An Intelligent Expert System for Difficulty Scoring in Common Bile Duct Stone Extraction During Endoscopic Retrograde Cholangiopancreatography (ERCP)

Mona Mohammad-Asghari (MSc)[1], Taha Pishro Dabaghiyan (MSc)[1], Tahereh Mahmoudi (PhD)[2,3], Ramin Niknam (MD)[4,5*], Hossein Parsaei (PhD)[2], Mohammad Mehdi Movahedi (PhD)[2], Kamran Bagheri Lankarani (MD)[4], Fardad Ejtehadi (MD)[5], Seyed Ebrahim Fallahzadeh Abarghooei (MD)[5], Seyed Alireza Taghavi (MD)[5], Gholam Reza Sivandzadeh (MD)[5], Saghar Alihosseini (MD)[5], Seyed Ali Malek-Hosseini (MD)[6]

[1]Research Committee, Shiraz University of Medical Sciences, Shiraz, Iran
[2]Department of Medical Physics and Engineering, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran
[3]Nanomedicine and Nanobiology Research Center, Shiraz University of Medical Sciences, Shiraz, Iran
[4]Health Policy Research Center, Shiraz University of Medical Sciences, Shiraz, Iran
[5]Gastroenterohepatology Research Center, Shiraz University of Medical Sciences, Shiraz, Iran
[6]Shiraz Transplant Center, Abu Ali Sina Hospital, Shiraz University of Medical Sciences, Shiraz, Iran

ABSTRACT

Background: Extraction of Common Bile Duct (CBD) stones via Endoscopic Retrograde Cholangiopancreatography (ERCP) can be technically challenging, with difficulty influenced by anatomical and procedural factors. Accurate automated assessment of these features may facilitate objective difficulty scoring and procedural planning. This study presents an intelligent Expert System for Difficulty Scoring (ESDS) that integrates anatomical features automatically extracted from fluoroscopic images with demographic and procedural parameters to predict the technical difficulty scoring.

Objective: To develop and validate an automated expert system for objective prediction of the technical difficulty of CBD stone extraction during ERCP using imaging and clinical features.

Material and Methods: This retrospective study used 52 ERCP cases to develop the ESDS. Image preprocessing was applied to enhance contrast and preserve structural details. The CBD, stones, and duodenoscope were delineated, and imaging-based features were extracted using customized image processing algorithms. These included stone diameter, proximal and distal CBD diameters, distal CBD angulation, proximal and distal CBD arm lengths. Demographic and procedural variables were also incorporated. An if then rule-based framework integrated these features to generate difficulty scores.

Results: The automated feature extraction method was validated against expert annotations using Bland Altman analysis, demonstrating excellent agreement between automated and manual measurements. The average Intra-Class Correlation Coefficient (ICC) across the evaluated anatomical features was approximately 99.62%, with a Root Mean Square Error (RMSE) of 0.36.

Conclusion: By enabling objective difficulty scoring, it has the potential to improve ERCP quality, reduce procedure time, and lower risks associated with stone extraction.

Keywords

Endoscopic Retrograde Cholangiopancreatography (ERCP); Common Bile Duct (CBD); Duodenoscopes; Expert Systems; Difficulty Scoring; Artificial Intelligence

*Corresponding author:
Ramin Niknam
Health Policy Research Center, Shiraz University of Medical Sciences, Shiraz, Iran
E-mail: niknamramin@yahoo.com

Received: 13 September 2025
Accepted: 24 August 2026

Copyright: © Journal of Biomedical Physics and Engineering
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 Unported License, (http://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited non-commercially.

Introduction

Common Bile Duct Stones (CBDS), defined as the presence of stones within the Common Bile Duct (CBD), are a frequent cause of biliary obstruction, affecting approximately 1–15% of patients with gallstones [1]. This condition not only complicates the management of cholelithiasis but also contributes to serious clinical consequences, including recurrent cholangitis, pancreatitis, and obstructive jaundice. In addition, it places a considerable financial burden on healthcare systems due to repeated interventions and prolonged hospitalizations [2].

Endoscopic Retrograde Cholangiopancreatography (ERCP) is the preferred minimally invasive approach for CBDS removal, offering reduced risk and faster recovery compared to surgery [2,3]. The effectiveness of ERCP depends on achieving complete stone removal and ensuring full bile duct clearance. Failure to extract all stones can result in their recurrence and related complications, with reported rates of CBDS recurrence after ERCP with Endoscopic Sphincterotomy (EST) ranging from 4% to 25% [4]. Therefore, removing stones is crucial for improving patient outcomes and recovery. The evaluation of ERCP outcomes depends on both the procedure's success and complication rates, which can vary widely due to multiple factors. Among these, technical difficulty plays a key role in determining ERCP success [5]. A major factor contributing to this difficulty is the specific anatomical and morphological features of the stones and CBD. Studies indicate that stones are larger than 1 cm in diameter, a distal CBD diameter below 15 mm, and a distal CBD angle under 135° increase extraction challenges due to the narrow bile duct opening after sphincterotomy [5-7]. Hence, the endoscopist's ability to accurately assess the number and anatomical details of stones, including size, is vital for successful outcomes. Nonetheless, assessing these features from fluoroscopic or endoscopic images is challenging not only for less experienced practitioners but also for expert endoscopists. Advanced medical image processing approaches, particularly feature extraction techniques, provide promising strategies to overcome these challenges. Endoscopists benefit from prior knowledge of stone characteristics, such as size and location. There is a recognized need for an automated system to analyze features of common bile duct stones using fluoroscopic images during surgery.

Huang et al. [8] reported in a retrospective study of fluoroscopic images of patients undergoing ERCP that several factors influence the technical difficulty of endoscopic extraction of CBD stones, including the size and number of stones, the diameter, arm, and distal angle of the common bile duct. Williams et al. [9] worked on the management of CBDS in an updated guideline in 2017. For challenging CBDS, endoscopic sphincterotomy with Endoscopic Large Papillary Balloon Dilation (EPLBD) may be the preferred initial procedure (Wongkanong et al. [10], 2022).

However, the relationship between CBD stone-related parameters and the need for adjunctive techniques such as EPLBD remains difficult to determine objectively. Demographic and clinical characteristics included age, sex, Body Mass Index (BMI), underlying disease, acute cholangitis, and previous cholecystectomy [10]. In a study [11] in 2023, reviewing data from patients who underwent ERCP, they concluded that stones larger than 2 cm, square-shaped stones, cases with ampullary dilation, and cases requiring mechanical lithotripsy were more difficult than others. In a study [12] in 2015, other clinical factors including; age, sex, number of CBD stones, stone diameter, external pancreaticobiliary tract, internal pancreaticobiliary tract, transverse diameter of stone, and stone type (bilirubin/cholesterol) were introduced as indicators affecting the difficulty of stone extraction. Although previous studies have identified multiple anatomical and morphological factors that contribute to the technical difficulty of CBDS

extraction, these parameters are still evaluated manually by endoscopists. Such manual assessments are inherently subjective, time-consuming, and often inconsistent, particularly when performed under the time constraints of ERCP procedures. Moreover, quantitative measurements from fluoroscopic images can be challenging, even for expert endoscopists, due to variability in image quality, projection angle, and lack of standardized reference scales. As a result, there is currently no widely available automated system to provide objective and quantitative difficulty assessment during ERCP.

To address these limitations, we propose an automated assistant system as Expert System for Difficulty Scoring (ESDS) using both anatomical and clinical features. At first, several imaging-based features such as stone size, diameters of the distal, proximal, and mid-portions of the CBD, as well as the distal and proximal CBD arms and the distal CBD angulation were extracted using image processing techniques. Then these anatomical features were integrated with critical clinical features to develop expert difficulty scoring system. The remainder of this paper is organized as follows: the Materials and Methods section details our data and modeling approach, the Results section presents our findings and performance evaluation, and the final "Discussion and Conclusion" section explores the clinical implications and summarizes our key contributions.

Material and Methods

Data Description

This retrospective study was conducted at Namazi Hospital of Shiraz University of Medical Sciences in Shiraz, Iran.

Consecutive patients were enrolled from July 2023 to January 2025. A total of 422 cholangiograms from ERCP procedures were initially identified from the available database. After applying the predefined inclusion and exclusion criteria and assessing the availability of complete clinical, procedural, and fluoroscopic information, 52 cases were ultimately included in the study and used to develop the proposed difficulty scoring system. The equipment used in this study comprised the following: duodenoscope (TJF-Q180V; Olympus Medical Systems, Japan) and flat-panel detector (C-arm 8000; Ziehm Imaging, Germany).

In this study, specific inclusion and exclusion criteria were established to ensure the selection of appropriate cases for analysis. Patients eligible for inclusion were those who had undergone endoscopic ERCP with either a confirmed or suspected diagnosis of CBDS. Additionally, only patients whose CBD stones were clearly visible on cholangiograms were considered for the study. For exclusion criteria, patients presenting with biliary sludge were excluded, as were pregnant or lactating women to avoid any potential risks to maternal or fetal health. Individuals with history of biliary injury, previous biliary surgery, suspected or confirmed malignancy including pancreatic head carcinoma, cholangiocarcinoma, and papillary tumor were also not included in the study, given the potential for confounding factors that could affect image interpretation. Furthermore, patients who underwent ERCP exclusively for endoscopic stent placement, replacement, or removal were excluded, as these procedures did not contribute to the primary objective of stone detection. Lastly, cases where cholangiopancreatography was unsuccessful whether due to failed cannulation, intubation were omitted to ensure the quality and consistency of the collected data. All collected fluoroscopic images have a resolution of 576×576 pixels, with a physical size of 20.32×20.32 cm and an 8-bit grayscale format (LUT: 0-255).

In this research, data were gathered by four experts, each possessing more than 10 years of experience and having performed over 1,000 ERCP procedures. In addition, an endoscopist with approximately 3 years of experience

and 500 ERCP procedures also participated in datacollection. The study received approval from the Ethics Committee of Shiraz University of Medical Sciences Institutional Review Board. It adhered to the principles outlined in the Declaration of Helsinki. To protect patient privacy, all identifying information was removed from the data prior to analysis, ensuring confidentiality throughout the study.

Feature Extraction from Fluoroscopic Images

Based on previous studies and the consensus of our research team, a set of anatomical and morphological features of the CBD and stones were extracted from fluoroscopic ERCP images to develop the difficulty scoring system. These features included the diameter of the duodenoscope, stone size, diameters of the distal, proximal, and mid-portions of the CBD, the distal and proximal CBD arms, and the distal CBD angulation.

To ensure the reliability of these measurements, each parameter was first obtained manually and subsequently calculated automatically using image processing methods. This dual approach allowed validation of the automated extraction process by comparing the results with manual measurements, thereby confirming the accuracy of the image processing algorithms. Manual measurements were performed using Digimizer software, while automated calculations were carried out with customized image processing algorithms developed for this study.

The specific methods applied for feature extraction are described in detail below.

The **Diameter of the Distal CBD** was defined as the diameter of the 1-cm proximal segment from the opening of the major ampulla. To calculate this parameter, grayscale transverse images of the bile duct were generated, and the midpoints of the duct were identified. A one-pixel-width curve was fitted through these points to represent the duct centerline, and pixel distances were converted into millimeters using the known duodenoscope diameter as a reference. A point located 10 mm proximal to the distal end of the CBD was selected, together with five points above and five points below it, to construct a vertical line across 11 points. The mean overlap between these vertical lines and the duct contour was taken as the distal CBD diameter.

The **Diameter of the Proximal CBD** was defined as the diameter of the 1-cm segment distal to the beginning of the intrahepatic ducts. The calculation method was similar to that of the distal CBD. A point located 10 mm downward from the proximal end of the CBD, along with five points above and five points below it, was used to construct vertical lines. The mean overlap between these lines and the duct contour represented the proximal CBD diameter.

The **Diameter of the Mid-CBD** was defined as the diameter measured at the midpoint of the CBD centerline. At this halfway point, a vertical line was drawn, and its overlap with the duct contour was measured and taken as the mid-CBD diameter.

The **Distal CBD Arm** was defined as the distance between the ampullary orifice and the angular point of the distal CBD. This parameter was calculated by fitting a straight line from a pixel point located 10 mm upward from the distal end of the CBD to the midpoint of the CBD centerline. The length of this line was taken as the distal CBD arm.

The **Proximal CBD Arm** was defined as the distance between the orifice of the intrahepatic ducts and the angular point of the distal CBD. For calculation, a straight line was fitted from a pixel point located 10 mm downward from the proximal end of the CBD to the midpoint of the CBD centerline. The length of this line was considered the proximal CBD arm.

The **Distal CBD Angulation** was defined as the first angulation along the course of the CBD starting from the ampullary orifice. It was calculated as the angle formed between the distal and proximal CBD arms determined

in the previous steps.

Finally, the **Diameter of the Duodenoscope** was determined as a reference scale. It was first identified from the known scope size on the cholangiograms. Alternatively, a quantitative method was used by selecting the midpoint of the duodenoscope centerline together with five points above and five points below it. A vertical line was drawn across these 11 points, and the average overlap with the duodenoscope contour was calculated. This value was taken as the duodenoscope diameter in millimeters. Figure 1 illustrates the exact locations of the eight extracted features on a sample image from the database, as annotated by an expert.

Accurate conversion of measurements from pixels to millimeters is essential, as the extracted anatomical features must ultimately be expressed in real-world units. For this purpose, the duodenoscope, with a known actual diameter of 11.3 mm, was used as an internal reference within cholangiogram images.

A key challenge is that in X-ray imaging the apparent size of an object depends on its distance from the detector, with closer objects appearing larger than their true dimensions. This effect may distort the measured size of bile duct structures or stones, particularly in patients with higher body mass. In Figure 2, two metal coins of a certain diameter are placed on the patient's bed, and an image is taken of them. Then, the coins are placed 5 cm above the bed surface using a small

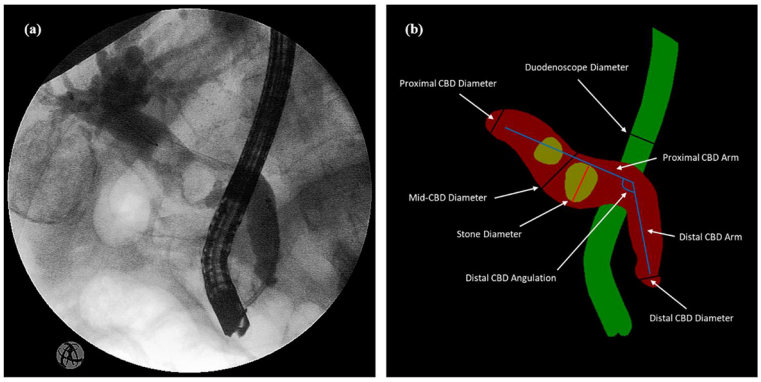


Figure 1: (**a**) Original Image, (**b**) Illustration of the eight extracted features labeled on a sample image by an expert endoscopist.

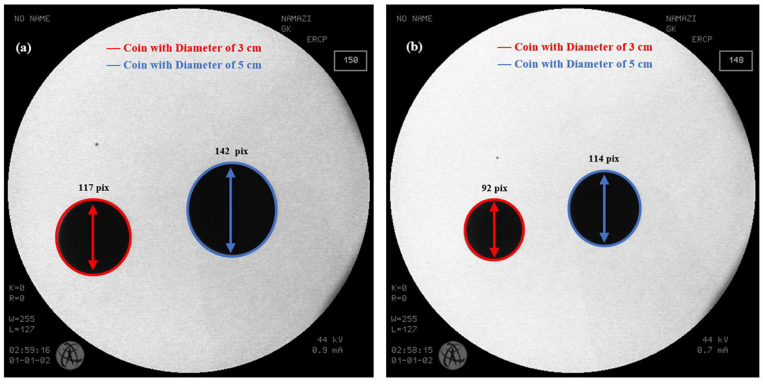


Figure 2: Comparison of the effect of the height of the object on its size in the image. Image (**a**) was recorded at a height of 5 cm from the patient's bed and image (**b**) was recorded at the same height as the patient's bed.

wooden platform, and the image is taken again. The noticeable difference in size of the coins with a fixed size in the two images confirms the need to calculate the image-specific scale for each patient individually.

To reduce this error, the pixel-to-millimeter scale was calculated separately for each image, based on the measured diameter of the duodenoscope in that image. Using this image-specific scale, pixel-based measurements of stone size, CBD diameters (distal, proximal, and middle), and lengths of the distal and proximal CBD arms were converted into millimeters.

Clinical and Procedural Feature Acquisition

Demographic and procedural characteristics that may influence the difficulty of stone extraction were recorded. Information recorded by the endoscopist in real time during the procedure included: Needle-knife precut, presence of diverticulum, spastic duodenum, papilla appearance, CBD sphincterotomy, pancreatic sphincterotomy, sphincter balloon dilatation, CBD cannulation, CBD dilation, CBD guidewire attempt, intrahepatic bile duct dilation, pancreatic contrast injection, pancreatic guidewire attempt, stone extraction attempt, complete clearance, stone/sludge color, and stent insertion. Post-procedure data included patient history (such as jaundice or cholangitis), stone extraction time, total procedure time, and medications used for managing complications. A summary of all clinical-demographic features is shown in Table 1.

Designing an Expert System for Difficulty Scoring

Accurate assessment of the difficulty of stone extraction during ERCP is essential, as it guides endoscopists in selecting the most appropriate therapeutic strategies [5,7,13-15]. In this study, we developed a novel expert-based scoring system designed to predict the technical difficulty of CBD stone extraction, particularly in cases with normal CBD anatomy. The system was constructed through a structured consensus process. Following an extensive literature review, five senior endoscopists (each with experience in over 1000 ERCP procedures) participated in several dedicated meetings. Using a modified Delphi approach, the experts defined and validated scoring criteria for a wide range of anatomical and procedural variables. Summary of these features are presented in Table 2.

Each parameter was assigned a score based on expert consensus which is shown in Table 2. The Expert System was then designed to apply If Then rules to these scores (Table 3). Specifically, the system aggregates the values to generate a cumulative difficulty index: “If the total score ranges from 0 to 4, the case is classified as easy”, “If the total score ranges from 5 to 9, the case is classified as intermediate”, and “If the total score is ≥10, the case is classified as difficult”. In addition, the system incorporates rules for exceptional scenarios, ensuring flexibility in clinical application, also presented in Table 3. These scenarios included: (1) If the stone size is extremely large, then the case is classified as difficult, regardless of the cumulative score. (2) If the distal CBD is severely narrowed or (3) has a sharp angulation, then the case is automatically classified as difficult.

This rule-based framework allows the expert system to mimic the decision-making process of experienced endoscopists, while maintaining transparency and interpretability. Unlike black-box machine learning models, the If Then logic ensures that every decision can be traced back to explicit rules derived from clinical expertise. Thus, the system provides a reliable, flexible, and clinically interpretable tool for predicting the technical difficulty of ERCP stone extraction.

Statistical Analysis

All statistical analyses were conducted using

Table 1: Summary of all clinical-demographic information of patient

Feature Type	Feature Name	Criteria
Historical	Session Number	1st, 2nd, 3rd, 4th,…
	Jaundice	Yes, No
	Acute Cholangitis	Yes, No
	Acute Pancreatitis	Yes, No
	Previous Pancreatitis	Yes, No
	Previous Cholecystectomy	Yes, No
During the Session	Pre-cut	Yes, No
	Papilla Appearance	Normal, Prominent, Ulcerative
	Diverticulum Presence	Beside, Within, No
	Stricture of the Junction of the CBD and Papilla	Yes, No
	CBD Sphincterotomy	Complete, Partial, No
	Pancreatic Sphincterotomy	Yes, No
	Sphincter Balloon Dilation Usage	Yes, No
	CBD Cannulation Successful	Yes, No
	CBD Guidewire Attempts Number	1,2,3,4,…
	Pancreatic Contrast Injection	Yes, No
	Pancreatic Guidewire Attempts Number	1,2,3,4,…
	CBD Dilation	Yes, No
	Intra-hepatic Bile Duct Dilation	Yes, No
	Stone Extraction Attempts Number	1,2,3,4,…
	Complete Stone Clearance	Yes, No
	Stone/Sludge Color	Yellowish, Pigmented
	CBD Stenting	Yes, No
	PD Stenting	Yes, No

CBD: Common Bile Duct, PD: Pancreatic Duct

Python (version 3.10.1) and IBM SPSS Statistics (version 26). Descriptive statistics were used to summarize demographic, anatomical, and procedural variables, with continuous variables expressed as mean ± standard deviation (SD) and categorical variables presented as frequencies and percentages. Agreement between automated and manual feature extraction methods was assessed using Bland–Altman analysis and the intra-class correlation coefficient (ICC). Root mean square error (RMSE) and coefficient of variation (CV) were also calculated to quantify measurement accuracy and variability. For comparisons among the easy, intermediate, and difficult ERCP groups, one-way analysis of variance (ANOVA) was used for continuous variables, whereas the chi-square test was applied to categorical variables. Pearson correlation analysis was performed to assess the relationships between extracted imaging features and procedural variables, including stone extraction time and the number of extraction attempts. A two-sided P-value<0.05 was considered statistically significant.

Results

All image-processing algorithms were

Table 2: Summary of all extracted features of patient image and during procedure information

Feature Name	Scoring Criteria		
	0	1	2
CBD Guidewire Attempts Number	X < 3 times	X = 3 times	X > 3 times
Pancreas Manipulation	No	•Pancreatic Contrast Injection •Pancreatic Guidewire Attempt	Pancreatic Stenting
Stricture of Junction of CBD and Papilla	No	Yes	---
Stone Extraction Attempts Number	X < 4 times	4 times ≤ X ≤ 6 times	•X > 6 times •Without Extraction Attempts
Stone (CBD) Clearance	Complete	Partial + CBD Stenting	•No •No + CBD Stenting
Stone Extraction Time	X < 10 minutes	10 minutes ≤ X ≤ 20 minutes	•X > 20 minutes •Without Extraction Attempts
Stone Number	X < 3	X = 3	X > 3
Largest Stone Diameter	X ≤ 15 mm	15 mm < X < 30 mm	X > 30 mm
Distal CBD Angulation	X ≥ 130 °	100 < X < 130 °	X ≤ 100 °
Proximal CBD Arm	X ≥ 40 mm	20 mm < X < 40 mm	X ≤ 20 mm
Distal CBD Arm	X ≥ 50 mm	25 mm < X < 50 mm	X ≤ 25 mm
Proximal CBD Diameter	X ≥ 12 mm	6 mm < X < 12 mm	X ≤ 6 mm
Mid-CBD Diameter	X ≥ 16 mm	8 mm < X < 16 mm	X ≤ 8 mm
Distal CBD Diameter	X ≥ 12 mm	6 mm < X < 12 mm	X ≤ 6 mm

CBD: Common Bile Duct

Table 3: Rule-Based Framework for Difficulty Classification in Endoscopic retrograde cholangiopancreatography (ERCP) Stone Extraction

Difficulty Class	Criteria
Easy	0 ≤ Total Score < 5
Intermediate	5 ≤ Total Score <10
Difficult	Total Score ≥ 10 or Stone size ≥ 30 mm or Distal CBD diameter ≤ 6 mm or Distal CBD angulation ≤ 100°

CBD: Common Bile Duct

implemented using Python version 3.10.1. The proposed intelligent system was evaluated using the 52 included ERCP cases to assess the accuracy and reliability of automated extraction of anatomical and morphological features from fluoroscopic images. Agreement between automated and expert-derived measurements was assessed using Bland Altman analysis and quantitative reliability metrics. The Bland–Altman results for the eight extracted features are shown in Figure 3.

For further evaluation, quantitative metrics were computed to assess and validate the accuracy of the automated algorithms, as summarized in Table 4.

A comprehensive analysis of the difficulty scoring system is presented in Table 5, which includes the mean ± standard deviation of the extracted features across the study population

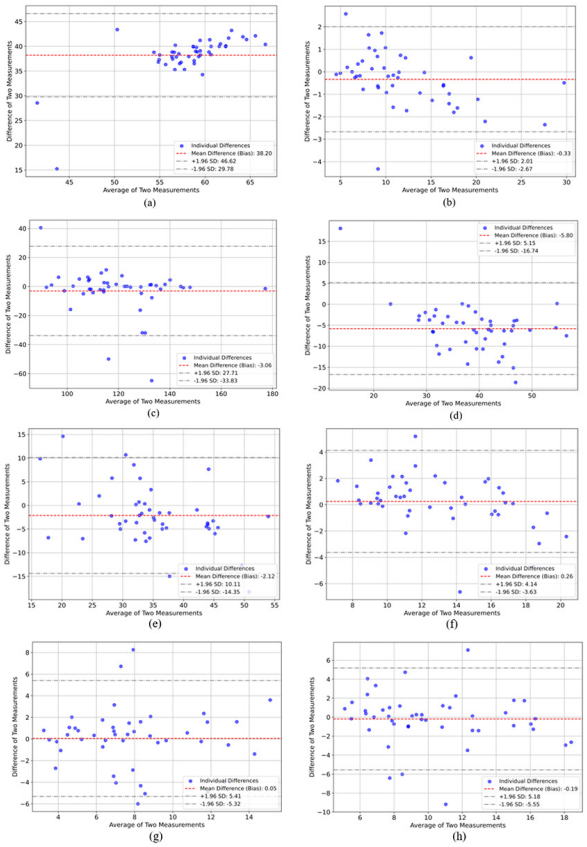


Figure 3: Bland-Altman plots for each parameter; (**a**) Duodenoscope Diameter (pix), (**b**) Largest Stone Size (mm), (**c**) Distal CBD (Common Bile Duct) Angulation (°), (**d**) Distal CBD Arm (mm), (**e**) Proximal CBD Arm (mm), (**f**) Mid-CBD Diameter (mm), (**g**) Distal CBD Diameter (tapering) (mm), (**h**) Proximal CBD Diameter (mm)

and the relationship between the estimated difficulty scores and each individual parameter.

Discussion

ERCP remains an indispensable procedure for the diagnosis and treatment of CBD stones. However, the procedural success, efficiency, and patient safety are heavily reliant on the endoscopist's ability to rapidly and accurately identify the location, size, and characteristics of stones, as well as the anatomical specifics of the bile duct.

This study evaluated anatomical and procedural predictors of ERCP difficulty. Several biliary anatomical variables showed significant differences across difficulty groups, while demographic factors and stone size alone were not independently predictive.

Sex, age, largest stone size alone, distal CBD angulation alone, guidewire attempts, and total extraction time were not significantly different between groups (P>0.05).

The findings highlight that duct geometry, rather than stone size alone, drives procedural difficulty. A narrow distal CBD significantly limits clearance space, reducing the ability to open baskets or sweep stones effectively. The stone-to-duct ratio therefore provides a more accurate measure of difficulty than stone size alone. Similarly, longer proximal and distal CBD arms facilitate maneuverability, providing mechanical advantage during extraction, while short arms increase procedural complexity. Interestingly, distal CBD angulation was not significant; however, the composite index involving angulation and stone size achieved significance. This suggests that angulation effects are context-dependent, becoming important when combined with other constraints such as small distal duct diameter. The higher number of stone extraction attempts in difficult cases likely reflects the anatomical constraints rather than operator technique alone, emphasizing the clinical relevance of these structural features.

These observations are consistent with prior studies reporting that distal CBD diameter, stone-to-duct ratio, and distal taper are critical predictors of difficult stone clearance [16-18]. Previous literature also supports the role of short distal arms and acute CBD angulation as independent risk factors for procedural difficulty [19,20]. While some studies have linked large stone size to complexity, recent evidence emphasizes that relative size (stone-to-duct ratio) and outlet constraints exert a greater effect on clearance difficulty than absolute stone dimensions [18,21]. Our findings align with this paradigm.

Table 4: Numerical metrics results of automatic algorithm

Feature Name	Mean Difference	Standard Deviation of Differences	Root Mean Square Error (RMSE)	Coefficient of Variation (CV)%	ICC%	CI95%	*P* value
Largest Stone Diameter	0.09	0.07	0.11	0.28	99.98	[1.0, 1.0]	<0.001
Distal CBD Angulation	0.37	0.12	0.38	0.33	99.94	[1.0, 1.0]	<0.001
Proximal CBD Arm	0.27	0.13	0.30	0.46	99.94	[1.0, 1.0]	<0.001
Distal CBD Arm	0.44	0.15	0.46	0.32	99.87	[1.0, 1.0]	<0.001
Proximal CBD Diameter	0.42	0.16	0.45	0.36	99.36	[0.99, 1.0]	<0.001
Mid-CBD Diameter	0.45	0.14	0.47	0.31	99.15	[0.99, 1.0]	<0.001
Distal CBD Diameter	0.33	0.15	0.36	0.44	99.13	[0.98, 1.0]	<0.001

CBD: Common Bile Duct, ICC: Intra-Class Correlation Coefficient, CI95%: Confidence Interval in 95%

Pre-procedural measurement of distal CBD diameter and calculation of stone-to-duct ratio on imaging could guide case triage. High ratios should prompt anticipation of advanced techniques such as endoscopic papillary large balloon dilation (EPLBD) or mechanical lithotripsy, ensuring equipment and expertise are available during the index ERCP. Similarly, cases with short arms or narrow ducts may benefit from early escalation rather than repeated standard extraction attempts, reducing procedure time and trauma. These findings also underscore the importance of scheduling anatomically complex cases with experienced endoscopists or in tertiary centers.

The small sample size and group imbalance (Easy group n = 5) may limit statistical power and external validity. Definitions of certain composite variables, particularly angulation-based indices, should be standardized for reproducibility. Additionally, multiple comparisons were performed; while many findings were highly significant, the potential for type I error should be acknowledged.

Future work should focus on prospective validation of simplified difficulty prediction scores incorporating distal CBD diameter, stone-to-duct ratio, and arm length. Identification of clinically actionable thresholds for these variables would enable standardized risk stratification. Integration of automated imaging analysis for diameter and angulation measurement, particularly in AI-assisted ERCP planning tools, represents a promising avenue for improving prediction accuracy and workflow efficiency.

A cornerstone of our research was the accurate extraction of quantitative anatomical and stone-related features directly from fluoroscopic images. These included stone diameter, distal and proximal CBD diameter, distal and proximal CBD arm length, and distal CBD angle. A small distal CBD diameter can necessitate extensive sphincterotomy or adjunctive techniques like balloon dilation, prolonging the procedure and increasing the risk of

complications. Our system's ability to provide precise, millimeter-level measurements, corrected for magnification artifacts using the duodenoscope as a reference, moves beyond subjective visual estimation. This objective data empowers endoscopists, particularly those with less experience, to make more informed decisions regarding pre-procedural planning, choice of extraction tools, and anticipation of potential challenges, ultimately contributing to a more streamlined and successful ERCP.

The integration of these extracted features with demographic and procedural characteristics into a novel difficulty scoring system further distinguishes our approach. Developed through a rigorous modified Delphi consensus process with five highly experienced endoscopists, this scoring scale offers a nuanced

Table 5: Statistical Analysis and Evaluation of Extracted Feature Metrics and Difficulty Scoring System Outcomes

Characteristics	Total, n = 52	Difficult group = 24	Intermediate group = 23	Easy group = 5	*P* value
Sex, M/F, n	24/28	9/15	12/11	3/2	0.492
Age, mean (SD), years	61.87 (18.44)	57.2 (18.11)	62.82 (15.93)	65.8 (21.28)	0.461
Largest stone size, mean (SD), mm	12.27 (5.46)	12.51 (6.50)	11.68 (5.23)	13.83 (4.63)	0.665
Distal CBD Angulation, mean (SD), degree	119.99 (16.66)	117.76 (20.78)	123.35 (20.82)	115.19 (8.39)	0.844
Proximal CBD Arm, mean (SD), mm	35.89 (8.78)	34.31 (11.22)	34.61 (8.22)	49.34 (6.90)	**0.010**
Distal CBD Arm, mean (SD), mm	42.03 (11.23)	39.96 (11.30)	41.23 (8.51)	55.60 (13.89)	**0.043**
Proximal CBD Diameter, mean (SD), mm	10.57 (4.11)	9.11 (4.04)	11.23 (4.90)	14.48 (3.39)	**0.042**
Mid-CBD Diameter, mean (SD), mm	13.06 (3.25)	11.81 (4.16)	13.56 (4.24)	16.67 (1.35)	**0.028**
Distal CBD Diameter, mean (SD), mm	7.91 (2.71)	5.95 (3.08)	9.00 (2.36)	12.22 (2.70)	**<0.001**
CBD Guidewire Attempts Number, mean (SD)	2.82 (2.02)	3.37 (3.30)	2.52 (2.21)	1.61 (0.54)	0.370
Stone Extraction Attempts Number, mean (SD)	2.84 (0.97)	3.33 (1.23)	2.47 (1.23)	2.23 (0.44)	**0.016**
Stone Extraction Time, mean (SD), minutes	15.23 (11.19)	18.25 (13.57)	13.56 (11.13)	8.42 (8.87)	0.143
Stone-to-distal-CBD diameter ratio, mean (SD)	1.77 (1.14)	2.34 (1.39)	1.30 (0.53)	1.20 (0.55)	**<0.001**
Composite index involving stone size, distal CBD diameter and angulation, mean (SD)	0.24 (0.13)	0.20 (0.13)	0.26 (0.11)	0.35 (0.11)	**0.003**

CBD: Common Bile Duct, SD: Standard Deviation

classification of extraction difficulty (easy, intermediate, difficult). Importantly, the system incorporates "special conditions" (e.g., stone size >30 mm, distal CBD diameter <6 mm, distal CBD angulation <100°), which automatically classify a case as "difficult" regardless of the cumulative score from other factors. This sophisticated tiered classification directly addresses the complexities encountered in real-world ERCP scenarios, providing a flexible yet robust framework that can guide endoscopists in anticipating and managing challenging cases. This proactive identification of difficulty aligns with the findings of Huang et al. [8], who emphasized the influence of stone and duct characteristics on technical difficulty and the need for a smart auxiliary system.

The principal contribution of this work is the development of an integrated expert system that combines automated image-derived measurements with demographic and procedural variables to predict procedural difficulty. Unlike previous works that might focus on specific aspects like segmentation or clinical scoring, our system uniquely combines automated, precise fluoroscopic feature extraction with a multi-factor difficulty scoring scale. The meticulous creation of a gold standard database from 422 ERCP cases, annotated by highly experienced specialists, provides a solid foundation for the developed algorithms. Moreover, the ingenious use of the duodenoscope's fixed diameter as an internal calibration reference for pixel-to-millimeter conversion effectively mitigates magnification artifacts common in X-ray imaging, ensuring the accuracy and clinical utility of the quantitative measurements. This pragmatic solution addresses a significant challenge in translating image-based measurements to real-world anatomical sizes.

Despite these significant advancements, our study is not without limitations. The dataset, while substantial, was collected from a single hospital in Shiraz, Iran. This regional specificity means that the generalizability of our difficulty scoring system and predictive models to diverse global populations, different anatomical variations, or varying ERCP practices (e.g., types of equipment, endoscopist training styles) would require further validation through large-scale, multi-center prospective studies. Additionally, while the selected eight imaging features are clinically significant, there might be other subtle or dynamic fluoroscopic indicators that could further refine the prediction of difficulty or complications, warranting future exploration. The current system provides quantitative data and predictions, but its full clinical utility would be realized with the development of an intuitive, real-time graphical user interface integrated directly into the ERCP suite, allowing immediate feedback to the endoscopist. Future research could also explore the application of more advanced deep learning architectures for end-to-end learning from raw fluoroscopic images, potentially capturing more complex, subtle patterns than explicitly defined features.

Conclusion

In conclusion, our automated image-processing and expert-based assistant system represents a pivotal advancement in the field of ERCP for CBD stone management. By providing objective and quantitative measurements of stone and CBD characteristics and supporting difficulty classification, the proposed framework may assist clinical decision-making and procedural planning during ERCP. It stands to significantly enhance the endoscopist's ability to make informed decisions, optimize procedural planning, reduce operative time, and ultimately elevate patient safety and outcomes. As the landscape of interventional endoscopy evolves, such intelligent auxiliary systems will play an increasingly vital role in achieving higher standards of care.

Acknowledgment

The authors would like to thank Shiraz University of Medical Sciences for supporting

this study.

Authors' Contribution

MM. Asghari and T. Pishro Dabaghiyan were involved in data collection, algorithm implementation, and manuscript preparation. T. Mahmoudi and R. Niknam contributed to study supervision, conceptualization, methodology design, and final manuscript review. H. Parsaei and MM. Movahedi participated in supervision and methodological development. K. Bagheri Lankarani, F. Ejtehadi, SA. Malek-Hosseini, SE. Fallahzadeh Abarghooei, S. Alireza Taghavi, and GhR. Sivandzadeh contributed to conceptualization, data collection, and annotation verification. S. Alihosseini was involved in data collection. All authors read and approved the final version of the manuscript.

Ethical Approval

This study was approved by the Institutional Review Board of Shiraz University of Medical Sciences (IR.SUMS.MED.REC.1403.031) and was conducted in accordance with the tenets of the Declaration of Helsinki. All methods were performed in accordance with the relevant guidelines and regulations.

Informed Consent

Informed consent was obtained from all individual participants included in the study.

Funding

This work was supported by Shiraz University of Medical Sciences as part of a Master's thesis project (Grant No. 29384).

Conflict of Interest

None

References

1. McNicoll CF, Pastorino A, Farooq U, Froehlich MJ, St Hill CR. Choledocholithiasis. InStatPearls [Internet]. StatPearls Publishing; 2023.
2. Galeazzi M, Mazzola P, Valcarcel B, Bellelli G, Dinelli M, Pasinetti GM, Annoni G. Endoscopic retrograde cholangiopancreatography in the elderly: results of a retrospective study and a geriatricians' point of view. *BMC Gastroenterol.* 2018;**18**(1):38. doi: 10.1186/s12876-018-0764-4. PubMed PMID: 29540171. PubMed PMCID: PMC5853060.
3. Evans N, Buxbaum JL. Endoscopic treatment of ERCP-related duodenal perforation. *Techniques in Gastrointestinal Endoscopy.* 2019;**21**(2):83-90. doi: 10.1016/j.tgie.2019.04.002.
4. Nzenza TC, Al-Habbal Y, Guerra GR, Manolas S, Yong T, McQuillan T. Recurrent common bile duct stones as a late complication of endoscopic sphincterotomy. *BMC Gastroenterol.* 2018;**18**(1):39. doi: 10.1186/s12876-018-0765-3. PubMed PMID: 29544453. PubMed PMCID: PMC5856321.
5. Schutz SM, Abbott RM. Grading ERCPs by degree of difficulty: a new concept to produce more meaningful outcome data. *Gastrointest Endosc.* 2000;**51**(5):535-9. doi: 10.1016/s0016-5107(00)70285-9. PubMed PMID: 10805837.
6. McHenry L, Lehman G. Difficult bile duct stones. *Curr Treat Options Gastro.* 2006;**9**(2):123-32. doi: 10.1007/s11938-006-0031-6.
7. Kim HJ, Choi HS, Park JH, Park DI, Cho YK, Sohn CI, et al. Factors influencing the technical difficulty of endoscopic clearance of bile duct stones. *Gastrointest Endosc.* 2007;**66**(6):1154-60. doi: 10.1016/j.gie.2007.04.033. PubMed PMID: 17945223.
8. Huang L, Lu X, Huang X, Zou X, Wu L, Zhou Z, et al. Intelligent difficulty scoring and assistance system for endoscopic extraction of common bile duct stones based on deep learning: multicenter study. *Endoscopy.* 2021;**53**(5):491-8. doi: 10.1055/a-1244-5698. PubMed PMID: 32838430.
9. Williams E, Beckingham I, El Sayed G, Gurusamy K, Sturgess R, Webster G, Young T. Updated guideline on the management of common bile duct stones (CBDS). *Gut.* 2017;**66**(5):765-82. doi: 10.1136/gutjnl-2016-312317. PubMed PMID: 28122906.
10. Wongkanong C, Patumanond J, Ratanachu-Ek T, Junrungsee S, Tantraworasin A. Common bile duct stone-related parameters and their potential prognostic values for stone extraction requiring endoscopic papillary large balloon dilation. *JGH Open.* 2022;**7**(1):16-23. doi: 10.1002/jgh3.12845. PubMed PMID: 36660050. PubMed PMCID: PMC9840188.
11. Wongkanong C, Patumanond J, Ratanachu-Ek T, Junrungsee S, Tantraworasin A. Development of clinical prediction rule for the requirement of endoscopic papillary large balloon dilation

(EPLBD) on endoscopic CBD stone clearance. *PLoS One.* 2023;**18**(3):e0282899. doi: 10.1371/journal.pone.0282899. PubMed PMID: 36928213. PubMed PMCID: PMC10019717.

12. Han JY, Lee DH, Jeong S, Choi HJ, Moon JH, Han JH, et al. Clinical Features and Outcomes of Endoscopic Treatment for Stones in Stemware-Shaped Common Bile Ducts: A Multicenter Data Analysis. *Gut Liver.* 2015;**9**(6):800-4. doi: 10.5009/gnl14433. PubMed PMID: 26087795. PubMed PMCID: PMC4625711.
13. Ragunath K, Thomas LA, Cheung WY, Duane PD, Richards DG. Objective evaluation of ERCP procedures: a simple grading scale for evaluating technical difficulty. *Postgrad Med J.* 2003;**79**(934):467-70. doi: 10.1136/pmj.79.934.467. PubMed PMID: 12954961. PubMed PMCID: PMC1742772.
14. Williams EJ, Ogollah R, Thomas P, Logan RF, Martin D, Wilkinson ML, Lombard M. What predicts failed cannulation and therapy at ERCP? Results of a large-scale multicenter analysis. *Endoscopy.* 2012;**44**(7):674-83. doi: 10.1055/s-0032-1309345. PubMed PMID: 22696192.
15. Haseeb A, Freeman ML. Endoscopic Papillary Large Balloon Dilation Versus Endoscopic Sphincterotomy for Treatment of Bile Duct Stones. *Curr Treat Options Gastro.* 2019;**17**(2):221-30. doi: 10.1007/s11938-019-00234-5. PubMed PMID: 30945095.
16. Garg PK, Tandon RK, Ahuja V, Makharia GK, Batra Y. Predictors of unsuccessful mechanical lithotripsy and endoscopic clearance of large bile duct stones. *Gastrointest Endosc.* 2004;**59**(6):601-5. doi: 10.1016/s0016-5107(04)00295-0. PubMed PMID: 15114300.
17. Üsküdar O, Parlak E, Dışıbeyaz S, Köksal AS, Cıçek B, Kılıç ZM, et al. Major predictors for difficult common bile duct stone. *Turk J Gastroenterol.* 2013;**24**(5):423-9. doi: 10.4318/tjg.2013.0511. PubMed PMID: 24557966.
18. El Menabawey T, Mulay A, Graham D, Phillpotts S, Sethi A, Webster GJ. Predictors of success of conventional ERCP for bile duct stones and need for single-operator cholangioscopy. *Endosc Int Open.* 2023;**11**(10):E943-51. doi: 10.1055/a-2164-8557. PubMed PMID: 37818453. PubMed PMCID: PMC10562053.
19. Ryu S, Jo IH, Kim S, Kim YJ, Chung WC. Clinical Impact of Common Bile Duct Angulation on the Recurrence of Common Bile Duct Stone: A Meta-analysis and Review. *Korean J Gastroenterol.* 2020;**76**(4):199-205. doi: 10.4166/kjg.2020.76.4.199. PubMed PMID: 33100315. PubMed PMCID: PMC12286447.
20. Shimatani M, Hatanaka H, Kogure H, Tsutsumi K, Kawashima H, Hanada K, et al. Diagnostic and Therapeutic Endoscopic Retrograde Cholangiography Using a Short-Type Double-Balloon Endoscope in Patients With Altered Gastrointestinal Anatomy: A Multicenter Prospective Study in Japan. *Am J Gastroenterol.* 2016;**111**(12):1750-8. doi: 10.1038/ajg.2016.420. PubMed PMID: 27670601.
21. Beran A, Abboursheid T, Ali AH, Nayfeh T, Albunni H, Vargas A, et al. Predictors of Post-endoscopic Retrograde Cholangiopancreatography Pancreatitis: A Comprehensive Systematic Review and Meta-analysis. *Clin Gastroenterol Hepatol.* 2025;**23**(11):1905-16. doi: 10.1016/j.cgh.2024.11.014. PubMed PMID: 39694210.