Deep Learning Detection of Common Bile Duct Stones on Segmented Intraoperative Cholangiograms

Matthew B. Bloom MD[1], Sayeh Mirshojae BS[1], Yimeng He MS[2], Xiuzhen Huang PhD[2], Edward Phillips, MD[1]
Departments of Surgery[1] and Computational Biomedicine[2], Cedars-Sinai Medical Center, Los Angeles, CA

Introduction

Intraoperative cholangiography (IOC) is performed during cholecystectomy to define biliary anatomy and detect common bile duct (CBD) stones. However, with increasing reliance on perioperative ERCP for the management of choledocholithiasis, many surgeons now perform IOC infrequently. As a result, interpretation often occurs under time pressure without immediate expert consultation, particularly during emergency or nighttime operations.

Artificial intelligence-based image analysis may provide real time decision support for surgeons performing IOC. This study evaluates the feasibility of deep learning based automated detection of CBD stones on intraoperative cholangiography.

Methods

- A retrospective dataset of 1,156 intraoperative cholangiogram images, each consisting of a single representative frame from a unique patient undergoing cholecystectomy at a tertiary care center, was assembled (954 without stones, 202 with stones).
- Images were annotated by trained reviewers to identify biliary anatomy and contrast filling defects consistent with common bile duct stones.
- Data were divided at the patient level to prevent data leakage, ensuring that images from each patient appeared in only one dataset. The data were used to train and evaluate a convolutional neural network for automated stone detection.
- Dataset split
  Training set: **816 images**
  Validation set: **204 images**
  Independent test set: **125 images**
- Model performance was evaluated on the independent test set using receiver operating characteristic (ROC) analysis and standard classification metrics.

Results

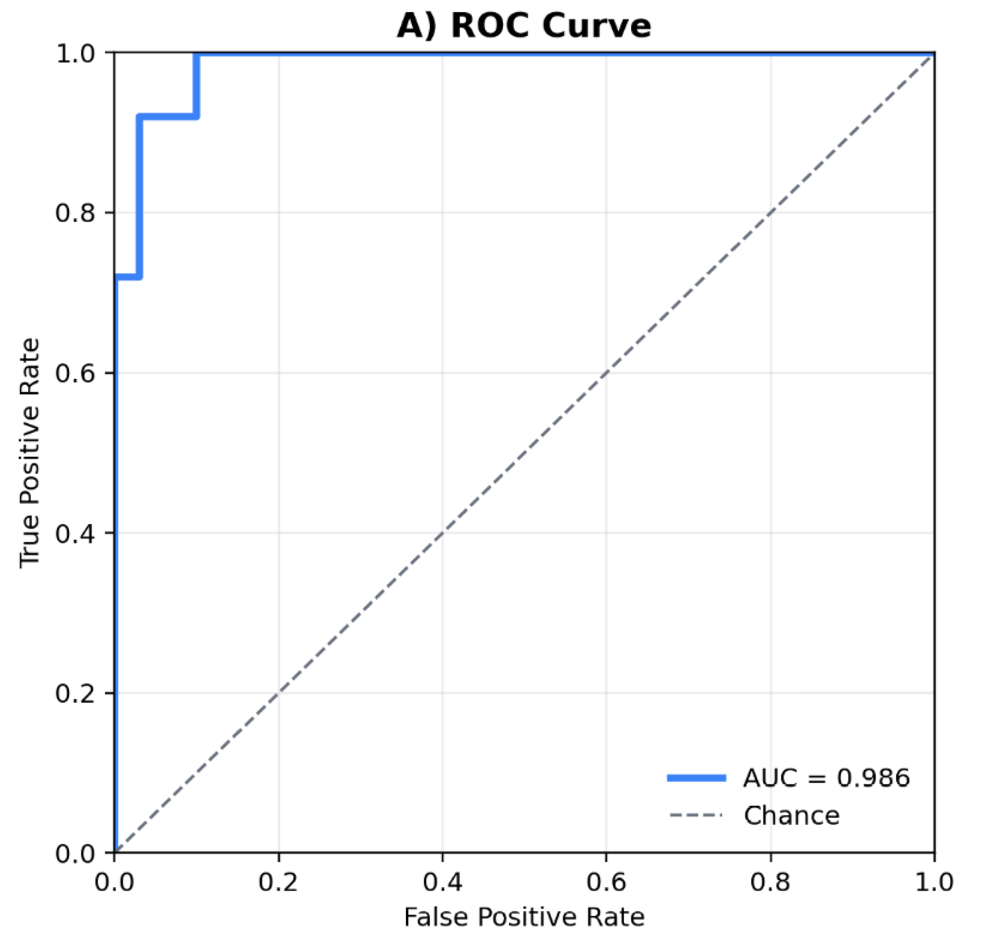

Figure A: ROC curve showing CNN performance for CBD stone detection IOC (AUC = 0.986).

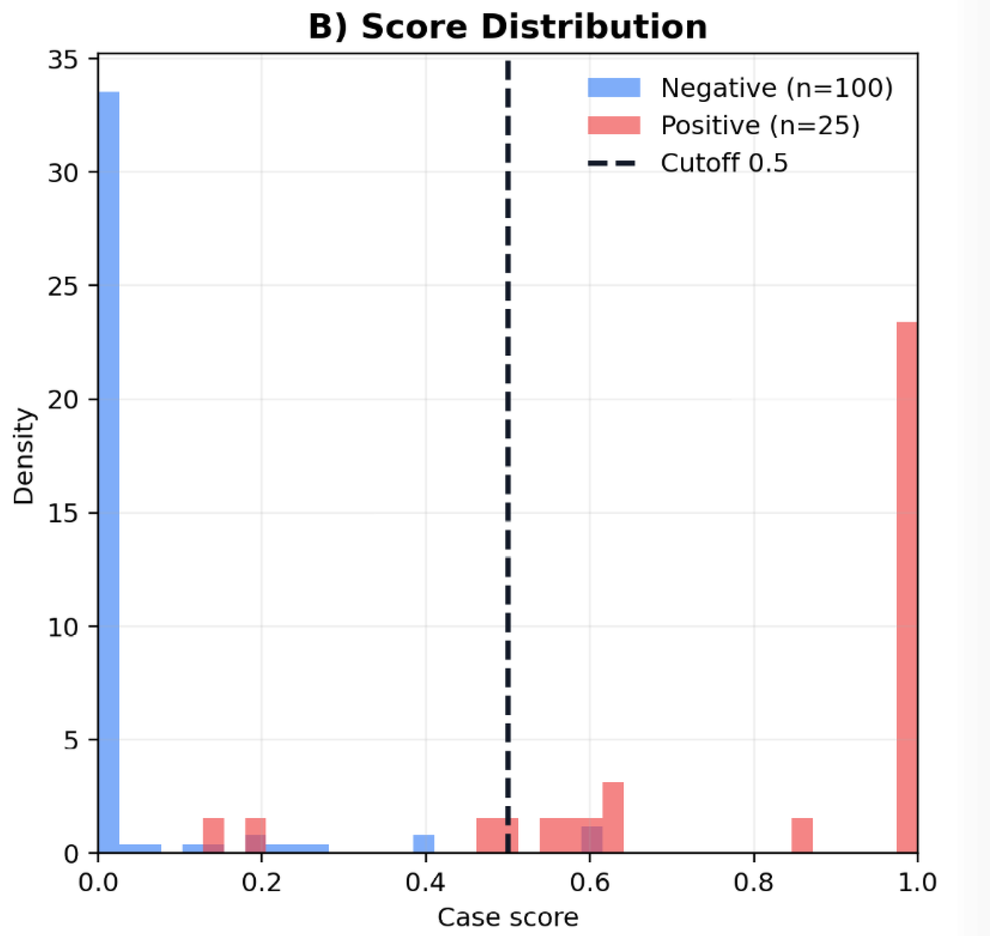

Figure B: Distribution of CNN prediction scores IOCs classified as negative (n = 100) and positive (n = 25) for CBD stones. The dashed vertical line indicates the classification cutoff (0.5).

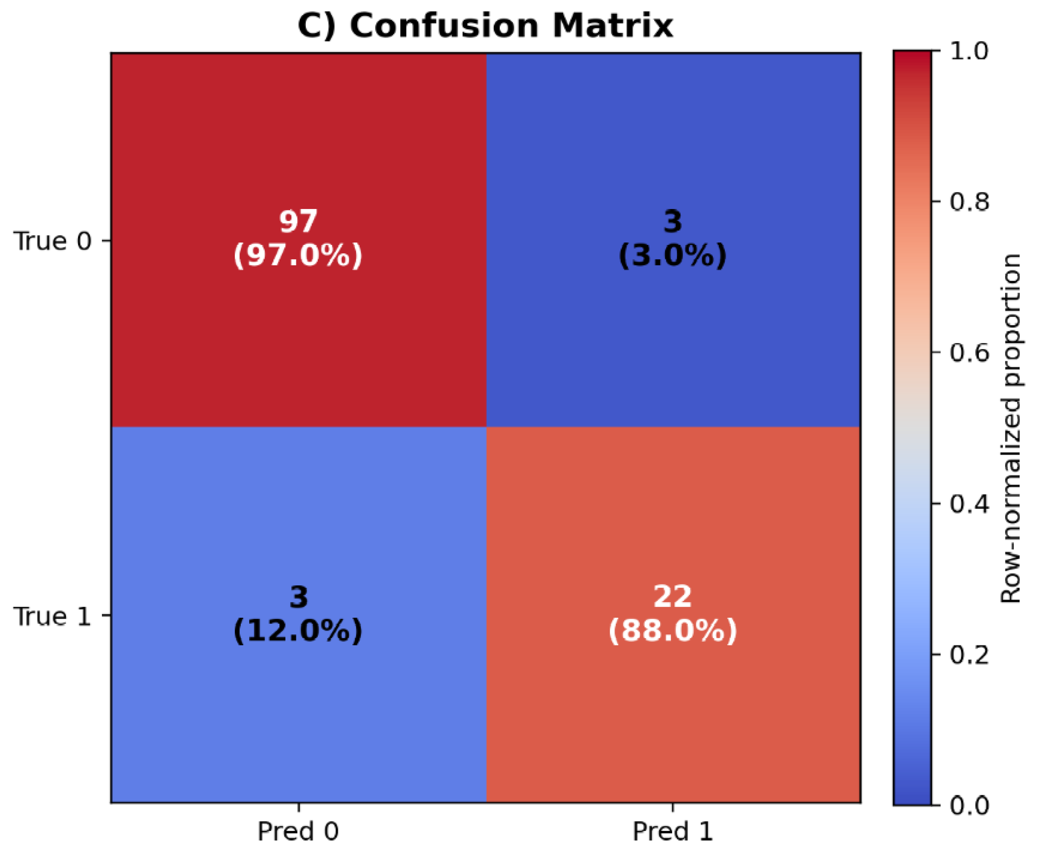

Figure C: Confusion matrix showing classification outcomes of the CNN model for detection of CBD stones on IOCs. Cells display the number and percentage of cases for each predicted and true class.

Key Findings:

On the independent test set (n = 125), the model demonstrated strong performance for detection of CBD stones.

ROC analysis showed **AUC = 0.986**. Using a probability threshold of 0.5, classification performance was:

- **Sensitivity: 88%**
- **Specificity: 97%**
- **Overall accuracy: 95%**

Prediction score distributions showed clear separation between stone-positive and stone-negative cholangiograms.

Results (cont.)

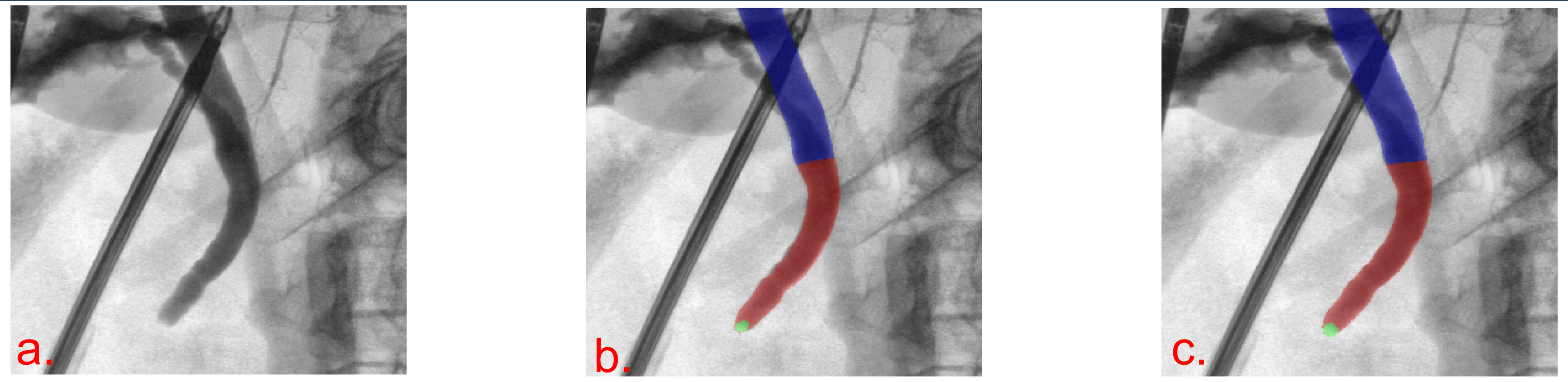

Figure 1.0: Representative cases showing common bile duct stones on IOC. (a) Source (b) Ground truth (c) Prediction.

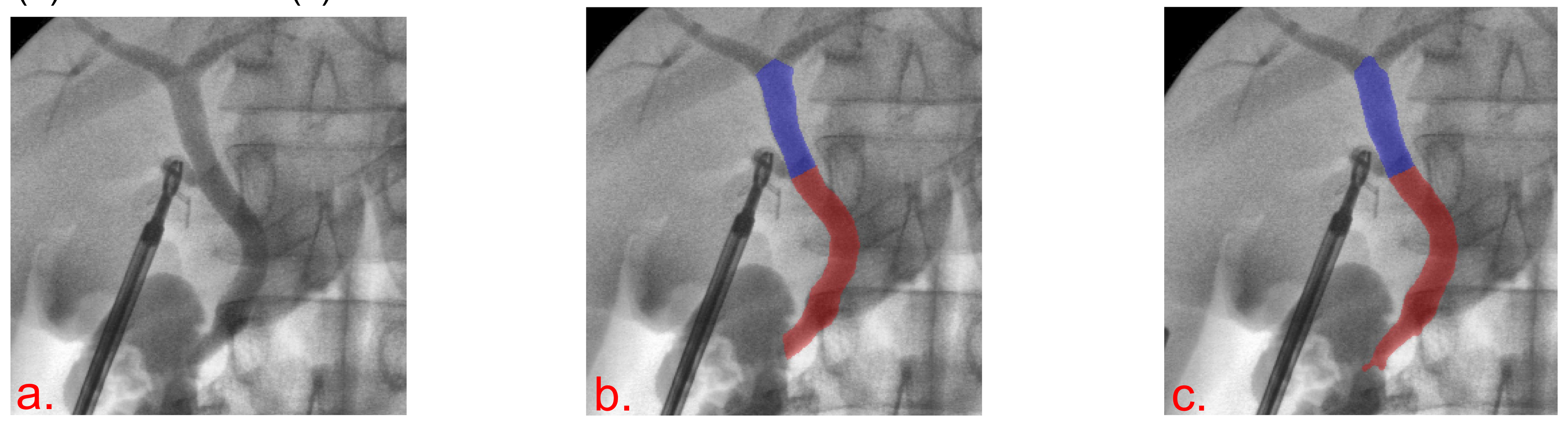

Figure 2.0: Representative cases showing no stone IOC. (a) Source (b) Ground truth (c) Prediction.

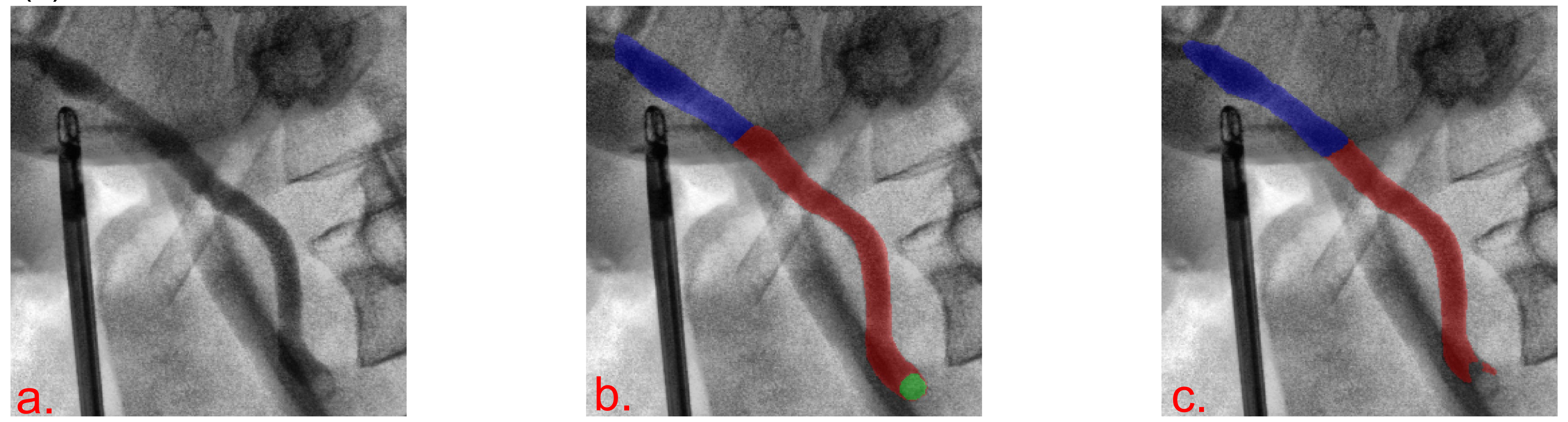

Figure 3.0: Representative cases showing false negative IOC. (a) Source (b) Ground truth (c) Prediction.

Conclusion

Deep learning–based analysis of intraoperative cholangiograms demonstrated accurate detection of common bile duct stones. These findings suggest that AI-assisted image interpretation may support intraoperative decision-making during cholecystectomy.