UNIVERSITAS
GADJAH MADA
Reinforcement Learning dengan Representasi Visual Fitur Acak dan Ruang Laten Pada Lingkungan Atari
NOVANTIO BANGUN HERMAWAN, Dr. Yohanes Suyanto, M.I.Kom
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR ISI

Halaman Judul **ii**

Halaman Pengesahan **iii**

Halaman Pernyataan **iv**

PRAKATA **v**

INTISARI **xii**

ABSTRACT **xiii**

I PENDAHULUAN **1**
1.1 Latar Belakang Masalah 1
1.2 Perumusan Masalah 3
1.3 Batasan Masalah 3
1.4 Tujuan Penelitian 4
1.5 Manfaat Penelitian 4
1.6 Kontribusi Penelitian 5
1.7 Keaslian Tesis 5

II TINJAUAN PUSTAKA **6**

III METODOLOGI PENELITIAN **10**
3.1 Deskripsi Umum Penelitian 10
3.2 Alat dan Bahan Penelitian 11
3.2.1 Perangkat Keras 11
3.2.2 Perangkat Lunak 11
3.3 Environment Penelitian 12
3.4 Rancangan Usulan Penelitian 13
3.4.1 Pra-pemrosesan Observasi Visual 14
3.4.2 Proses Representasi Fitur Visual 16
3.4.3 Gambaran Umum Arsitektur Frozen Random Intrinsic . 18
3.4.4 Konfigurasi Pelatihan Agen 18
3.5 Rancangan Pengujian dan Evaluasi 19

UNIVERSITAS
GADJAH MADA
Reinforcement Learning dengan Representasi Visual Fitur Acak dan Ruang Laten Pada Lingkungan Atari
NOVANTIO BANGUN HERMAWAN, Dr. Yohanes Suyanto, M.I.Kom
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

3.5.1 Analisis Kurva Pembelajaran (Learning Curve) 20
3.5.2 Analisis Performa Agen 21
3.5.3 Analisis Penggunaan Sumber Daya Komputasi 22
3.5.4 Pengujian Komparasi Model 22
3.5.5 Pengujian Ablation Study 24
3.5.6 Metrik Evaluasi 24
3.6 Arsitektur Model 25
3.6.1 CNN Baseline 25
3.6.2 CNN + RND 26
3.6.3 CNN + ICM 28
3.6.4 CNN + LSTM 29
3.6.5 CNN + Transformer 30
3.6.6 Vision Transformer (ViT) 32
3.7 Strategi Pelatihan Frozen Random Instrinsic(FRI) 34
3.7.1 Metode Pelatihan FRI 34
3.7.2 Arsitektur FRI 35
3.7.3 Justifikasi Jumlah Langkah Tiap Fase 37
3.7.4 Mekanisme Pencegahan Interferensi Gradien 38
3.7.5 Rasionalisasi Pemilihan Model Pembanding 38

IV IMPLEMENTASI **40**
4.1 Implementasi Environment 40
4.2 Implementasi PPO 41
4.2.1 Inisialisasi Model PPO 41
4.3 Implementasi CNN 42
4.4 Implementasi Frozen CNN 45
4.4.1 Pembekuan Parameter CNN (*Frozen CNN*) 45
4.5 Implementasi RND 47
4.5.1 Jaringan Dasar RND 48
4.5.2 Modul RND Lengkap 49
4.6 Implementasi ICM 52
4.6.1 Encoder ICM 52
4.6.2 Forward Model dan Inverse Model 53
4.7 Implementasi Usulan RND + ICM 56
4.7.1 Arsitektur Frozen Random Intrinsic 56

UNIVERSITAS
GADJAH MADA
Reinforcement Learning dengan Representasi Visual Fitur Acak dan Ruang Laten Pada Lingkungan Atari
NOVANTIO BANGUN HERMAWAN, Dr. Yohanes Suyanto, M.I.Kom
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

4.7.2 Normalisasi Welford untuk Intrinsic Reward 58
4.7.3 Callback RND + ICM Paralel 60
4.7.4 Fungsi Pelatihan Lengkap 62
4.8 Implementasi Pembanding CNN + LSTM 65
4.8.1 Arsitektur CnnLstmExtractor 65
4.8.2 Konfigurasi Policy CNN + LSTM 67
4.9 Implementasi Pembanding CNN + Transformer 68
4.9.1 Modul Self-Attention Spasial 68
4.9.2 Arsitektur CNN + Transformer 70
4.10 Implementasi Pembanding Vision Transformer 72
4.10.1 Patch Embedding 72
4.10.2 Blok ViT 74
4.10.3 Arsitektur AtariViT 76

V HASIL PENELITIAN DAN PEMBAHASAN 79
5.1 Perbandingan Model pada Breakout 79
5.1.1 Hasil Kuantitatif 79
5.1.2 Analisis Kurva Pelatihan 79
5.2 Perbandingan Model pada Pong 81
5.2.1 Hasil Kuantitatif 81
5.2.2 Analisis Kurva Pelatihan 82
5.3 Ablation Study 83
5.3.1 Ablation Study pada Breakout 83
5.3.2 Ablation Study pada Pong 85
5.4 Diskusi dan Analisis Komparatif 86
5.4.1 Keunggulan Pendekatan Frozen Random Intrinsic 86
5.4.2 Perbandingan Efisiensi Komputasi 87
5.4.3 Analisis Kegagalan Transformer pada Atari 87
5.4.4 Generalisasi antar Lingkungan 88
5.4.5 Ringkasan Temuan 88

VI KESIMPULAN DAN SARAN 89
6.1 Kesimpulan 89
6.2 Saran 90

A NASKAH PUBLIKASI 96

UNIVERSITAS
GADJAH MADA
Reinforcement Learning dengan Representasi Visual Fitur Acak dan Ruang Laten Pada Lingkungan Atari
NOVANTIO BANGUN HERMAWAN, Dr. Yohanes Suyanto, M.I.Kom
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR TABEL

2.1 Ringkasan Penelitian Terkait Reinforcement Learning dan Representasi Laten 8

3.1 Spesifikasi Perangkat Keras 12
3.2 Spesifikasi Perangkat Lunak 12
3.3 Wrapper preprocessing lingkungan Atari dan fungsinya 16
3.4 Konfigurasi hiperparameter PPO untuk seluruh model 19
3.5 Konfigurasi Model Eksperimen 23
3.6 Desain Ablation Study 24
3.7 Metrik evaluasi yang digunakan dalam penelitian 25
3.8 Parameter arsitektur CNN Baseline 26
3.9 Arsitektur jaringan RND (*target* dan *predictor*) 27
3.10 Arsitektur komponen ICM 28
3.11 Parameter arsitektur CNN + LSTM 30
3.12 Parameter arsitektur CNN + Transformer 32
3.13 Parameter arsitektur Vision Transformer (ViT) 34
3.14 Hasil eksperimen pendahuluan pemilihan koefisien pembobot *intrinsic reward*. 36

4.1 Konfigurasi Hiperparameter PPO 42
4.2 Arsitektur Jaringan RND 48

5.1 Hasil perbandingan seluruh model pada Breakout (300k langkah) 79
5.2 Hasil perbandingan seluruh model pada Pong (3M langkah) . . 81
5.3 Hasil ablation study pada Breakout 83
5.4 Hasil ablation study pada Pong 85
5.5 Perbandingan efisiensi komputasi antar model pada Breakout . 87
5.6 Ringkasan perbandingan kinerja seluruh model pada dua lingkungan 88

UNIVERSITAS
GADJAH MADA
Reinforcement Learning dengan Representasi Visual Fitur Acak dan Ruang Laten Pada Lingkungan Atari
NOVANTIO BANGUN HERMAWAN, Dr. Yohanes Suyanto, M.I.Kom
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR GAMBAR

3.1 Diagram Alur Tahapan Penelitian 11
3.2 Arsitektur Usulan Penelitian 14
3.3 Alur Visual Prapemrosesan Observasi Atari 15
3.4 Diagram Alir Perbandingan CNN Baseline vs Frozen CNN . . . 17
3.5 Pembentukan Latent Space pada PPO Random CNN 18
3.6 Contoh Kurva Pembelajaran Agen RL 21
3.7 Arsitektur CNN Baseline. 26
3.8 Alur CNN + RND: cabang kebijakan (kiri) dan cabang *intrinsic reward* RND (kanan). 27
3.9 Alur modul ICM: encoder ϕ bersama, forward model untuk *intrinsic reward*, and inverse model untuk supervisi encoder. . . 29
3.10 Arsitektur CNN + LSTM: setiap frame diproses secara terpisah oleh CNN berbagi bobot, hasilnya menjadi urutan input LSTM. 30
3.11 Arsitektur CNN + Transformer: CNN mengekstraksi fitur lokal, Transformer menangkap relasi global antar token spasial. . 31
3.12 Arsitektur Vision Transformer (ViT) untuk Atari: dari patch embedding hingga representasi akhir melalui token [CLS]. . . . 33
3.13 Alur arsitektur Frozen Random Intrinsic (FRI) pada model usulan. 35
3.14 Arsitektur model usulan CNN + RND + ICM dengan tiga komponen paralel: CNN sebagai encoder kebijakan, RND sebagai cabang kebaruan state, dan ICM sebagai cabang kebaruan transisi. 36

5.1 Kurva *rollout/ep_rew_mean* seluruh model pada Breakout selama 300k langkah pelatihan. 80
5.2 Kurva *rollout/ep_rew_mean* seluruh model pada Pong selama 3M langkah pelatihan. 82
5.3 Kurva *rollout/ep_rew_mean* ablation study pada Breakout. Garis biru tua adalah fase 1 (CNN Init), sedangkan garis lainnya adalah variasi fase 2 yang dimulai dari checkpoint fase 1. 84

UNIVERSITAS
GADJAH MADA
Reinforcement Learning dengan Representasi Visual Fitur Acak dan Ruang Laten Pada Lingkungan Atari
NOVANTIO BANGUN HERMAWAN, Dr. Yohanes Suyanto, M.I.Kom
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

5.4 Kurva *rollout/ep_rew_mean* ablation study pada Pong. Garis biru tua adalah fase 1, sedangkan garis lainnya merupakan variasi fase 2 yang dimulai dari checkpoint yang sama. 85