Seeing Clearly: Retrospective AI-Based Detection of Lens Contamination in Minimally Invasive Surgery

Faiza Quadri[1]; Rishit Yokananth[1]; Eshan Sayani[2]; Sriya Dommaraju[1]; Kathleen Schmidt[3]; Maansi Srinivasan[4]; Lisa Du[1]; Tony Zhu[1]; John Uecker, MD[1,3]; Christopher Idelson, PhD[3]

[1]The University of Texas at Austin, TX; [2]Texas A&M School of Medicine, TX; [3]ClearCam Inc.; [4]Texas A&M University School of Engineering Medicine, TX

Introduction

- Minimally invasive surgery (MIS) relies on clear laparoscopic visualization, yet lens contamination events (LCEs) such as fogging, blood, and debris frequently obstruct the surgical field[1-3]
- Visualization quality in MIS is typically assessed subjectively by clinicians, leading to variability between operating rooms and inconsistent recognition of compromised visualization[4-5]
- This study evaluates the performance of an artificial intelligence model, LUCID, designed to automatically classify laparoscopic lens clarity as clean or contaminated

Methods

Study Design & Dataset	• **Study type:** Retrospective analysis of laparoscopic surgical video • **Data source**: De-identified laparoscopic videos collected from hospitals in Central Texas • **Dataset:** 77 surgical cases totaling ~96 hours of video (44.5 hours gynecology, 51.6 hours general surgery). • Procedures included hysterectomy, cholecystectomy, appendectomy, oophorectomy, hernia repair, splenectomy, and colon resection • Video frames standardized to 224x224 resolution • **Ethics:** De-identified surgical videos were analyzed and did not require IRB review
Video Annotation	• Trained annotators labeled frames as “clean” or “dirty” based on presence of lens contamination events (blood, condensation, smearing, debris) • Frames with unobstructed visualization labeled clean • Frames with visible contamination labeled dirty • Two independent annotators reviewed frames to reduce subjectivity and improve labeling reliability
Model Development	• **Deep learning model:** Proprietary transformer-based image classification model • Frames resized to 224x224 pixels and normalized to [0,1] pixel range • Data augmentation during training included horizontal flipping and ± 15° rotation
Model Validation	• Holdout test dataset created from original dataset • **Test set:** 9 laparoscopic cases totaling 261 minutes • **Average case length:** 29.0 ± 12.3 minutes • Model predictions compared with human-annotated ground truth

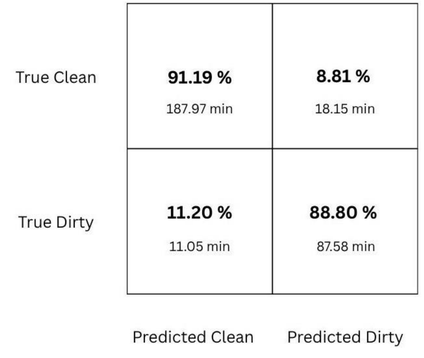
Performance Metrics	• Model performance evaluated using:
	  • **Accuracy:** overall classification correctness
	  • **Precision:** reliability of contamination predictions
	  • **Sensitivity:** ability to detect true contamination events
	  • **Specificity:** ability to correctly identify clean frames

Results

- **AI Model Performance:** Across the holdout dataset (9 laparoscopic cases, 261 minutes of video), the LUCID model achieved **90.42% accuracy**, with 94.45% precision, 91.19% sensitivity, and 88.80% specificity in detecting lens contamination **(Figure 1)**
- **Classification Outcomes:** The model correctly classified **61.68%** of frames as **clean** and **28.74%** as **contaminated**, while 5.96% were false negatives and 3.63% were false positives **(Figure 1)**
- **Case-Level Performance**: Average case accuracy was **90.77% ± 5.25%,** demonstrating consistent performance across procedures of varying duration and surgical type **(Figure 2)**
- **Visualization Variability Across Cases:** Suboptimal visualization varied substantially across surgeries, with predicted contaminated frames ranging from 1.33% to 98.95% per case. On average, **44.35% ± 43.15%** of frames were predicted as **contaminated**, compared to a ground truth prevalence of 32.37% **(Figure 2)**

Results

Figure 1: Confusion matrix comparing predicted and actual classification for dirty and clean classification after deployment of the AI model



Among laparoscopic surgeries analyzed, lens contamination and suboptimal visualization were common. The LUCID AI model demonstrated high accuracy in identifying compromised visualization, suggesting feasibility for objective assessment of intraoperative visual clarity.

Figure 2: Case data with video length, lens clarity with model classification and ground truth labels, and model accuracy per case

Case ID	Case Length (min)	True Clean (%)	Predicted Clean (%)	True Dirty (%)	Predicted Dirty (%)	Case Accuracy (%)
HD	25.32	0.00	1.05	100	98.95	98.95
GH	21.37	100.00	98.67	0	1.33	98.67
GA	29.07	100.00	92.83	0	7.17	92.83
GE	89.15	100.00	91.18	0	8.82	91.18
JD	22.23	0.00	11.39	100	88.61	88.61
SB	31.9	100.00	88.35	0	11.65	88.35
KH	34.63	100.00	87.87	0	12.13	87.87
DB	5.8	0.00	12.93	100	87.07	87.07
QA	45.28	0.00	16.56	100	83.44	83.44

Discussion

Objective Visualization Assessment	• Evaluation of visualization quality in MIS currently relies largely on subjective clinician judgment, leading to variability across operating rooms • LUCID introduces a quantitative and automated approach to detect lens contamination events during surgery • Objective visualization assessment may help standardize recognition of suboptimal visualization across surgical teams
Model Performance Interpretation	• LUCID demonstrated strong classification performance, achieving 90.42% overall accuracy on the holdout dataset • Performance metrics were balanced across measures, including 94.45% precision, 91.19% sensitivity, and 80.80% specificity • Case-level accuracy (90.77% ± 5.25%) closely aligned with overall model accuracy, suggesting consistent performance across procedures
Clinical Implications	• Lens contamination events disrupt workflow and contribute to operative inefficiency and safety risks during MIS • Prior literature estimates that ~40% of MIS operating time occurs under suboptimal visualization conditions[2] • Automated detection of compromised visualization may help support intraoperative awareness and improve surgical workflow efficiency
Role in Future Surgical AI Systems	• AI models for visualization-based tasks rely on clear image data to maintain optimal accuracy and performance[6] • Systems capable of detecting visualization degradation in real time may help ensure reliable input for downstream AI-assisted surgical technologies • Objective monitoring of visualization quality may therefore serve as an enabling layer for future intelligent surgical systems

Strengths and Limitations

Strengths

- Utilizes intraoperative laparoscopic videos rather than simulated datasets
- Incorporates frame-level annotation with independent human reviewers to establish ground truth for model validation
- Evaluates model performance using holdout set to enable assessment of generalization on unseen surgical cases

Limitations

- Small holdout test dataset limits ability to fully evaluate performance across diverse surgical contexts
- Video data derived from limited geographical region, restricting generalizability
- Manual annotation of frames introduces variability in labeling despite oversight

Conclusion

- LUCID achieved strong performance in detecting laparoscopic lens contamination, supporting feasibility of automated visual assessment during MIS
- Automated detection of lens contamination may help reduce workflow disruption, cognitive burden, and risks associated with compromised visualization
- Continued development of visualization monitoring systems may enable standardized assessment of surgical video quality and support the next generation of AI-assisted surgery

Disclosure

Dr. Uecker and Dr. Idelson are co-founders and equity holders of ClearCam Inc. Ms. Schmidt is an employee of ClearCam Inc., and Ms. Srinivasan was previously employed by the company. All other authors report no competing financial interests.

References

1. Ameerah et al. Surg Endosc. 2025. 2. Dhingra et al. J Soc Laparosc Robot Surg. 2025. 3. Bonrath et al. BMJ Qual Saf. 2015. 4. Dae et al. J Robot Surg. 2022. 5. Venkatayogi et al. J Robot Surg. 2023. 6. Arakaki et al. JMA J. 2025.