

MIND THE (DOMAIN) GAP: EVALUATING AND IMPROVING ROBUSTNESS OF ARTIFICIAL INTELLIGENCE SYSTEMS IN ENDOSCOPY

Fons van der Sommen

UEG Week Copenhagen 2023

Poster

October 15, 2023

Artificial intelligence systems for endoscopic applications show substantial performance degradation when exposed to clinically prevalent image artifacts, but specific training strategies can improve robustness.

KEY POINTS

- Three AI systems trained for Barrett's neoplasia detection and polyp characterization experienced mean performance losses of 7.4% with mild image degradation and 16.6% with moderate degradation when tested on images with clinically prevalent artifacts including blur, inadequate exposure, and color variability.
- Domain-specific pretraining and ensemble model approaches substantially reduced performance loss across all degraded test sets compared to standard training protocols.
- AI systems trained on high-quality data sets are prone to domain-shift, performing poorly when real-world endoscopy conditions differ from training data.
- Image degradation severity was clinically calibrated by expert endoscopists to represent quality levels encountered in daily clinical practice.
- This work is relevant to developers and clinical implementers of AI systems in gastrointestinal endoscopy who need to ensure robust performance across diverse clinical conditions.

VIEW THIS CONTENT ON GUTFLIX

<https://gutflīx.eu/search/d/01308ad0-743c-11ee-8122-0242ac140004>



AI-generated summary

This summary was generated by an AI large language model based on the content transcript. It is for informational purposes only and should not be considered a substitute for clinical judgment. Always rely on your professional expertise and the full clinical context when making clinical decisions.