# Reflective Essay: EAGER Framework Development Journey

**Bayram Tosun**
**241032251**
**Luk Arnaut**
**MSc Computer Science**

## Introduction

Undertaking the development of the EAGER framework for my MSc dissertation presented significant challenges that extended beyond initial expectations. The project began with a seemingly straightforward objective: applying reinforcement learning to improve deepfake detection. However, the implementation process revealed complex technical obstacles, resource limitations, and broader ethical considerations that fundamentally shaped my understanding of AI development.

The deepfake detection problem poses a unique and pressing challenge in today's technological landscape. On one hand, deepfake generation techniques have advanced at an unprecedented pace, buoyed by substantial funding and extensive research support. State-of-the-art synthesis models now produce hyper-realistic and temporally coherent manipulated content that can easily bypass traditional detection systems. On the other hand, many deepfake detection methods are built upon approaches developed several years ago, largely relying on convolutional neural networks and older artifact cues. These methods were conceived under considerably different threat models and with datasets that no longer encapsulate the complexity of current deepfake generation techniques (Rana & Bansal, 2024;, Saif & Tehseen, 2022). This dual limitation—in both methodologies and training data—suggested that the problem required not just technical improvements but a fundamental reconceptualization of the detection approach. This realization led to the core innovation of EAGER: reformulating deepfake detection as a sequential decision-making problem where an intelligent agent learns to investigate strategically rather than exhaustively.

## Critical Analysis of Project Strengths and Limitations

### Strengths Realized Through Implementation

The integration of DINOv3 from Meta (Siméoni et al., 2025) proved transformative for the project. Released during the implementation phase, this self-supervised vision transformer represented a significant advancement over previous feature extraction methods I had explored. DINOv3's ability to extract rich, meaningful features from frames without requiring specific deepfake training data demonstrated the power of self-supervised learning in this domain. The decision to incorporate this model, despite requiring substantial modifications to the existing architecture, ultimately enhanced the system's performance considerably.

The reinforcement learning agent developed capabilities that distinguished it from traditional CNN-based approaches. Instead of a fixed, sequential processing of frames, an RL agent can learn a policy to make intelligent decisions during inference. For instance, in a hybrid supervised and reinforcement learning framework, an RL agent can "meticulously choose and apply top-k augmentations to the test samples via a learned policy" (Nadimpalli and Rattani, 2022). Rather than processing frames sequentially without adaptation, the agent demonstrated strategic thinking during both training and inference. It learned to examine frames selectively, developing patterns that suggested genuine investigation rather than mechanical processing. Achieving 95.83% accuracy while analysing only 20% of available frames validated this strategic approach and demonstrated that intelligent frame selection could maintain detection quality while significantly reducing computational requirements.

The three-phase training methodology emerged through iterative development rather than initial design yet proved remarkably effective. Each phase built naturally upon the previous one, creating a learning progression that allowed for stable improvement. The successful deployment through Django, with processing times of 10-15 seconds per video, demonstrated that the research innovations could translate into practical applications without extensive re-engineering.

### Limitations Discovered Through Experience

The computational requirements for training posed substantial challenges throughout the project. The system demanded high-end GPU resources operating continuously for extended periods—resources that remain inaccessible to many researchers. Multiple training interruptions due to resource constraints or system failures highlighted the reproducibility concerns inherent in computationally intensive research. These limitations raise important questions about the accessibility and democratization of AI research.

Face detection emerged as a critical bottleneck that I had initially underestimated. The challenges extended beyond simple detection failures to complex scenarios involving multiple subjects, face tracking across frames, and handling partial occlusions. When videos contained two people, the detection system would inconsistently switch between faces, disrupting the temporal analysis crucial for deepfake detection. Attempting to implement consistent single-face tracking across frames revealed the complexity of what seemed like a solved problem. These issues affected approximately 8% of processed frames, but their impact on system reliability was disproportionately significant.

The dataset limitation proved particularly constraining. Training on deepfakes generated by 2019-2020 techniques while attempting to detect outputs from contemporary models like Sora (OpenAI, 2024) and Gemini Veo2 (van den Oord, 2024) highlighted a fundamental challenge in defensive AI development. The generational gap between available training data and current threats raised serious questions about the relevance of our detection approaches. This limitation wasn't merely technical but represented a structural problem in how the field approaches the rapidly evolving deepfake landscape.

The implementation of GRPO revealed significant limitations in existing reinforcement learning frameworks. StableBaselines3 (Raffin et al., 2021), despite being a widely used library, proved incompatible with the gradient flow modifications required for group-based advantage computation. The discovery that gradients remained zero despite apparently successful training runs consumed three weeks of development time. This silent failure mode—where training appeared to proceed normally while no learning occurred—necessitated implementing GRPO from scratch using TorchRL. This experience highlighted the gap between theoretical algorithm development and practical implementation in existing frameworks.

## Theoretical Foundations Versus Practical Realities

### The Gap Between Mathematical Elegance and Implementation Complexity

The theoretical elegance of reinforcement learning algorithms contrasts sharply with their implementation complexity. The mathematical formulations—PPO's clipped surrogate objective, advantage estimation, and policy gradients—present clear optimization objectives. However, translating these concepts into functioning systems revealed numerous practical challenges rarely addressed in academic literature. The agent's behaviour during early training phases demonstrated the difficulty of reward engineering. Initial implementations resulted in degenerate strategies: the agent would consistently predict one class to minimize risk or repeatedly examine the same frame to exploit reward loopholes.

Each iteration of reward function refinement revealed new ways the agent could satisfy the mathematical objectives while violating the intended behaviour. The eventual emergence of the "NEXT-NEXT-NEXT" pattern—where the agent learned to examine three consecutive frames before making decisions—represented a successful balance between exploration and efficiency. However, reaching

this point required dozens of iterations, each addressing previously unforeseen exploits in the reward structure.

### Unexpected Discoveries and Adaptations

The implementation of Monte Carlo Dropout for uncertainty estimation presented both theoretical elegance and practical challenges (Gal and Ghahramani, 2015). The Bayesian interpretation—approximating posterior distributions through multiple stochastic forward passes—provided robust confidence estimates crucial for preventing premature classification decisions. However, the computational cost of twenty forward passes nearly doubled inference time, creating tension between uncertainty quality and system efficiency. The improved confidence calibration justified this cost for high-stakes applications, but the trade-off remains a consideration for broader deployment.

The experience with DINOv3 challenged my assumptions about pre-trained models. Previous attempts with ResNet (He et al., 2015), EfficientNet (Tan and Le, 2019), and other vision transformers had yielded moderate results. DINOv3's performance wasn't merely incrementally better—it fundamentally transformed the system's capabilities. The features extracted by this self-supervised model appeared particularly suited for deepfake detection, despite no specific training for this task. This discovery emphasized the importance of architectural selection over training paradigm in transfer learning applications.

## Future Development Pathways and Unrealized Ambitions

### Immediate Extensions with Additional Time

Given additional development time, several critical improvements would take priority. The confidence calculation system, while functional, requires refinement for production deployment. Current uncertainty estimates perform well in typical cases but lack robustness in edge scenarios where appropriate uncertainty expression is crucial.

The dataset limitation demands innovative solutions. Creating synthetic training data using contemporary generation models would help bridge the gap between available data and current threats. While this approach raises its own challenges regarding data authenticity and distribution shift, it represents a practical path toward maintaining detection relevance.

The potential for human feedback reinforcement learning remains particularly compelling. Incorporating human judgment into the reward signal could address the challenge of defining appropriate behaviour programmatically. Combined with continual learning capabilities that allow adaptation to new deepfake techniques without complete retraining, this approach could enable detection systems to evolve alongside generation methods.

Deepfake generation often involves independent manipulation of audio and visual streams, which can lead to subtle inconsistencies and uncertainties in learned representations (Zou et al., 2024). By analysing the relationship between audio and visual signals, models can detect when these signals don't align as they would in authentic media (Koutlis and Papadopoulos, 2024).

### Long-term Research Directions

The selective attention mechanism developed for EAGER has potential applications beyond deepfake detection. Various computer vision tasks could benefit from strategic, RL-guided analysis rather than exhaustive processing. Medical imaging, security screening, and quality control represent domains where efficient, targeted analysis could provide value. However, each application would require careful adaptation and domain-specific considerations rather than direct transfer of the existing framework.

The concept of hierarchical decision-making—where agents operate at multiple temporal and spatial scales—represents a natural evolution of the current approach. Such systems could adapt their investigation strategies based on content complexity, potentially improving both efficiency and

accuracy. While implementing this vision exceeded the dissertation timeline, it represents a promising direction for future research.

## Legal, Social, Ethical, and Sustainability Considerations

### Ethical Implications of Detection Technology

Developing deepfake detection technology necessitates confronting complex ethical considerations. Creating improved training datasets requires generating or collecting more sophisticated deepfakes, which inherently involves using people's facial data Many existing datasets for computer vision, including those used for deepfakes, have faced criticism for collecting data without informed consent, leading to their withdrawal or modification (Andrews et al., 2023) (Hanley et al., 2020). The current approach to dataset creation often overlooks these considerations, focusing on technical performance over ethical implications.

The current Django implementation provides basic functionality, offering links to authorities when deepfakes are detected. However, this represents minimal engagement with the broader privacy and security implications of biometric processing. The system processes facial data without comprehensive privacy safeguards beyond avoiding permanent storage. The decision not to implement user tracking or personalization features reflected uncertainty about ethical implementation rather than a principled privacy stance.

The dual-use potential of detection technology remains a persistent concern. Systems designed to identify deepfakes could enable surveillance applications, support authoritarian control, or be reverse engineered to improve deepfake generation. The decision to open-source the code prioritized academic transparency and reproducibility, but the potential for misuse requires ongoing consideration.

## Personal Growth and Professional Development

### Technical Skills Evolution

This project fundamentally transformed my approach to technical problem-solving. The development process established a rigorous cycle of research, implementation, failure analysis, and iteration. Each challenge—from debugging gradient flow issues to implementing custom GRPO modifications—expanded my technical capabilities beyond initial expectations. The necessity of building solutions from scratch when existing tools proved insufficient developed both my programming skills and my ability to navigate complex codebases.

The project revealed that proposed solutions often lead to unexpected discoveries. The initial goal of applying reinforcement learning to deepfake detection evolved into implementing custom optimization algorithms, integrating state-of-the-art vision models, and developing web deployment systems. Each expansion of scope brought new learning opportunities and challenges that enriched the research experience.

### Research Methodology and Critical Thinking

The iterative development of reward functions provided invaluable lessons in systematic problem-solving. The agent's behaviour served as an unforgiving critic of logical flaws in reward design. When the agent discovered exploits in reward structures, it forced more careful consideration of incentive mechanisms. When learning failed despite seemingly correct implementations, it necessitated questioning fundamental assumptions about the problem formulation.

The project would have benefited from more comprehensive validation methodologies implemented earlier in the development process. While the achieved accuracy metrics are encouraging, more extensive cross-validation across diverse data distributions would strengthen confidence in the system's generalization capabilities. This reflection doesn't diminish the project's achievements but highlights areas for methodological improvement in future research.

## Conclusion: Synthesis and Forward Vision

The EAGER project represents both a successful technical achievement and a learning experience that extends beyond its immediate contributions. Demonstrating that reinforcement learning agents can achieve 95.83% accuracy while processing only 20% of video frames validates the core hypothesis about strategic investigation. However, the implementation challenges encountered throughout development provide equally valuable insights into the practical realities of AI system development.

The gap between theoretical elegance and implementation complexity isn't a failure of theory or practice but rather reflects the inherent challenges of advancing the field. Each obstacle—from framework limitations to dataset constraints—provided learning opportunities that traditional coursework couldn't offer. These experiences emphasize the importance of practical implementation in understanding both the potential and limitations of AI techniques.

Looking forward, the project reinforces my commitment to developing AI systems that balance technical innovation with practical considerations. The computational costs of training, privacy implications of deployment, and ethical considerations of dual-use technology aren't peripheral concerns but central challenges that must be addressed in responsible AI development.

This dissertation journey has evolved my perspective from implementing existing techniques to critically examining fundamental assumptions in AI development. The technical achievements—accuracy improvements, efficiency gains, successful deployment—are important, but the questions raised about sustainable, ethical, and adaptive AI development will guide my future research. The challenge of deepfake detection serves as a microcosm of broader challenges in AI: building systems that remain effective as threats evolve, balancing performance with accessibility, and ensuring that defensive technologies don't enable new forms of harm. These considerations will continue to shape my approach to AI research and development beyond this project.

**References**


Rana, P. and Bansal, S. (2024). Exploring deepfake detection: techniques, datasets and challenges. International Journal of Computing and Digital Systems, 15(1), 769-781. https://doi.org/10.12785/ijcds/160156

Saif, S. and Tehseen, S. (2022). Deepfake videos: synthesis and detection techniques – a survey. Journal of Intelligent &Amp; Fuzzy Systems, 42(4), 2989-3009. https://doi.org/10.3233/jifs-210625

Siméoni, O. et al. (2025) “DINOv3.” doi:10.48550/ARXIV.2508.10104.

Nadimpalli, A.V. and Rattani, A. (2022) “On Improving Cross-dataset Generalization of Deepfake Detectors,” in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), p. 91. doi:10.1109/cvprw56347.2022.00019.

OpenAI. (2024). Sora System Card. Available at: https://www.openai.com/index/sora-system-card/ (Accessed: 20 August 2025).

Van den Oord, A. (2024). Google Labs: Video Image Generation Update, December 2024. Available at: https://blog.google/technology/google-labs/video-image-generation-update-december-2024/ (Accessed: 20 August 2025).

Raffin, A. *et al.* (2021) “Stable-Baselines3: Reliable Reinforcement Learning Implementations,” *Journal of Machine Learning Research*, 22(268), p. 1. Available at: https://jmlr.org/papers/volume22/20-1364/20-1364.pdf (Accessed: August 2025).

Gal, Y. and Ghahramani, Z. (2015b) “Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning,” arXiv (Cornell University) [Preprint]. doi:10.48550/arXiv.1506.02142.

He, K. et al. (2015) “Deep Residual Learning for Image Recognition.” doi:10.48550/ARXIV.1512.03385.

Tan, M. and Le, Q.V. (2019) “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” arXiv [Preprint]. doi:10.48550/ARXIV.1905.11946.

Zou, H. et al. (2024) “Cross-Modality and Within-Modality Regularization for Audio-Visual Deepfake Detection,” in ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), p. 4900. doi:10.1109/icassp48485.2024.10447248.

Koutlis, C. and Papadopoulos, S. (2024) “DiMoDif: Discourse Modality-information Differentiation for Audio-visual Deepfake Detection and Localization.” doi:10.48550/ARXIV.2411.10193.

Andrews, J.T.A. et al. (2023) “Ethical Considerations for Responsible Data Curation.” doi:10.48550/ARXIV.2302.03629.

Hanley, M. et al. (2020) “An Ethical Highlighter for People-Centric Dataset Creation,” arXiv (Cornell University) [Preprint]. doi:10.48550/arxiv.2011.13583.