Evaluating Uncertainty in Deep Q-Network Ensembles for Trustworthy Anomaly Detection in Medical Imaging

Zeduo Zhang[1,2], **Yalda Mohsenzadeh**[1,2]

[1]Department of Computer Science, Western University, London, Ontario, Canada,
[2]Vector Institute for Artificial Intelligence, Toronto, Ontario, Canada
zzhan762@uwo.ca, ymohsenz@uwo.ca

Abstract

Reliable anomaly detection is crucial for safe AI deployment in clinical imaging, yet most systems offer limited insight into prediction uncertainty or failure modes—key factors in medical decision-making. We analyze the uncertainty characteristics of a patch-level Deep Q-Network anomaly detection framework (DQN_AD) for brain MRI, trained with few annotated abnormal cases and designed to generalize to highly imbalanced clinical datasets. Our study links uncertainty to model errors, calibration, anomaly scores, spatial correspondence with ground truth, and selective evaluation. Results show that high-uncertainty predictions consistently coincide with error-prone regions, providing a strong signal for identifying potential failures. This study establishes the foundation for uncertainty-aware, reinforcement learning–based anomaly detection models that enhance reliability, interpretability, and clinical usability in large-scale MRI analysis.

Introduction

Reliable anomaly detection in medical imaging is essential for clinical safety and decision support. Despite advances in unsupervised (Zhang and Mohsenzadeh 2025a; Schlüter et al. 2022) and semi-supervised methods (Zhang et al. 2023; Yao et al. 2023), most systems remain opaque. They offer limited insight into prediction uncertainty or potential failure modes. These are critical limitations in high-stakes settings such as brain MRI, where both false negatives and false positives carry serious consequences.

The previous work, DQN_AD (Zhang and Mohsenzadeh 2025b), adapts Deep Q-Networks for patch-level anomaly localization in medical images. By representing each state as a CNN feature patch, DQN_AD achieves fine-grained, visually interpretable localization while learning from only a few annotated abnormal cases. This design addresses the cost and scarcity of clinical annotations, generalizes well under severe class imbalance, and has demonstrated competitive performance with high computational efficiency—making it well-suited for large-scale MRI analysis.

Given its accuracy, interpretability, and scalability, assessing the reliability of DQN_AD through uncertainty analysis is crucial for clinical adoption. In this study, we extend DQN_AD to an ensemble of independently trained agents, allowing pixel- and image-level uncertainty estimation via variance and entropy between ensemble members. We systematically analyze how these uncertainty measures relate to model errors, calibration, anomaly scores, and spatial correspondence with ground truth. Our findings reveal that high-uncertainty predictions consistently align with error-prone regions, providing a reliable signal for identifying potential failures.

These insights form the foundation for developing uncertainty-aware DQN_AD models, where uncertainty can be incorporated into reward design or environment sampling to guide exploration toward ambiguous regions. Such integration could further enhance robustness, interpretability, and trustworthiness, paving the way for clinically reliable, decision-supportive anomaly detection in brain MRI.

Related Work

Anomaly Detection

Deep learning has enabled effective anomaly detection (AD) by modeling normal data and flagging deviations. Approaches include reconstruction-based (Lu et al. 2023), embedding-based (Li et al. 2023), and self-supervised methods (Bozorgtabar and Mahapatra 2023), but most lack robustness to distribution shifts and do not quantify model uncertainty—limiting their use in safety-critical domains. Recently, reinforcement learning (RL)-based AD, such as DQN_AD (Zhang and Mohsenzadeh 2025b), has been proposed to localize anomalies via sequential decision-making, yet typically ignores uncertainty estimation.

Explainability and uncertainty-awareness are crucial for medical anomaly detection, where model decisions must be interpretable and trustworthy to ensure clinical adoption. These aspects not only support safety and transparency but also help mitigate risks in real-world deployment scenarios (Campagner et al. 2025; Houssein et al. 2025). In anomaly detection, combining the outputs (e.g., heatmaps) from multiple predictors can highlight regions of disagreement between models, which often correspond to ambiguous or abnormal areas (Roy 2025). These regions of high disagreement are particularly valuable for interpretability, as they direct attention to uncertain zones in the input.

This work analyzes the uncertainty characteristics of

DQN_AD as a step toward integrating uncertainty awareness into decision-making, aiming for clinically reliable, decision-supportive DQN-based systems.

Method

DQN-based Anomaly Detection

We address semi-supervised anomaly detection, where training data is dominated by normal images and only a few annotated anomalies are available, which is a common setting in medical imaging.

The previous DQN_AD framework (Zhang and Mohsenzadeh 2025b) formulates anomaly localization as sequential patch-level decisions, using multi-scale CNN features as states and a task-specific reward to encourage accurate detection while limiting false positives. Each locally aware patch feature serves as the state for an RL agent, which selects an action indicating normal or anomalous. Sparse rewards are given for correct predictions on the few labeled anomalies, and the agent outputs Q-values for each patch. Aggregating patch scores yields the final anomaly map. A guided sampling strategy balances labeled anomaly and unlabeled normal patches: after correct predictions, the agent explores dissimilar regions; after errors, it focuses on similar regions, mitigating class imbalance.

Ensemble of Q-Network

While the DQN-based anomaly detection framework performs well in semi-supervised settings, it lacks mechanisms for uncertainty quantification, which is vital for interpretability and safety in deployment. We address this by training an ensemble of independently initialized DQN agents in parallel, each interacting with the environment using its own replay buffer to avoid cross-agent data sharing, thus preserving learning diversity. This mechanism can be directly leveraged in future work to develop uncertainty-aware DQN-based anomaly detection methods—where uncertainty signals inform reward shaping, guide exploration toward challenging regions, and ultimately yield models that are more accurate, robust, and clinically reliable. During inference, all agents evaluate the image, and their predictions are aggregated to compute mean, variance, and entropy per patch. These measures provide robust anomaly localization and interpretable uncertainty maps, enhancing reliability for clinical use.

We extract two complementary measures at the pixel level (Zahari, Cox, and Obara 2024). Standard deviation (std) quantifies the dispersion of anomaly scores across agents, with high std values indicating strong inter-agent disagreement. Such disagreement often arises in structurally ambiguous regions, under domain shifts, or in patterns that are underrepresented in training data, making std maps valuable both for uncertainty visualization and as modifiers for prediction confidence. Entropy, computed from the ensemble-averaged softmax probabilities, captures prediction ambiguity at each pixel. High entropy values correspond to flatter probability distributions and are frequently observed near lesion boundaries or in gradual transitions between normal and abnormal tissue. Each agent outputs an anomaly probability $p^{(k)}(x)$ per pixel. We take a binary vote $\hat{y}^{(k)}(x)$ with a fixed threshold 0.5. Let $m(x) = \frac{1}{K}\sum_{k=1}^{K} \hat{y}^{(k)}(x)$ be the fraction of "anomaly" votes. The *vote entropy* is

$$H_{\text{vote}}(x) = -\,m(x)\log m(x) - (1-m(x))\log(1-m(x)), \tag{1}$$

At the image level, we take the std across agents of their per-image maximum anomaly score to measure global disagreement, and the spatial mean of the pixel-wise (vote) entropy to summarize overall predictive ambiguity.

models	AUROC	AUPRC	Sensitivity	DICE
DQN_AD	**88.7/96.7**	**94.8/66.2**	86.6/**63.0**	84.2/**49.7**
DQN-E	85.8/96.4	93.1/63.9	**89.4**/61.2	**88.0**/47.6

Table 1: DQN_AD vs. DQN-E. Metrics are reported as image / pixel.

Results

Datasets & Experiment Setup

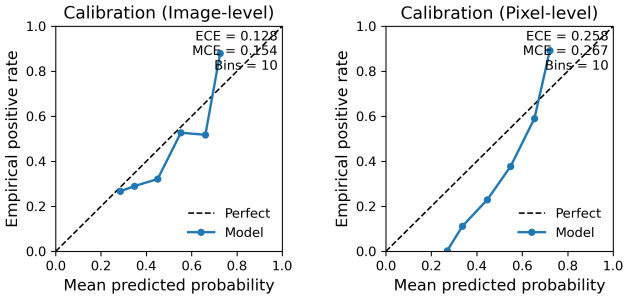
We use the IXI (healthy) and BraTS 2021 (glioblastoma) MRI datasets, following the training/testing split and preprocessing pipeline of DQN_AD (Zhang and Mohsenzadeh 2025b). IXI volumes are used for training and BraTS for validation/testing in a semi-supervised setting with 10 labeled anomaly cases and abundant normal samples. Ten bootstrapped DQN agents are trained under the same architecture and hyperparameters as DQN_AD. Unlike the original paper which selects the operating threshold on the validation set, we use a fixed probability threshold of 0.5 for all threshold-based metrics and agent votes (no test-time tuning).

Baseline Performance and Rationale for Uncertainty Analysis

Table 1 compares the baseline DQN_AD with its ensemble variant (DQN-E) at image and pixel levels using AUROC, AUPRC, sensitivity, and DICE. DQN_AD attains higher AUROC/AUPRC, indicating stronger discriminative ability, while DQN-E improves pixel-level sensitivity (86.6→89.4) and DICE (84.2→88.0), suggesting enhanced segmentation stability via ensemble averaging. Given these results, DQN_AD is used as the primary backbone for uncertainty analysis, with ensemble outputs exploited to interpret prediction variability.

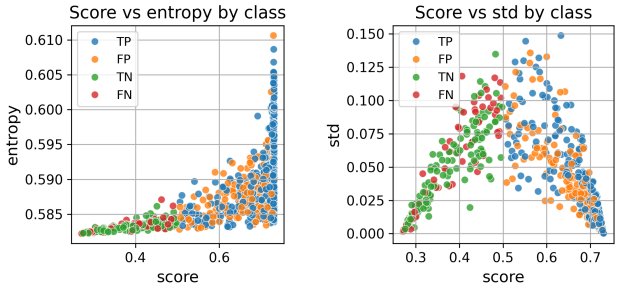
Reliability of Predictions

Figure 1 reports reliability diagrams (Kocak et al. 2025) at the image and pixel levels (10 bins). The model shows moderate image-level calibration (Expected Calibration Error (ECE) = 0.128, Maximum Calibration Error (MCE) = 0.154). Deviations concentrate in the mid-confidence range (0.4–0.6), indicating mild under-confidence, whereas predictions remain close to the diagonal for high-confidence scores (> 0.7). In contrast, pixel-level calibration is weaker (ECE = 0.258, MCE = 0.267), with systematic gaps across



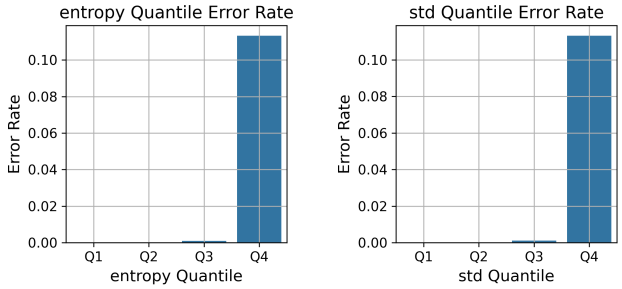
(a) (b)

Figure 1: Calibration curves for the DQN_AD model at (a) image level and (b) pixel level.



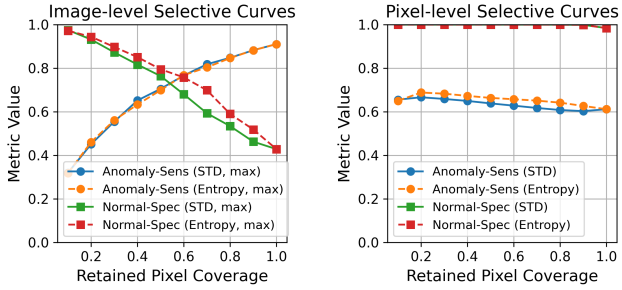
(a) (b)

Figure 2: Anomaly score versus uncertainty at the image level.



(a) (b)

Figure 3: Pixel error rate across uncertainty quantiles.



(a) (b)

Figure 4: Selective evaluation: retain low-uncertainty pixels first; image-level metrics computed after pixel filtering and max aggregation.

bins, likely reflecting heavy class imbalance and spatial correlation. Implications. For clinical use, image-level triage would benefit from post-hoc calibration (e.g., temperature scaling or isotonic regression) to improve decisions in borderline cases; this is left as future work. Pixel-level heatmaps are still informative for delineating lesion extent, but their probabilities should be treated cautiously. These findings further indicate the need for developing uncertainty-aware strategies to enhance the reliability of both image- and pixel-level anomaly detection.

Structural Relationship Between Score and Uncertainty

For the image-level analysis in Fig. 2, uncertainty (std) peaks at intermediate anomaly scores, indicating higher ambiguity in borderline cases. TP tends to show slightly higher uncertainty than FP in image-level std, while FN and TN exhibit similar uncertainty patterns. In image-level entropy, TP and FP are generally higher than TN and FN. These findings suggest that uncertainty captures meaningful distinctions between prediction outcomes.

Error-Uncertainty Coupling and Selective Evaluation

The error–uncertainty analysis in Fig. 3 shows that at the pixel level, both measures display a highly concentrated error distribution, where the majority of misclassified pixels fall into the highest-uncertainty quartile (Q4) and errors are negligible in Q1–Q3. This sharp separation highlights the strong discriminative capability of pixel-level uncertainty estimates and suggests that excluding or down-weighting high-uncertainty pixels could substantially improve segmentation performance, enabling selective evaluation or human-in-the-loop refinement strategies.

In Fig. 4, at the pixel level, specificity remains nearly flat across coverage because the majority of low-uncertainty background pixels are consistently classified as normal; anomaly sensitivity decreases slightly when boundary pixels (with higher uncertainty) are progressively included. At the image level, adding high-uncertainty pixels before aggregation (max) increases sensitivity, since many such pixels in positive images coincide with lesion regions, while decreasing specificity in negatives due to amplified noise/artifacts. This divergence highlights that pixel-level uncertainty chiefly stabilizes background, whereas image-level aggregation magnifies extreme, ambiguous evidence.

Spatial Correspondence Between Uncertainty and Anomalous Regions

Table 2 compares ensemble-derived mean, std, and entropy scores with ground-truth anomalies at both pixel and image levels. The mean score, defined as the ensemble-averaged anomaly probability, serves as a baseline. High correlation between uncertainty measures (std or entropy) and ground

	mean	std (max)	std (mean)	entropy
AUROC	85.8/96.4	18.7/90.6	76.9/90.6	73.6/90.3
AUPRC	93.1/63.9	50.5/23.1	84.4/23.1	80.9/21.0

Table 2: AUROC and AUPRC of ensemble-derived signals (image / pixel) with respect to ground truth. "max/mean" denote image-level aggregation.

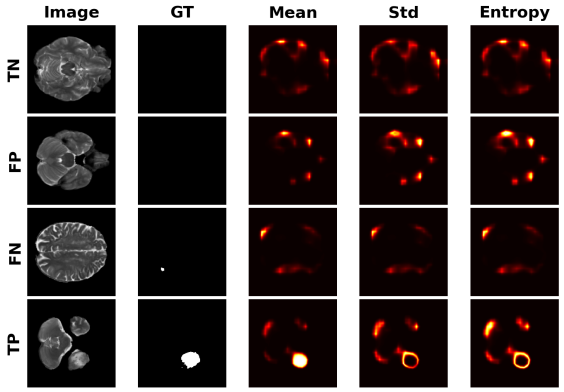


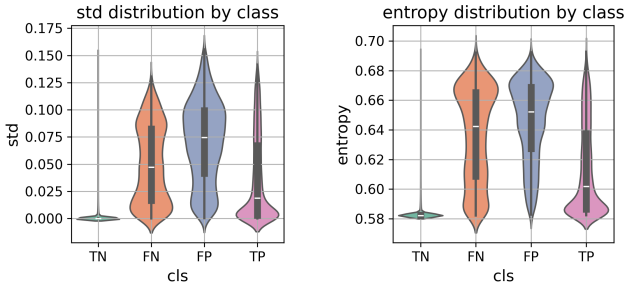
Figure 5: **Uncertainty maps.** Examples across TN/FP/FN/TP with mean score, std, and entropy.

truth suggests that uncertainty can be a reliable indicator of anomalous regions. For std, aggregation strategy matters: std(mean) achieved higher image-level AUROC (76.9) than std(max) (18.7), because max pooling emphasizes the most confident pixels, which are often lesion cores where model disagreement is minimal. Entropy, computed from pixel-wise predictive distributions and averaged over the image, also correlated strongly with anomaly locations, supporting its role as an explanatory signal. For consistency with the base DQN_AD model and sharper localization, subsequent analyses adopt max-based aggregation, with mean-based scores retained as a complementary reference.

Qualitative heatmaps (Fig. 5) show that both std and entropy concentrate in regions corresponding to FP and FN cases, especially along lesion boundaries. Lesion cores exhibit high confidence with low std, explaining the limited AUROC of max-based std. This balance characterized by high confidence in characteristic pathology regions while retaining uncertainty signals for ambiguous boundaries, aligns with real-world MRI workflows, where annotations are costly and diagnostic reliability is critical.

Uncertainty Distribution Across Prediction Outcomes

In Fig. 6, FP and FN pixels exhibited substantially higher uncertainty than TP and TN pixels, reflecting greater disagreement among ensemble agents in incorrectly predicted regions. FP pixels showed higher uncertainty than FN pixels, suggesting that false positives often stem from ambiguous or artifact-like patterns interpreted inconsistently by agents. TN and TP pixels had low uncertainty, indicating strong inter-agent consensus; however, TP uncertainty was slightly higher than TN, possibly due to greater variability



(a) (b)

Figure 6: uncertainty distribution by prediction class at pixel level.

in correctly identified anomalous pixels compared to uniform background regions. From a clinical perspective, these findings indicate that uncertainty cues, particularly entropy for FP and std for FN, could be operationalized to flag cases for targeted review, reducing both missed detections and unnecessary follow-ups.

Conclusion

This study demonstrates that DQN_AD is not only highly discriminative for semi-supervised medical anomaly detection, but also reliably aware of where errors are likely to occur. Pixel-level uncertainty enables precise lesion refinement, while image-level uncertainty supports triage and targeted review, which is critical for safe deployment in high-stakes clinical settings. Entropy and std provide complementary cues for identifying false positives and false negatives, respectively, offering actionable strategies such as:

- excluding or down-weighting high-uncertainty pixels to improve segmentation stability,
- applying uncertainty gating or calibration for borderline image-level predictions, and
- using uncertainty maps to guide targeted human review.

Looking forward, we will investigate robustness to clinically relevant noise and artifacts, assess the impact of ensemble size and replay-buffer designs, and compare alternative uncertainty estimation techniques (e.g., Monte Carlo Dropout, distributional RL). We also plan to develop an uncertainty-aware DQN_AD that incorporates uncertainty signals directly into the reward function or environment sampling, enabling adaptive decision-making in ambiguous cases. Combining this with robustness-enhancement strategies such as adversarial training and artifact-aware sampling could yield models that are both uncertainty-calibrated and resilient, advancing the trustworthiness of DQN-based anomaly detection in real-world MRI workflows.

References

Bozorgtabar, B.; and Mahapatra, D. 2023. Attention-conditioned augmentations for self-supervised anomaly detection and localization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 14720–14728.

Campagner, A.; Biganzoli, E. M.; Balsano, C.; Cereda, C.; and Cabitza, F. 2025. Modeling Unknowns: A Vision for Uncertainty-Aware Machine Learning in Healthcare.

Houssein, E. H.; Gamal, A. M.; Younis, E. M.; and Mohamed, E. 2025. Explainable artificial intelligence for medical imaging systems using deep learning: a comprehensive review. *Cluster Computing*, 28(7): 469.

Kocak, B.; Klontzas, M. E.; Stanzione, A.; Meddeb, A.; Demircioğlu, A.; Bluethgen, C.; Bressem, K.; Ugga, L.; Mercaldo, N.; Díaz, O.; et al. 2025. Evaluation metrics in medical imaging AI: fundamentals, pitfalls, misapplications, and recommendations. *European Journal of Radiology Artificial Intelligence*, 100030.

Li, J.; Wang, X.; Zhao, H.; Wang, S.; and Zhong, Y. 2023. Anomaly segmentation for high-resolution remote sensing images based on pixel descriptors. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 4426–4434.

Lu, R.; Wu, Y.; Tian, L.; Wang, D.; Chen, B.; Liu, X.; and Hu, R. 2023. Hierarchical vector quantized transformer for multi-class unsupervised anomaly detection. *Advances in Neural Information Processing Systems*, 36: 8487–8500.

Roy, D. 2025. Bayesian Autoencoder for Medical Anomaly Detection: Uncertainty-Aware Approach for Brain 2 MRI Analysis. *arXiv preprint arXiv:2504.15562*.

Schlüter, H. M.; Tan, J.; Hou, B.; and Kainz, B. 2022. Natural synthetic anomalies for self-supervised anomaly detection and localization. In *European Conference on Computer Vision*, 474–489. Springer.

Yao, X.; Li, R.; Zhang, J.; Sun, J.; and Zhang, C. 2023. Explicit boundary guided semi-push-pull contrastive learning for supervised anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 24490–24499.

Zahari, R.; Cox, J.; and Obara, B. 2024. Uncertainty-aware image classification on 3D CT lung. *Computers in Biology and Medicine*, 172: 108324.

Zhang, H.; Wu, Z.; Wang, Z.; Chen, Z.; and Jiang, Y.-G. 2023. Prototypical residual networks for anomaly detection and localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 16281–16291.

Zhang, Z.; and Mohsenzadeh, Y. 2025a. Efficient slice anomaly detection network for 3D brain MRI Volume. *PLOS Digital Health*, 4(6): e0000874.

Zhang, Z.; and Mohsenzadeh, Y. 2025b. Semi-Supervised Anomaly Detection in Brain MRI Using a Domain-Agnostic Deep Reinforcement Learning Approach. *arXiv preprint arXiv:2508.01137*.