SIIM CMIMI24
Oct 21-22 | Boston, MA
Conference On Machine Intelligence In Medical Imaging

Analysis of Out-of-Distribution Factors to Detect Iris and Pupil Using Cataract Surgical Images

Mahtab Faraji, PhD Student, University of Illinois Chicago; Rogerio Garcia Nespolo, PhD;
Homa Rashidisabet, PhD Candidate; Daniel Wang, MD; Alexis Warren, MD; Hesham Gabr, MD;
Yannek Leiderman, MD, PhD; Darvin Yi, PhD

Introduction/Background

Deep Learning (DL) has substantial potential in ophthalmology, particularly for disease classification and tasks like iris and pupil detection in cataract surgery. DL algorithms typically assume that train and test samples share the same distribution. However, in practice, test samples often differ in distribution, which is called Out-of-Distribution (OOD), affecting generalizability. The OOD can result from factors like differences in data acquisition and preprocessing methods between train and test datasets. This study investigates the impact of two possible OOD-causing factors on a YOLOv5 model's performance in detecting the iris and pupil in cataract surgery images.

Methods/Intervention

Figure 1 presents a block diagram of the proposed method. Cataract surgical images were divided into train, validation, and test sets. The test set underwent two transformations to simulate OOD: 1) adding Gaussian noise at various levels and 2) converting images to grayscale. We trained a YOLOv5 model for iris and pupil detection using the train set and evaluated it with the validation set. The model was then tested on both the original and transformed test data. Performance was assessed using the Mean Average Precision (mAP) metric.

Results/Outcome

Tables 1 and 2 illustrate the YOLOv5 model's performance in detecting the iris and pupil in noisy and grayscale images, respectively, across four different datasets. The first columns of both tables present results for the original test data, with values representing average mAP for iris and pupil detections. Table 1 shows a progressive decline in mAP as noise levels increase from 2% to 15%. Notably, dataset_3 remains stable up to 4% noise, while dataset_4's performance drops to nearly zero at 15% noise. Similarly, Table 2 reveals a 3%-6% reduction in mAP across all datasets when images are transformed to grayscale.

Conclusion

Our findings demonstrate that noise and grayscale conversion significantly impact DL model performance. This underscores the necessity of considering these factors when deploying DL models in real-world scenarios.

Statement of Impact

This study examines how OOD factors, such as noise and grayscale conversion, affect DL models for iris and pupil detection in cataract surgery images.

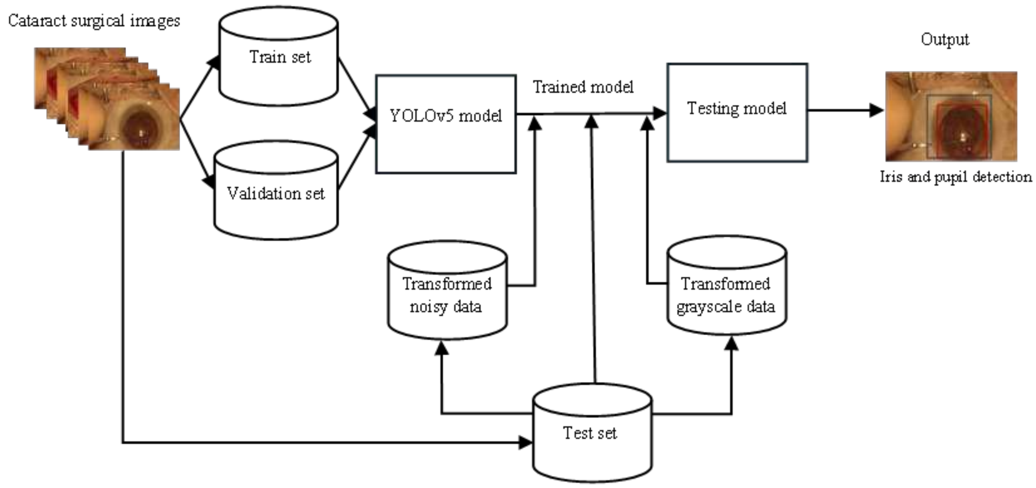

Fig 1. A block diagram of the proposed method.

Dataset	Original images	Transformed data					
		Gaussian noise value					
		2%	4%	6%	8%	10%	15%
Dataset_1	0.80	0.77 (3.75%)	0.76 (5.0%)	0.76 (5.0%)	0.76 (5.0%)	0.74 (7.5%)	0.65 (18.75%)
Dataset_2	0.76	0.75 (1.31%)	0.72 (5.26%)	0.70 (7.89%)	0.63 (17.10%)	0.52 (31.57%)	0.31 (59.21%)
Dataset_3	0.77	0.77 (0%)	0.77 (0%)	0.75 (2.59%)	0.75 (2.59%)	0.75 (2.59%)	0.65 (15.58%)
Dataset_4	0.75	0.75 (0%)	0.71 (5.33%)	0.65 (13.33%)	0.57 (24.0%)	0.38 (49.33%)	0.0 (100%)

Table 1. Average mAP of iris and pupil for the original and noisy test dataset. The values in parentheses indicate the percentage drop in performance compared to the original images.

Dataset	Original images	Transformed data by grayscaling
Dataset_1	0.80	0.77 (3.75%)
Dataset_2	0.76	0.73 (5.26%)
Dataset_3	0.77	0.74 (3.89%)
Dataset_4	0.75	0.71 (5.33%)

Table 2. Average mAP of iris and pupil for the original and grayscaled test dataset. The values in parentheses indicate the percentage drop in performance compared to the original images.

Keywords

Deep learning; Out-of-Distribution; Cataract surgery images; YOLOv5