

## CONTENTS

<b>Halaman Pengesahan</b>	<b>ii</b>
<b>Halaman Pernyataan</b>	<b>iii</b>
<b>PREFACE</b>	<b>iv</b>
<b>CONTENTS</b>	<b>v</b>
<b>LIST OF FIGURES</b>	<b>ix</b>
<b>LIST OF TABLES</b>	<b>xi</b>
<b>ABSTRACT</b>	<b>xii</b>
<b>I Introduction</b>	<b>1</b>
1.1 Research Background . . . . .	1
1.2 Research Problem . . . . .	2
1.3 Research Goals and Objectives . . . . .	2
1.4 Research Scope and Limitations . . . . .	3
1.5 Benefits of the Research . . . . .	3
1.6 Research Writing System . . . . .	3
<b>II Literature Review</b>	<b>5</b>
2.1 Feature Extraction Methods . . . . .	5
2.2 Modeling Speech Emotion Recognition . . . . .	6
<b>III Theoretical Basis</b>	<b>8</b>
3.1 Feature Extraction Techniques . . . . .	8
3.1.1 Mel-Spectrogram . . . . .	8
3.1.2 Mel-Frequency Cepstral Coefficients . . . . .	9
3.2 Data Analysis Methods . . . . .	10
3.2.1 Classical Multidimensional Scaling . . . . .	10
3.2.2 Linear Discriminant Analysis . . . . .	10
3.3 Deep Learning Methods . . . . .	11
3.3.1 1-D Convolutional Neural Networks . . . . .	11

3.3.2	Bi-Directional Gated Recurrent Unit . . . . .	13
3.3.3	Attention Mechanism . . . . .	14
3.4	Full Model Selection . . . . .	14
3.4.1	Overview . . . . .	14
3.4.2	Particle Swarm Optimization . . . . .	14
3.5	Model Evaluation Metrics . . . . .	16
3.5.1	Metrics . . . . .	16
3.5.2	Effect Size Analysis . . . . .	17
3.5.3	ANOVA Test . . . . .	17
<b>IV</b>	<b>Methodology</b>	<b>19</b>
4.1	Pipeline . . . . .	19
4.2	Dataset Selection and Preparation . . . . .	20
4.3	Feature extraction . . . . .	21
4.4	Data Analysis and Feature Selection . . . . .	23
4.4.1	Plotting Emotions . . . . .	23
4.4.2	Feature Selection . . . . .	24
4.5	Model architectures . . . . .	24
4.5.1	Common Architectural Elements . . . . .	25
4.5.2	1D CNN Model . . . . .	25
4.5.3	Bi-GRU Model . . . . .	26
4.5.4	Fusion Model . . . . .	27
4.6	Model Training . . . . .	28
4.7	Full Model Selection . . . . .	29
4.8	Model Evaluation . . . . .	31
<b>V</b>	<b>Implementation</b>	<b>33</b>
5.1	Data Processing Pipeline . . . . .	33
5.2	Feature Extraction Implementation . . . . .	34
5.3	Data Analysis . . . . .	35
5.3.1	Classical Multidimensional Scaling . . . . .	35
5.3.2	Linear Discriminant Analysis and Feature Importance . . . . .	36
5.4	Model Implementation . . . . .	36
5.4.1	Common Building Blocks . . . . .	36
5.4.2	Convolutional Neural Network Model . . . . .	37
5.4.3	Bi-Directional Gated Recurrent Unit Model . . . . .	38

5.4.4	Fusion Model . . . . .	39
5.5	Training Infrastructure . . . . .	40
5.5.1	Data Loading . . . . .	40
5.5.2	Training Loop . . . . .	43
5.5.3	Validation Loop . . . . .	44
5.5.4	Training Execution . . . . .	44
5.6	Particle Swarm Optimization Implementation . . . . .	45
5.6.1	Search Space Definition . . . . .	45
5.6.2	Particle Representation . . . . .	45
5.6.3	Fitness Evaluation . . . . .	46
5.6.4	Velocity and Position Updates . . . . .	47
5.6.5	Swarm Management . . . . .	48
5.6.6	Results Tracking and Analysis . . . . .	49
5.7	Evaluation Framework . . . . .	49
5.7.1	Performance Metrics . . . . .	49
5.7.2	Statistical Validation . . . . .	50
5.7.3	Computational Efficiency Tracking . . . . .	50
<b>VI</b>	<b>Results and Discussion</b>	<b>51</b>
6.1	Data Analysis Results . . . . .	51
6.1.1	Classical Multidimensional Scaling Results . . . . .	51
6.1.2	Linear Discriminant Analysis Results . . . . .	53
6.2	Speech Emotion Recognition Results . . . . .	55
6.2.1	Training and Evaluation Metrics . . . . .	55
6.2.2	Loss Graphs . . . . .	57
6.3	Full Model Selection Results . . . . .	58
6.4	Final Model Evaluation . . . . .	59
6.5	Statistical Comparisons . . . . .	63
6.6	Comparisons to Related Work . . . . .	65
6.7	Evaluation . . . . .	66
6.7.1	Research Objective Achievement . . . . .	66
6.7.2	Limitations of the Research . . . . .	67
<b>VII</b>	<b>Conclusion</b>	<b>69</b>
7.1	Summary of Achievements . . . . .	69
7.2	Future Works . . . . .	69



7.3	Final Thoughts . . . . .	70
	<b>References</b>	<b>71</b>
A	<b>Classification Reports</b>	<b>77</b>
B	<b>Confusion Matrices for Models</b>	<b>78</b>