

CONTENTS

ENDORSEMENT PAGE	ii
PERNYATAAN BEBAS PLAGIASI	iii
PREFACE	v
NOMENCLATURE AND ABBREVIATION	vi
ABSTRACT	vii
INTISARI	viii
CONTENTS	ix
LIST OF FIGURES	xii
LIST OF TABLES	xiii
CHAPTER I Introduction	1
1.1 Background	1
1.2 Research Problems	5
1.3 Research Novelty	5
1.4 Research Objective	7
1.5 Research Benefits	8
1.6 Thesis Outline	8
CHAPTER II Literature Review and Theoretical Basis	9
2.1 Literature Review	9
2.1.1 Deep Learning in Eye Movement Classification	9
2.1.2 Tree-structured Parzen Estimator (TPE) as Hyperparameter Optimizer in Deep Learning	12
2.2 Theoretical Basis	14
2.2.1 Eye Movement Classification	14
2.2.2 Feature Extraction	15
2.2.2.1 Temporal Scale	16
2.2.2.2 Speed	17
2.2.2.3 Direction	17
2.2.2.4 Displacement	18
2.2.2.5 Standard Deviation	18
2.2.2.6 Acceleration	18
2.2.3 Windowing Method	19
2.2.4 Deep Learning	20
2.2.4.1 One-Dimensional Convolutional Neural Network (1D-CNN)	21
2.2.4.2 Long Short-Term Memory (LSTM)	22
2.2.4.3 Hybrid 1D-CNN–BiLSTM Architecture	24
2.2.4.4 Temporal Convolutional Network (TCN)	24
2.2.5 Hyperparameters	26
2.2.5.1 Architectural Hyperparameters	26
2.2.5.2 Flow Control Hyperparameters	28
2.2.6 Hyperparameter Optimization (HPO)	29
2.2.6.1 Tree-structured Parzen Estimator (TPE)	30

2.2.7	Model Evaluation	31
2.2.7.1	Leave-One-Video-Out (LOVO)	32
2.2.7.2	F1-Score	32
2.2.8	Statistical Test	34
2.2.8.1	Shapiro-Wilk Test	34
2.2.8.2	Repeated Measures ANOVA	35
2.2.8.3	Friedman Test	35
2.2.8.4	T-Test	36
2.2.8.5	Wilcoxon Test	36
2.2.8.6	One-Tailed Test	36
2.3	Hypotheses	37
CHAPTER III Methodology		38
3.1	Tools and Materials	38
3.1.1	Tools	38
3.1.2	Materials	39
3.2	Research Flow	40
3.3	Experiment Design	40
3.3.1	Data Preprocessing	42
3.3.2	Deep Learning Architecture	44
3.3.3	Hyperparameter Optimization Specification	45
3.3.4	Statistical Analysis	47
CHAPTER IV Results and Discussion		50
4.1	Results	50
4.1.1	Replication Results	50
4.1.1.1	Fixation Score	51
4.1.1.2	Statistical Comparison of Fixation Scores	52
4.1.1.3	Saccades Score	54
4.1.1.4	Statistical Comparison of Saccades Scores	55
4.1.1.5	Smooth Pursuit Score	56
4.1.1.6	Statistical Comparison of Smooth Pursuit Scores ..	58
4.1.1.7	Noise Score	59
4.1.1.8	Statistical Comparison of Noise Scores	61
4.1.1.9	Macro F1 Score	62
4.1.1.10	Statistical Comparison of Macro F1 Scores	64
4.1.1.11	Replication Score Summary	65
4.1.2	Hyperparameter Optimization Result	67
4.1.2.1	Comparison of Baseline and Optimized Hyperpa- rameters	67
4.1.2.2	Fixation Score	68
4.1.2.3	Saccades Scores	70
4.1.2.4	Smooth Pursuit Score	72
4.1.2.5	Noise Score	74
4.1.2.6	Macro F1-score	76
4.1.3	Comparison Summary	78

4.2	Discussion	79
4.2.1	Algorithm Comparison	79
4.2.2	The Importance of HPO	80
4.2.3	Challenges in Noise Classification	82
4.2.4	Ablation Study on Feature Contributions	83
4.2.5	Limitations	84
4.2.5.1	Computational Constraints and Experimental Scope	84
4.2.5.2	Justification and Limitations of the Hyperparameter Search Space	85
4.2.5.3	Model Efficiency and Deployment Considerations	85
4.2.5.4	Generalization	86
4.2.6	Practical Implication	86
CHAPTER V	Conclusion	88
5.1	Conclusions	88
5.2	Future Works	89
References	91

LIST OF FIGURES

Figure 1.1	Eye movement examples, with green-colored sample being fixation, orange-colored being saccades and blue-colored being smooth pursuit.	2
Figure 3.1	Flowchart of the research methodology.	41
Figure 3.2	Illustration of how temporal scale functioned.....	42
Figure 3.3	Example of extracted features.	43
Figure 3.4	Illustration of windowing method.	43
Figure 3.5	Architectural schemes of (a) 1D-CNN, (b) BiLSTM, (c) 1D-CNN–BiLSTM [19], and (d) TCN [18].	46
Figure 4.1	Comparison of Overall F1 Scores for Fixation, Saccades, Smooth Pursuit, Noise, and Macro F1 Score among Different Deep Learning Algorithm.....	66

LIST OF TABLES

Table 1.1	Comparison of prior EMC studies and the positioning of this research	6
Table 3.1	Snippet of raw gaze data entries from the GazeCom dataset. ...	39
Table 3.2	Summary statistics of annotated eye-movement classes in the GazeCom dataset. Percentages are computed w.r.t. the total number of rows (4,318,056) and total number of events (411,172).	40
Table 3.3	Optimized hyperparameters for TCN using Optuna.	45
Table 3.4	Auxiliary hypotheses	47
Table 4.1	Replication result fixation score	51
Table 4.2	Wilcoxon signed-rank test results for fixation score comparisons.	53
Table 4.3	Replication results for saccade classification (Macro-F1 score).	54
Table 4.4	Wilcoxon signed-rank test results for saccade score comparisons	56
Table 4.5	Replication result smooth pursuit score	57
Table 4.6	One-tailed paired <i>t</i> -test results for smooth pursuit score comparisons	59
Table 4.7	Replication results for noise classification F1-scores	60
Table 4.8	Wilcoxon signed-rank test results for noise F1-score comparisons	62
Table 4.9	Replication results: macro F1-scores across all eye-movement classes	63
Table 4.10	Pairwise one-tailed <i>t</i> -test results for macro F1 scores	65
Table 4.11	Hyperparameter Comparison	68
Table 4.12	Fixation F1-score comparison between baseline and optimized TCN	69
Table 4.13	Saccade F1-score (%) comparison between baseline and optimized TCN	70
Table 4.14	Smooth Pursuit F1-score comparison between original and optimized TCN	72
Table 4.15	Noise F1-score (%) comparison between baseline and optimized TCN	74
Table 4.16	Macro F1-score (%) comparison between Elmadjian TCN and TPE-optimized TCN	76
Table 4.17	Ablation study on the contribution of gaze features to the performance of the optimized model (F1-score in %).	83