph.....	15
Figure 3.2: LeNet-5 Architecture.....	16
Figure 3.3: Fully-connected Layer.....	17
Figure 3.4: 2D Convolutional Operation on Tensors.....	17
Figure 3.5: Pooling Operations.....	19
Figure 3.6: Basic and Bottleneck Blocks.....	21
Figure 3.7: Two's Complement & Fixed-Point Representation.....	22
Figure 3.8: Floating-point Representation (IEEE-754).....	22
Figure 3.9: Linear Quantization.....	24
Figure 3.10: Quantization-Aware Training Process.....	26
Figure 4.1: Research Flowchart.....	33
Figure 4.2: MNIST & CIFAR-10 Datasets Example.....	34
Figure 4.3: Scoring Metrics Workflow.....	37
Figure 4.4: Pruning Pipeline.....	37
Figure 4.5: Fault Injection Flowchart.....	41
Figure 4.6: Fault Injection in the Dequantizer Nodes.....	42
Figure 5.1: Implementation Workflow.....	53
Figure 5.2: GroupActivationImportance.....	56
Figure 5.3: QuantConv Pruner.....	63
Figure 5.4: QuantLinear (FC) Pruner.....	64
Figure 5.5: Pruner Main Function.....	65
Figure 5.6: Graph Interpreter.....	66
Figure 5.7: Node Executor.....	67
Figure 5.8: Fault Injection Functions.....	68
Figure 6.1: Accuracy to Pruning Ratio LeNet-5.....	77
Figure 6.2: Accuracy to Pruning Ratio ResNet-50.....	78
Figure 6.3: Accuracy to Pruning Ratio ResNet-18.....	78
Figure 6.4: g Distribution in FP32 Models.....	79
Figure 6.5: g Distribution in INT8 Models.....	79
Figure 6.4: ResNet50 Layer Vulnerability.....	84
Figure 6.5: ResNet18 Layer Vulnerability.....	84
Figure 6.6: LeNet-5 Layer Vulnerability.....	85
Figure 6.7: LeNet-5 Bit Vulnerability.....	85
Figure 6.8: ResNet18 Bit Vulnerability.....	86
Figure 6.9: ResNet50 Bit Vulnerability.....	86
Figure 6.10: Accelerator_FIT_Rate Values Across Models.....	87

LIST OF ALGORITHMS

Algorithm 1: Dependency Graph.....	27
Algorithm 2: Grouping.....	27
Algorithm 3: Assignment of group L_p -based importance scores.....	28
Algorithm 4: Assignment of neuron activation-based scores.....	28
Algorithm 5: Iterative Pruning.....	29