A Hybrid CNN and Reinforcement Learning Approach for Parameter Estimation from Magnetic Tweezer Images

Tautvydas Naudžiūnas, Guoda Perminaitė, Povilas Treigys

Institute of Data Science and Digital Technologies
Vilnius University

Problem

Many image-analysis algorithms depend heavily on correctly tuned parameters, but selecting these parameters becomes extremely challenging when there is no clear statistical relationship between the image appearance and the underlying algorithm settings. This difficulty can be seen in domains such as microscopy, where images are often low-resolution and noisy, causing conventional CNN-based regression models to struggle with reliable prediction performance. As a result, manual parameter selection is impractical, and standard supervised models fail to generalize. This study aims to develop a method capable of recovering algorithm parameters accurately, while avoiding the large data requirements of transformer-based approaches.

Objectives

The main objectives of this study are:

- Develop a robust method to predict algorithm parameters from noisy, low-resolution diffraction images.
- Improve prediction accuracy more than what supervised CNN regression can achieve on small datasets.
- Reduce prediction sensitivity to noise.
- Use reinforcement learning to refine parameter predictions based on an external performance metric rather than training directly on labels.

Proposed Solution

We introduce a hybrid CNN + Reinforcement Learning pipeline that is trained not only from the image features but also from real-world performance:

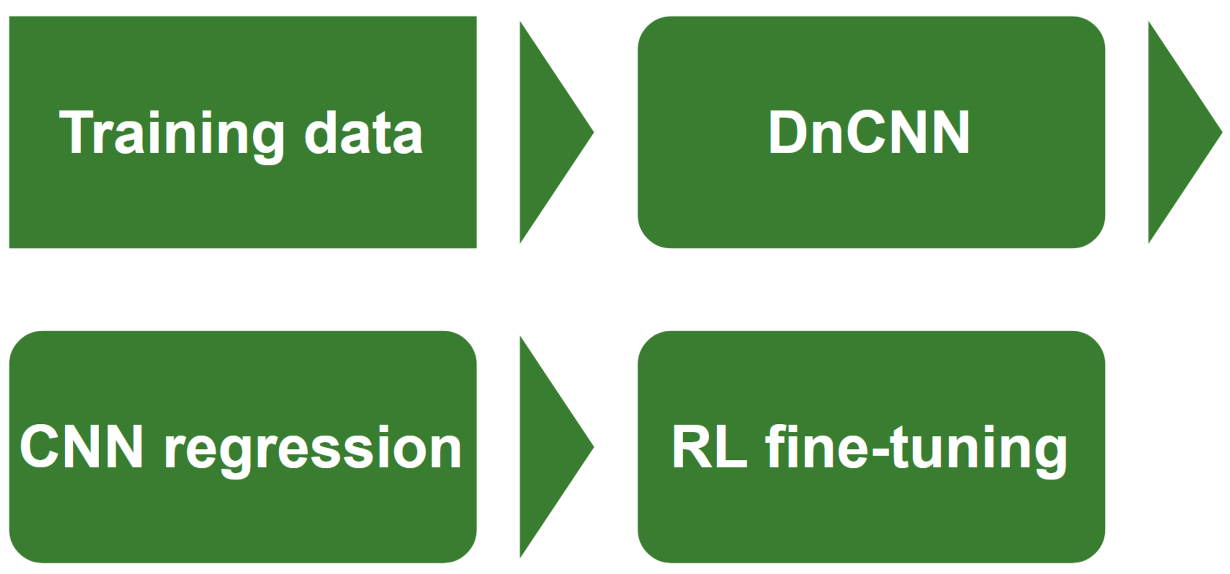


Fig. 1. Proposed training pipeline, RL uses the previous steps to further refine the model.

Method Overview

Stage 1 – Denoising (DnCNN)

- Removes structured and stochastic noise.
- Validation-loss-based tuning.

Stage 2 – Feature Extraction (CNN)

- Provides initial state for RL agent.
- Chosen over transformers (limited training data).
- Alone: insufficient parameter performance.

Stage 3 – Model Fine-Tuning (RL)

- RL agent predicts/refines parameter vector.
- DnCNN is frozen to preserve denoising.
- External algorithm evaluates predicted parameters.
- Reward guides model toward performing solutions.

Fig. 2. Method overview.

Why Our Model Succeeds

Why supervised loss fails:

- Euclidian distance or MSE do not reliably reflect real algorithm performance.
- Small parameter shifts sometimes cause large output changes; large shifts can cause minimal change.
- CNN trained on loss alone may appear optimal but perform poorly in practice.

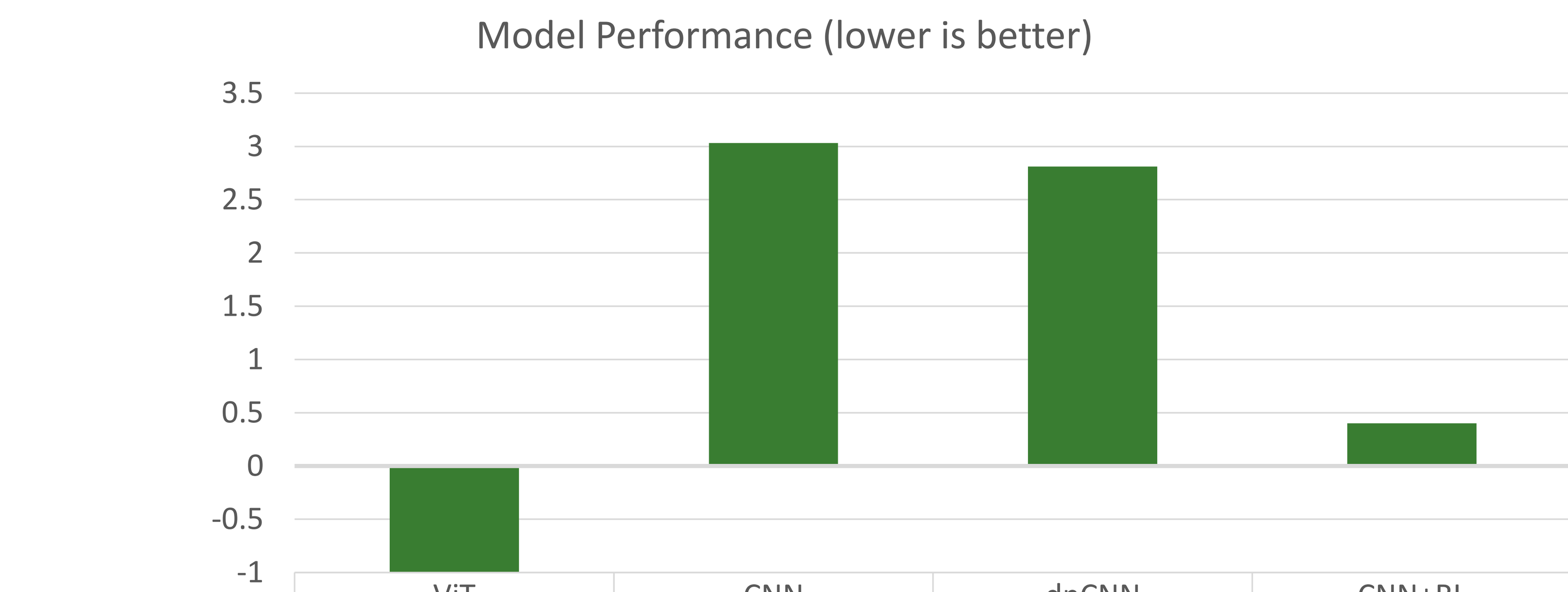
Non-unique solutions:

- Many test cases have multiple parameter sets that can achieve the required performance metric in the algorithm.
- Supervised learning forces the model to match a single label.

Fine-tuning using Reinforcement learning:

- Optimizes directly for performance, not for a loss function.
- Can converge to any high-performing parameter set.
- CNN gives a strong starting point, which greatly reduces RL exploration time.

Results



	ViT	CNN	dnCNN	CNN+RL
Avg. performance	-1	3.03	2.81	0.40
Avg. loss	0.0958	0.0058	0.0051	0.0185

Fig. 3. Average algorithm performance for each model. Our RL hybrid achieves the highest performance across 11 test videos. -1 means the test video has failed.

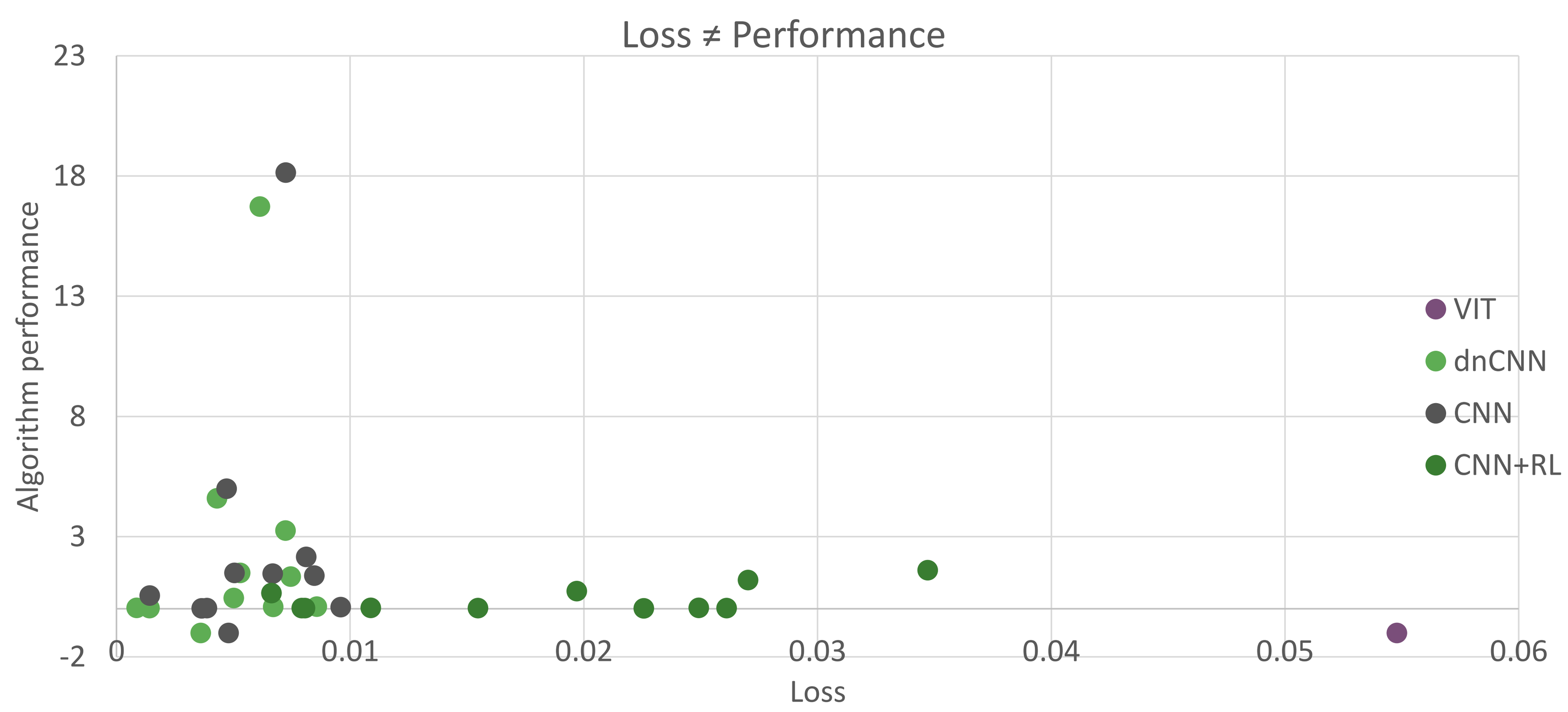


Fig. 4. Loss vs performance scatter plot. Low loss does not imply high performance for supervised models. Our model achieves consistently high performance despite higher loss.

- RL hybrid achieves highest performance across videos.
- Supervised models show low loss but have poor performance.
- Loss ≠ real performance. Our model provides consistent high reward.

Limitations

- Training time: RL requires a lot of interactions with the external algorithm, making training slower than purely supervised approaches.
- Parameter space size: With increasing the dimensionality of the parameter set, exploration becomes more challenging for basic policy gradient methods.
- Reward stability: performance metric fluctuations may affect RL convergence.

Acknowledgements: The authors are thankful for the high-performance computing resources provided by the ITAPC of VU.

Data Analysis Methods for Software Systems, 16th Conference, Druskininkai, Lithuania, November 27–29, 2025

Contact: Tautvydas Naudžiūnas, tautvydas.naudziunas@mif.vu.lt