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/

INTISARI

Reinforcement Learning dengan Representasi Visual Fitur Acak dan Ruang Laten pada Lingkungan Atari

Oleh

NOVANTIO BANGUN HERMAWAN
25/563864/PPA/07096

Pada pembelajaran *end-to-end*, CNN berperan ganda sebagai ekstraktor fitur visual sekaligus penentu kebijakan, sehingga representasi visual terus berubah mengikuti pembaruan kebijakan. Fitur yang tidak stabil ini menurunkan efektivitas eksplorasi dan membuat agen mudah terjebak pada optimum lokal.

Penelitian ini memisahkan pembelajaran visual dari pembelajaran kebijakan melalui arsitektur *Frozen Random Intrinsic* (FRI), yang menggabungkan *frozen* CNN, *Random Network Distillation* (RND), dan *Intrinsic Curiosity Module* (ICM) dalam kerangka *Proximal Policy Optimization* (PPO). CNN dilatih lebih dahulu lalu dibekukan sehingga representasi visual bersifat tetap, sementara RND (kebaruan *state*) dan ICM (kebaruan transisi) berjalan paralel sepanjang pelatihan sebagai sumber *reward* intrinsik. *Reward* dari kedua cabang dinormalisasi dengan metode Welford dan digabungkan dengan *reward* ekstrinsik sebagai masukan bagi PPO.

Eksperimen pada lingkungan Atari menunjukkan FRI mencapai *mean reward* 19,76 pada Breakout dan menjadi satu-satunya model yang memperoleh *reward* positif sebesar 2,15 pada Pong. Studi ablasi menunjukkan kombinasi RND dan ICM memberikan peningkatan performa terbesar dibandingkan model *baseline*. Pada skenario pelatihan dengan iterasi terbatas, FRI juga mengungguli model berbasis *Transformer* dan *Vision Transformer* yang mewakili pendekatan modern pemrosesan representasi visual.

Temuan ini menunjukkan bahwa integrasi *reward* intrinsik berbasis *state* dan transisi di atas representasi visual yang dibekukan mampu meningkatkan efektivitas eksplorasi, terutama pada lingkungan yang rentan terhadap optimum lokal.

Kata kunci : Reinforcement Learning, Proximal Policy Optimization, Representasi Visual, Latent Space, Game Atari.

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/

ABSTRACT

Reinforcement Learning with Random Feature Visual Representations and Latent Space in Atari Environments

By

NOVANTIO BANGUN HERMAWAN
25/563864/PPA/07096

In end-to-end learning, a CNN serves a dual role as both visual feature extractor and policy network, causing the visual representation to shift continuously as the policy is updated. These unstable features reduce exploration efficiency and make the agent prone to becoming trapped in local optima.

This study decouples visual representation learning from policy learning through the Frozen Random Intrinsic (FRI) architecture, which combines a frozen CNN, Random Network Distillation (RND), and an Intrinsic Curiosity Module (ICM) within the Proximal Policy Optimization (PPO) framework. The CNN is pre-trained and then frozen so that the visual representation remains fixed, while RND (state novelty) and ICM (transition novelty) run in parallel throughout training as sources of intrinsic reward. Rewards from both branches are normalized using Welford's method and combined with the extrinsic reward as input to PPO.

Experiments in Atari environments show that FRI attains a mean reward of 19.76 on Breakout and is the only model to obtain a positive reward of 2.15 on Pong. An ablation study indicates that the combination of RND and ICM yields the largest performance gain over the baseline models. Under a limited-iteration training scenario, FRI also outperforms Transformer- and Vision Transformer-based models, which represent modern approaches to visual representation processing.

These findings demonstrate that integrating state- and transition-based intrinsic rewards on top of a frozen visual representation improves exploration efficiency, particularly in environments susceptible to local optima.

Keywords: Reinforcement Learning, Proximal Policy Optimization, Visual Representation, Latent Space, Atari Games.