UNIVERSITAS GADJAH MADA
Kerangka Kerja Pembelajaran Penguatan Mendalam yang Dipandu oleh Pembelajaran Kurikulum untuk
Masalah Pemuatan Multi-Kontainer 3D dengan Kendala Kesatuan Keluarga dan Penyeimbangan Beban
Kreshnayogi Dava Berliansyach, Muhammad Ilham Amrizal, B.Eng., M.I.S., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

CONTENTS

ENDORSEMENT PAGE ii
ACKNOWLEDGEMENTS iii
CONTENTS iv
LIST OF TABLES vii
LIST OF FIGURES viii
ABSTRAK ix
ABSTRACT x
CHAPTER I Introduction 1
1.1 Research Background 1
1.2 Research Problem 3
1.3 Research Objective 3
1.4 Research Scope 4
1.5 Research Significance 4
1.6 Thesis Structure 5
CHAPTER II Literature Review 6
2.1 Three-Dimensional Container Loading Problem 6
2.2 Evolution of 3D-CLP Solution Methods 8
2.3 Reinforcement Learning for 3D-CLP 10
2.3.1 Foundational RL Approaches to 3D Packing 10
2.3.2 Transformer-Based Spatial Reasoning 10
2.4 Curriculum Learning in Reinforcement Learning 11
CHAPTER III Theorectical Framework 14
3.1 The Three-Dimensional Multi-Container Loading Problem (3D-MCLP) 14
3.1.1 Decision Variables and Objective Functions 14
3.1.2 Mathematical Constraints 16
3.2 Reinforcement Learning Theory 17
3.2.1 Markov Decision Processes (MDP) 17
3.2.2 Value Functions and Policy Representations 19
3.2.3 Policy Gradient Methods 20
3.2.4 Actor–Critic Algorithms 20
3.2.5 Proximal Policy Optimization (PPO) 21
3.2.6 Transformer Architectures for Spatial Reasoning 22
3.3 Curriculum Learning in Reinforcement Learning 23
3.3.1 The Concept of Curriculum Learning 23
3.3.2 Curriculum for Container Packing 24
3.3.3 Mathematical Impact on Optimization 25

UNIVERSITAS GADJAH MADA
Kerangka Kerja Pembelajaran Penguatan Mendalam yang Dipandu oleh Pembelajaran Kurikulum untuk
Masalah Pemuatan Multi-Kontainer 3D dengan Kendala Kesatuan Keluarga dan Penyeimbangan Beban
Kreshnayogi Dava Berliansyach, Muhammad Alfian Amrizal, B.Eng., M.I.S., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/.

CHAPTER IV Research Methodology 26
4.1 Conceptual Framework 26
4.2 MDP Formulation for Family Unity 28
4.2.1 State Space Representation ($\mathcal{S}$) 28
4.2.2 Action Space Decomposition ($\mathcal{A}$) 31
4.2.3 Constraint Enforcement via Action Masking 32
4.2.4 Composite Reward Function ($\mathcal{R}$) 33
4.3 Transformer Network Architecture 36
4.3.1 Encoders 36
4.3.2 Decoders 37
4.3.3 Value Network (The Critic) 39
4.4 Optimization Algorithm: Proximal Policy Optimization 39
4.4.1 PPO Objective Function 40
4.4.2 Total Loss Formulation 40
4.5 Curriculum Learning Strategy 41
4.6 Dataset 42
4.6.1 Procedural Training Dataset (The Learning Environment) 42
4.6.2 Evaluation Datasets 43
4.7 Experimental Setup 43
4.7.1 Implementation Environment 43
4.7.2 Hyperparameter and Reward Configuration 44
4.7.3 Baselines for Comparison 45
4.7.4 Evaluation Metrics 46
CHAPTER V IMPLEMENTATION 49
5.1 Codebase Architecture and Data Flow 49
5.2 Procedural Data Representation 49
5.3 Container State Downsampling 51
5.4 Action Space Masking 51
5.5 Hardware-Specific Algorithmic Adaptations 52
5.5.1 CPU-Isolation of the Physics Simulator 52
5.5.2 Scatter-Free Neural Network Workarounds 52
5.5.3 Dimensionality Limits and VRAM Optimization 52
5.5.4 Dynamic Item Slicing 53
5.6 Computational Scope and Benchmarking Justification 53
CHAPTER VI Results and Discussion 54
6.1 Training Convergence and Analysis 54
6.2 Validation and Policy Generalization 56
6.3 Real-life Performance Evaluation 59
6.4 Qualitative Analysis and 3D Visualization 62

UNIVERSITAS
GADJAH MADA
**Kerangka Kerja Pembelajaran Penguatan Mendalam yang Dipandu oleh Pembelajaran Kurikulum untuk
Masalah Pemuatan Multi-Kontainer 3D dengan Kendala Kesatuan Keluarga dan Penyeimbangan Beban**
Kreshnayogi Dava Berliansyach, Muhammad Alfian Amrizal, B.Eng., M.I.S., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

CHAPTER VII Conclusion and Future Work 65
7.1 Conclusion 65
7.2 Future Work 65
REFERENCES 67

Kerangka Kerja Pembelajaran Penguatan Mendalam yang Dipandu oleh Pembelajaran Kurikulum untuk
Masalah Pemuatan Multi-Kontainer 3D dengan Kendala Kesatuan Keluarga dan Penyeimbangan Beban
Kreshnayogi Dava Berliansyach, Muhammad Alfian Amrizal, B.Eng., M.I.S., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

LIST OF TABLES

Table 2.1 Definitions of Operational Constraints in 3D-CLP 7
Table 2.2 Comparative Analysis of Operational Constraints in Key Literature . 8

Table 3.1 Decision variables for the 3D-MCLP formulation. 15

Table 4.1 The five-phase state-complexity curriculum schedule. 41
Table 4.2 Software stack and core libraries utilized. 44
Table 4.3 Specifications of hardware for experimentation. 44
Table 4.4 Hyperparameters for training. 45
Table 4.5 Reward and curriculum-specific hyperparameters. 45

Table 6.1 Global Average Performance 60

Kerangka Kerja Pembelajaran Penguatan Mendalam yang Dipandu oleh Pembelajaran Kurikulum untuk
Masalah Pemuatan Multi-Kontainer 3D dengan Kendala Kesatuan Keluarga dan Penyeimbangan Beban
Kreshnayogi Dava Berliansyah, Wiwin Sulistyo, S.T., M.Sc., Muhammad Alfian Amrizal, B.Eng., M.I.S., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

LIST OF FIGURES

Figure 1.1 Loading families of boxes into a fleet. 2
Figure 3.1 Feasible orientations of item i placed at (x_i, y_i, z_i) 14
Figure 3.2 Example of Height Map and Weight Map representations. 18
Figure 3.3 A conceptual Actor–Critic loop for policy optimization. 21
Figure 3.4 Curriculum learning stages in reinforcement learning. 24
Figure 4.1 3-Point Support Visualization 26
Figure 4.2 High-Level Deep Reinforcement Learning Framework Overview .. 27
Figure 4.3 The Physical Geometry and Spatial Reasoning State 29
Figure 4.4 Illustration of Family Occupancy Vector ($\mathbf{m}_k$) 30
Figure 4.5 Top-down Illustration of the Container Center of Gravity ($\mathbf{g}_k$) 30
Figure 4.6 The Container-First action hierarchy. 32
Figure 4.7 The Proposed State Encoder 36
Figure 4.8 The Proposed Decoding Process 38
Figure 4.9 Problem Instance Generator Flow Chart 42
Figure 4.10 Family Utilization Ratio (FUR) Calculation Illustration 47
Figure 5.1 Codebase architecture and runtime data flow. 49
Figure 5.2 Example of Generated Data 50
Figure 6.1 Raw episodic training rewards across the five configurations 54
Figure 6.2 Value Loss (Critic Network) across 300 Epochs (Isolating Hard-Constrained Models) 55
Figure 6.3 Policy Loss (Actor Network) across 300 Epochs (Isolating Hard-Constrained Models) 56
Figure 6.4 Validation Total Reward performance on static $N = 50$ instances .. 57
Figure 6.5 Validation Family Utilization Ratio (FUR) across Models 57
Figure 6.6 Validation Container Cost minimization across Models 58
Figure 6.7 Validation Load Balance Deviation across Models 59
Figure 6.8 Family Utilization Ratio (FUR) across escalating relational complexity 61
Figure 6.9 Container Cost across escalating relational complexity 61
Figure 6.10 Load Balance Deviation (LBD) across escalating relational complexity 62
Figure 6.11 3D Visualization Grid comparing packing configurations 63