Research Article

Artificial Intelligence-Assisted Visual Sensing Technology under Duodenoscopy of Gallbladder Stones

Dongyan Li,[1] **Bin Du,**[1] **Youhui Shen,**[1] **and Lin Ge** [2]

[1]*Department of Endoscopy, Tianjin TEDA Hospital, Tianjin 300457, China*
[2]*Department of Biochemistry and Molecular Biology, Tianjin Medical University, Tianjin 300070, China*

Correspondence should be addressed to Lin Ge; qel@tmu.edu.cn

Academic Editor: Haibin Lv

Objective. This study is aimed at exploring the application effect of duodenoscopy assisted by visual sensing technology based on convolutional neural network (CNN) segmentation algorithm in the diagnosis and treatment of gallbladder stones, so as to provide safer and more effective treatment methods for patients with gallstones. *Methods*. 188 patients with gallstones and choledocholithiasis who were admitted to our hospital from January 2016 to April 2021 were selected as the research objects. Based on whether the patients were willing to use AI-assisted visual sensing technology during the treatment process, all patients were divided into two groups, namely, the AI group and the conventional group. Various surgical indicators of patients in two groups were compared. *Results*. The precision, recall, and mean intersection ratio of the M-Unet-based segmentation algorithm were 94.56%, 96.56%, and 98.92%, respectively. In the AI group, the operation time $(2.74 \pm 0.45 \, \text{h})$, postoperative drainage tube placement time $(4.31 \pm 1.15 \, \text{d})$, time required for recovery of gastrointestinal function $(1.74 \pm 0.54 \, \text{d})$, time required to get out of bed $(1.14 \pm 0.55 \, \text{h})$, and time spent in hospital $(9.94 \pm 1.45 \, \text{d})$ were all shorter compared with those in the conventional group, which were $3.21 \pm 0.32 \, \text{h}$, $12.14 \pm 2.98 \, \text{d}$, $2.89 \pm 0.67 \, \text{d}$, $2.09 \pm 0.87 \, \text{h}$, and $14.14 \pm 1.15 \, \text{h}$, showing statistical differences $(P < 0.05)$; the intraoperative blood loss $(79.74 \pm 6.45 \, \text{mL})$ and residual status of stones (0%) in the AI group were much lower than those in the conventional group $(P < 0.05)$. In addition, the incidence of complications (10.26%) and the indicators of postoperative gallbladder function of patients in the AI group were lower greatly than those in the conventional group $(P < 0.05)$. *Conclusion*. The visual sensing technology assisted by the CNN algorithm showed a good effect on image processing, and endoscopic technology can effectively improve the treatment effect of gallbladder stones combined with choledocholithiasis with the aid of this technology. Therefore, the conclusion in this study proved that visual sensing technology based on intelligent algorithms showed a good future in the medical field.

1. Introduction

The current society is in an era of rapid development of technology and civilization. People's living standards and quality of life have been greatly improved, which is a beneficial aspect of social development. However, it is not only the beneficial side that develops along with it, for example, various diseases also closely follow. Nowadays, people do not pay enough attention to dietary hygiene and dietary rules, so the probability of occurrence of digestive system diseases is becoming higher and higher. These digestive system diseases include gallbladder and biliary system diseases, espe-

cially stone diseases of the gallbladder and biliary tract [1, 2]. According to relevant research statistics, the incidence of gallbladder and biliary calculus diseases is basically about 10% among the adult population in China; and in this type of disease, the probability of patients with gallbladder stones combined with choledocholithiasis is about 12%~15% [3]. Therefore, a large number of clinical applications have been done to provide patients with gallbladder stones combined with choledocholithiasis with more convenient and effective treatment methods. After continuous efforts in researches of many experts and researchers, endoscopic technology has been introduced into the clinical treatment of gallbladder

stones and has been widely applied in clinical medicine today. However, laparoscopic technology and duodenoscopy are often used clinically in the treatment of this disease due to the complicated condition of gallbladder stones combined with choledocholithiasis [4]. A duodenoscope is mostly used to guide the implementation of bile duct stone removal. It refers to the method of pulling the bile duct stones out of the bile duct using net baskets, balloons, and other instruments under a duodenoscope. The main indications are choledocholithiasis, hepatolithiasis, and gallbladder stones [5]. Laparoscopy is mostly used to guide gallbladder resection. The main operation method is to insert a special catheter from the abdomen into the peritoneal cavity and then infuse about 2~5 L of carbon dioxide into the peritoneal cavity. When the abdominal pressure reaches a certain level, three incisions of 0.5~1.5 cm are made in the abdomen to dissect the triangle structure of the gallbladder. Then, the cystic duct and cystic artery are separated and clamped, and then, the entire gallbladder including the stones is resected [6, 7]. Although the treatment effect has made people feel satisfied, there is still a big gap between the effect observed through the display mirror and the effect under direct eyes. With the rapid development of science and technology these years, the artificial intelligence (AI) technology has been widely used in various fields, among which AI and visual sensing technology have developed rapidly. The main function of this technology is to add perception technology to the visual sensor so that it can segment, detect, and recognize the network architecture of the target object [8]. The basic theory of this technology is the convolutional neural network (CNN) algorithm [9], and visual sensing technology assisted by the CNN system has gradually reached or exceeded the recognition level of humans. In the past, this technology was mostly used for segmentation and identification of commodity goods [10] and measurement of pitting of gears in shooting ranges [11]. Its application in the field of medical examination is relatively small. To provide a safer and more effective treatment for patients with gallstones, the patients with gallstones and choledocholithiasis were selected as the research objects in this study. In addition, the laparoscopic technology, duodenoscopy, and visual sensing technology assisted by CNN segmentation algorithm were combined for research and analysis of the treatment of gallbladder stones combined with common bile duct stones, hoping that this method can give research basis for clinical treatment.

2. Materials and Methods

2.1. Visual Sensing Technology Based on CNN. The composition of the CNN system is based on the traditional neural network. CNN is a recognition process that gradually develops from a part to a whole in the aspect of biometric recognition [12]. Compared with the traditional neural networks [13, 14], the CNN system shows excellent features such as local receptive fields, shared weights, and multicore convolution. The existence of above advantages can not only greatly reduce the parameters in the neural network but also obtain different levels of characteristic information through the hierarchical network structure. The specific structure of the CNN system is shown in Figure 1. Its main components can be divided into a convolutional layer, an activation function, a convergence layer, and a fully connected layer. Its own parameters were optimized using an iterative optimization algorithm, and the network structure was converged, so as to further achieve the convergence of target loss function. Next, each layer of CNN would be explained and described in detail.

The convolutional layer is a type of local connection, which can reduce the number of model parameters. Each network neuron in the convolutional layer will only be connected to the local receptive field of the previous layer. However, the size of the local receptive field is mainly determined by the parameter size of the convolution kernel. The representative way of the two-dimensional convolution in the convolutional layer in CNN was as follows:

$$s(i, j) = (Q * W)(i, j) = \sum_m \sum_n Q(i + m, j + n) W(m, n). \quad (1)$$

In the above equation, $s(i, j)$ referred to result of the convolution of the target image; $(Q * W)$ was multiplication of the local area Q of the target image and the elements of each position in the convolution kernel W matrix; (i, j) represented the pixel at the target position of the image; and m and n referred to the number of parameters of the local area Q and the convolution kernel W of the target image, respectively.

The activation layer is generally connected after the convolutional layer, and the data is nonlinearly transformed by the activation function in the activation layer. The more commonly used activation functions include sigmoid activation function, tanh activation function, and ReLU activation function. The specific expressions of these three activation functions were as follows:

$$f(x) = \frac{1}{1 + e^{-x}}, \quad (2)$$

$$f(x) = \tanh(x), \quad (3)$$

$$f(x) = \max(x, 0). \quad (4)$$

In the expression equations of the above three activation functions, the ReLU activation function with relatively good effect was adopted in this study.

The main function of the pooling layer itself is to perform lower-level sampling on the feature map, so as to reduce the parameters required for calculation and overfitting and further improve the stability of CNN. There were two main pooling methods adopted in this study, namely, regional maximum pooling and average pooling.

The Softmax classifier belongs to a multinomial logistic regression model, which can be classified as a log-linear model. Most of the time, it is used in binary classification.
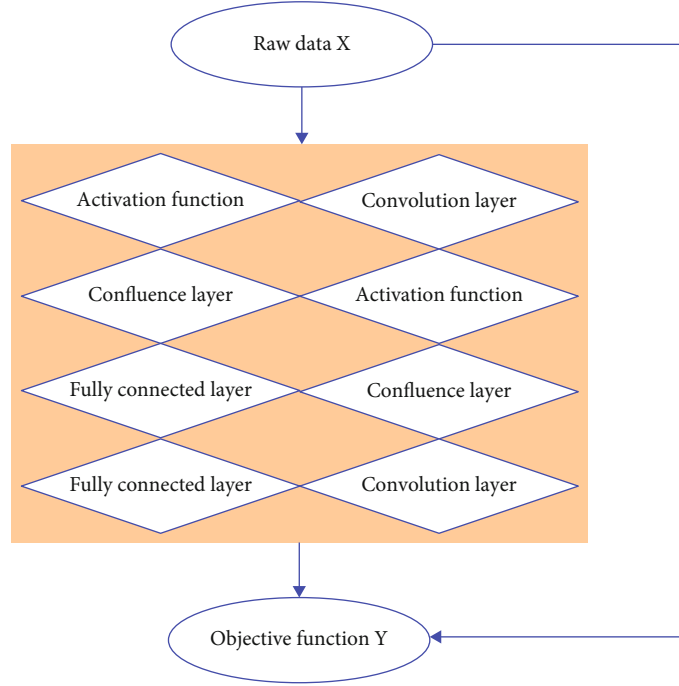
Figure 1: The specific structure of the CNN system.

The specific calculation process of the Softmax classifier can be expressed as follows:

$$P(Y|x) = \frac{p(w_k \cdot x)}{1 + \sum_{k=1}^{K-1} p(w_k \cdot x)}, \quad k = 1, 2, 3, \cdots, K-1, \quad (5)$$

$$P(Y|x) = \frac{1}{1 + \sum_{k=1}^{K-1} p(w_k \cdot x)}, \quad k = K, \quad (6)$$

$$P(Y|x) = \frac{p(w_k \cdot x)}{\sum_{k=1}^{K-1} p(w_k \cdot x)}, \quad k = 1, 2, 3, \cdots, K. \quad (7)$$

In the above three equations, W referred to the required parameter model; q was an input vector; $P(Y \mid q)$ referred to the probability of predicting the category of q as category Y, and the possible value of Y was $k = 1, 2, 3, \cdots, K$. The dataset required for training was given as $U = \{(x_1, y_1), (x_2, y_2), \cdots (x_m, y_m)\}$ ($x_i \in R_n$ and $y_i \in \{1, 2, \cdots, K\}$), and y_i and x_i corresponded to the category label. The maximum likelihood estimation method was adopted in this study to calculate the result of the model. Therefore, the function of x_i could be expressed as

$$\mathbb{R}_{j=1}^{K} 1\{Y = j\} P(Y = j | x_i ; w) = \mathbb{R}_{j=1}^{K} 1\{Y = j\} \frac{p\left(w_j^U \cdot x\right)}{\sum_{k=1}^{K} p\left(w_j^U \cdot x\right)}. \quad (8)$$

Then, the expression of its logarithmic function was

given as follows:

$$\sum_{j}^{k} 1\{Y = j\} \log \frac{p\left(w_j^U \cdot x\right)}{\sum_{k=1}^{K} p\left(w_j^U \cdot x\right)}. \quad (9)$$

The expression of $1\{\text{with true value was}\} = 1$, and the expression of $1\{\text{with false value was}\} = 0$. For the entire dataset U, its log likelihood function was written as follows:

$$\frac{1}{m} \left[\sum_{i=1}^{m} \sum_{j=1}^{K} 1\{y_i = j\} \right] \log \frac{p\left(w_j^U \cdot x\right)}{\sum_{k=1}^{K} p\left(w_j^U \cdot x\right)}. \quad (10)$$

Maximizing the log-likelihood function was equivalent to minimizing the loss function $J(w)$:

$$J(w) = \frac{1}{m} \left[\sum_{i=1}^{m} \sum_{j=1}^{K} 1\{y_i = j\} \right] \log \frac{p\left(w_j^U \cdot x\right)}{\sum_{k=1}^{K} p\left(w_j^U \cdot x\right)}. \quad (11)$$

In equation (11), when the value of $J(w)$ was the smallest, w was the model parameter to be sought; and $J(w)$ was a convex function. Then, the method was optimized to make the result converge to the best in the whole process.

Adam optimizer [15] combines the advantages of two types of optimization algorithms: AdaGrad optimization algorithm and RMSProp optimization algorithm. The optimization method is to use the dynamic changes of the first-order moment estimation (the average of all gradients) and the second-order moment estimation (the variance of

the gradient without centering) to adjust each parameter of learning rate.

Then, the expressions of the first-order moment m_u and the second-order moment o_u of the u-th iteration are shown in

$$m_u = \alpha m_{u-1} + (1-\alpha) g_u, \qquad (12)$$

$$o_u = \beta o_{u-1} + (1-\beta) g_u^2. \qquad (13)$$

In the above two equations, α and β were parameter values, which were defaulted to 0.9 and 0.99, respectively. Then, the required unbiased estimation of the first and second moments required proper processing of m_u and o_u. The specific processing process was as follows:

$$m_u' = \frac{m_u}{1-\alpha_u}, \qquad (14)$$

$$o_u' = \frac{o_u}{1-\beta^u}. \qquad (15)$$

After further improvement, the manifestation of the parameters can be expressed as below equation:

$$\chi_u = -\frac{\delta}{\sqrt{o_u' + \varepsilon}} m_u'. \qquad (16)$$

After Adam was corrected by equations (14) and (15), the range of the learning rate of each iteration can be determined, and the parameters were relatively stable.

The propagation algorithm in CNN can be classified into forward propagation algorithm and backward propagation algorithm.

The forward propagation algorithm [16] refers to the process in which information starts from the input layer and then propagates forward in a one-way propagation manner, penetrating the neural network to the output layer, and finally being output by the network. The information output from the $(h-1)$-th layer in the overpropagation process was the input information of the h-th layer. It was supposed that Z^h was the output information of the $(h-1)$-th layer, where $h = 1, 2, 3, \cdots, H$.

It was assumed that the h-th layer was a convolutional layer, and then, the algorithm at this layer can be expressed as follows:

$$Z^h = f^{h-1}\left(W^h Z^{h-1} + b^h\right). \qquad (17)$$

If the h-th layer was assumed as the pooling layer, then the algorithm at this layer can be written as follows:

$$Z^h = P\left(Z^{h-1}\right). \qquad (18)$$

If the h-th layer was assumed as the output layer, then

the algorithm at this layer can be written as follows:

$$Z^H = C\left(W^h Z^{h-1} + b^h\right). \qquad (19)$$

In equations (17)–(19), W^h and b^h represented the weight and bias of the h-th layer, respectively; f^{h-1} was the activation function in the $(h-1)$-th layer, which was the ReLU activation function mentioned above; P referred to the process of pooling the input through the fixed pooling function; and C was the Softmax classifier mentioned above.

The backpropagation algorithm [17, 18] refers to propagating the residuals backwards during the training process and then gradually correcting the parameters in the CNN, which greatly simplifies the solution of the partial derivatives of the variables in the multilayer composite function. The specific calculation method was shown in equations (20)–(25).

The expression equation of the loss function can be assumed as follows:

$$J(W, b\,;x, y) = \frac{1}{2}\left\|a_{W,b}(x) - y\right\|^2. \qquad (20)$$

The variables in equation (20) were specifically expressed as follows. x was an input information; W and b represented the parameters and bias of the network model, respectively; $a_{W,b}(x) - y$ referred to the probability that the network predicted the input x as the category y. The activation value was calculated based on the forward propagation algorithm, and then, the residual error ∂_i^h of the i-th neuron node in the h-th layer was calculated. The specific calculation equation was expressed as follows.

$$\partial_i^h = \frac{\kappa}{\kappa d_i^h} J(W, b\,;x, y). \qquad (21)$$

In the equation above, d_i^h referred to i-th unit of the h-th layer, including the weighted sum of the bias units, and had not been processed by the activation function.

If the h-th layer was a convolutional layer, then the $(h+1)$-th layer was the pooling layer, and the residual calculation could be realized with below equation:

$$\partial_i^h = \text{upsample}\left(\partial_i^{h+1}\right) \odot f'\left(d_i^h\right). \qquad (22)$$

In equation (22), $\odot$ referred to the dot product operation of matrix, upsample was to restore the residual matrix to before pooling, and $f'(d_i^h)$ was the derivative value of activation function f to d^h.

If the h-th layer was a pooling layer, the $(h+1)$-th layer was a convolutional layer, and the residual could be calculated with

$$\partial_i^h = \partial_i^{h+1} * \text{rot}180\left(W^{h+1}\right) \odot f'\left(d_i^h\right). \qquad (23)$$

In the equation above, $*$ was convolution, and rot180

referred to flipping up and down once and then flipping once left and right.

Finally, the partial derivatives of each layer could be written as

$$\frac{\kappa}{\kappa W_{ij}^h} J(W, b; x, y) = \partial_i^{(h+1)} Z_j^{(h)}, \tag{24}$$

$$\frac{\kappa}{\kappa b_i^{(h)}} J(W, b; x, y) = \partial_i^{(h+1)}. \tag{25}$$

A multipath dilated convolution (MPDC) [19] is proposed to further optimize the CNN image segmentation algorithm. The MPDC-Unet (M-Unet) was adopted for description in the following text (Figure 2), and the Soft Dice loss was applied for program supervision. In this network, its overall architecture was established in the way of encoding and decoding. It should encode (i.e., downsampling) and decode (i.e., upsampling) firstly and then get the segmentation result mask with the same size as the initial image. The encoding process used Max Pooling for double downsampling, and multiple convolutions were used between the two pooling layers to encode features. In the decoding stage, deconvolution was to gradually increase the resolution of the feature map, and the large convolution kernel was used to encode the context information. The low-level output feature map was firstly cropped to make it the same spatial size as the corresponding high-level feature map and then spliced in the channel dimension to achieve feature fusion so that the low-level and high-level features can be fused better.

The most commonly used loss function in semantic segmentation [20] is the pixel-wise cross-entropy loss. For each pixel, the predicted category was compared with the true value category, and the loss of all pixels was averaged, that was, the loss of each pixel in the image was treated equally. In this study, the Soft Dice coefficient was undertaken as the loss function, which could be expressed as follows:

$$\text{Soft Dice} = \frac{2\sum_i^M p_i g_i}{\sum_i^M p_i^2 + \sum_i^M g_i^2}. \tag{26}$$

In the equation above, p_i and g_i represented the value of the prediction mask and the true value mask, respectively. The cross-entropy loss function paid more attention to the global pixel classification including the background and foreground, which exerted a better effect on reducing missegmentation. Therefore, the Soft Dice loss and cross-entropy weighted summation were adopted in this study to replace the traditional cross-entropy loss function in M-Unet as shown in

$$V = w_1 \times \text{Soft Dice} + w_2 \times \text{Cross Entropy}. \tag{27}$$

In the above equation, when the values of w_1 and w_2 were both 0.5, the result was the best. The main process of image segmentation based on the M-Unet network model is shown in Figure 3.

The indicators commonly used to evaluate the effect of semantic segmentation included precision, recall, and mean intersection ratio. If there were two sets of pixel sets for the research target, then the actual set GT and the predicted set P can be defined as follows. The first referred to the correctly predicted pixels (true positive, TP), and TP referred to the intersection of the two sets; the second was false positive (FP), which referred to the predicted wrong pixel in the predicted pixel set; and the last was false negative (FN), which referred to the wrongly predicted pixels in the set of true value pixels. According to the above definitions, the higher the precision (that was, the ratio of the predicted correct pixel to the predicted total pixel), the higher the FP was (the fewer false predictions). The precision can be calculated with equation (28). The higher the recall (the ratio of the predicted correct pixel to the true value pixel) was, the less FN was, which meant the less missed detection of segmentation result. It can be calculated with equation (29). mIoU was the intersection ratio of two sets, which was used to judge the coincidence rate between the prediction result and the target true value. The higher the mean intersection ratio, the more accurate the segmentation algorithm. Its calculation expression was given as equation (30).

$$\text{Pre} = \frac{\text{TP}}{(\text{TP} + \text{FP})}, \tag{28}$$

$$\text{Rec} = \frac{\text{TP}}{(\text{TP} + \text{FN})}, \tag{29}$$

$$\text{mIoU} = \frac{\text{TP}}{(\text{TP} + \text{FP} + \text{FN})}. \tag{30}$$

In the above three equations, Pre, Rec, and mIoU referred to precision, recall, and mean intersection ratio, respectively.

2.2. Research Objects. The patients who were diagnosed with gallbladder stones combined with choledocholithiasis during January 2016 to April 2021 were randomly selected as the research objects with a total of 188 cases. Then, based on whether all patients were willing to accept the AI-assisted and visual sensing technology during the treatment process, all patients were divided into two groups: AI group and conventional group. According to statistics, there were 78 patients in the AI group, of which 38 were males and 40 were females. They were all between 26 and 75 years old, with an average age of 42.72 ± 2.87 years; and the course of disease was 1.17~5.67 months with the average value of 3.37 ± 0.78 months; and the number of stones was 17 cases of single stones and 61 cases of multiple stones. There were 110 patients in the conventional group, including 51 males and 59 females; they were aged 24~74 years with the average age of 42.45 ± 3.12 years; the disease course was also 1.16~5.68 months (with an average course of the disease of 3.44 ± 0.12 months); and there were 26 cases of single stones and 84 cases of multiple stones. The research had been applied to the medical ethics committee of our hospital, and the research permission had been obtained.
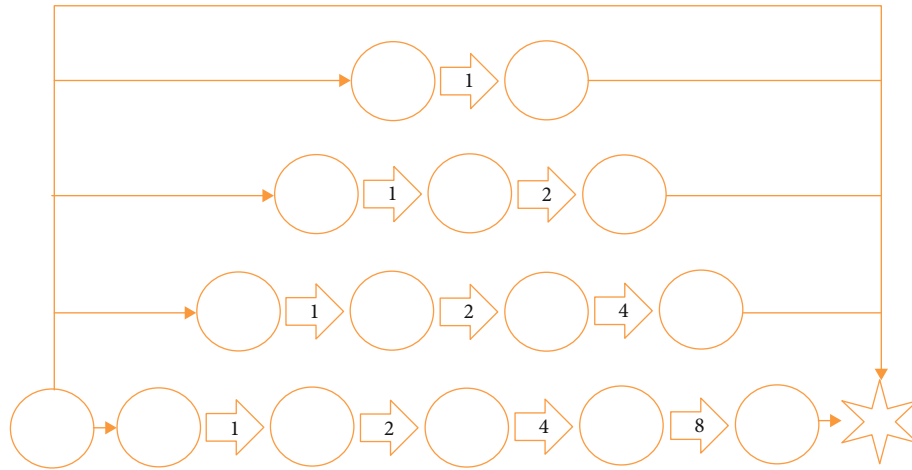
FIGURE 2: Module diagram of MPDC (the number in the arrow indicated the void ratio).
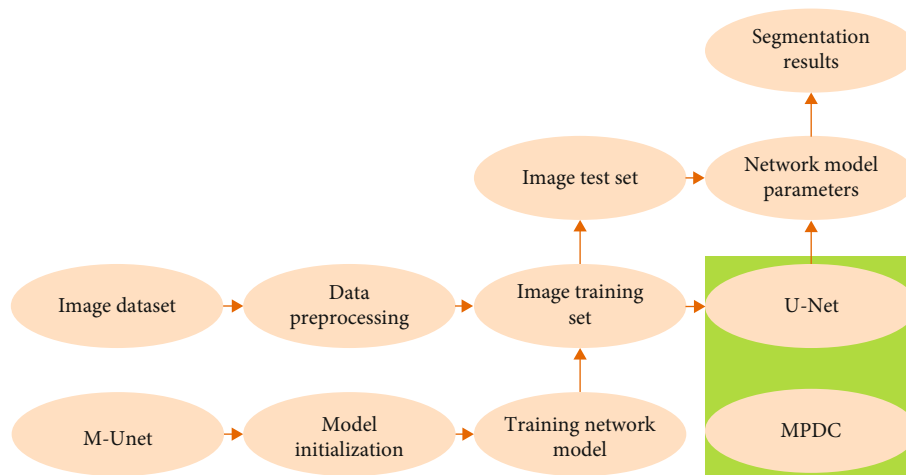

FIGURE 3: The main process of image segmentation based on the M-Unet network model.

The inclusion criteria were defined as follows: (a) all patients had been diagnosed with gallbladder stones combined with choledocholithiasis by CT or B-ultrasound and suffered from clinical manifestations such as jaundice, abdominal pain, nausea, and vomiting before the surgery; (b) after diagnosis, all patients were found to have indications for laparoscopy and duodenoscope diagnosis and treatment; (c) routine hematuria examinations of all patients showed that they were normal, and they could follow the doctor's instructions well; and (d) all patients and their family members had signed the informed consent forms.

The exclusion criteria were determined as follows: (a) patients with severe liver, kidney, and other organ failure diseases and endocrine diseases; patients who had taken drugs that may affect the research indicators within one month before surgery; patients who did not cooperate or comply during treatment and who did not cooperate with follow-up after surgery; patients who also received other treatment methods during treatment; and patients whose condition suddenly worsens or was transferred to hospital during treatment.

2.3. Treatment Methods. All patients were intubated in a supine position under general anesthesia. Then, the laparoscope was placed, and the specific operations were as follows. Firstly, 1 cm below the belly button was undertaken as the entrance of the laparoscope to make a transverse cut here. The incision was about 2~3 cm. Then, the abdomen was inflated and the pressure was maintained at 12~14 mmHg to separate the organs from the abdominal tissues and fully expose them. Under the above circumstances, the size and position of the gallbladder in the abdominal cavity and its adhesion to the surrounding tissues were observed. All patients were treated by laparoscopy combined with duodenoscope. The duodenoscope was inserted from the oral cavity, the duodenal papilla was found according to the duodenoscope, and the guide wire was inserted into the common bile duct along the nipple. During the surgery, patients in the conventional group were performed with endoscopy for retrograde cholangiopancreatography, while patients in the AI group received the visual sensor and the endoscope to directly observe the gallbladder and bile duct. The location and size of the stones in the bile duct were

observed. Afterwards, the papillary sphincter was incised (when the incision was about 1 cm and the stone was less than 1 cm, the balloon was to directly remove the stone; when the stone was larger than 1 cm, it had to be crushed firstly before removing). After the stone removal, it was observed again. If it was observed that the stones in the common bile duct had been removed, a plastic stent of the bile duct was placed for nasal drainage, and the patient's liver function and amylase indicators were monitored synchronously. After 3 days, laparoscopic cholecystectomy was performed for surgical treatment. The specific process is shown in Figure 4. Two weeks after the patient was discharged from the hospital, the biliary stent was removed by a duodenoscope, and the stone removal was observed by the above two methods.

2.4. Evaluation Indicators. The blood loss, operation time, postoperative drainage tube placement time, gastrointestinal function recovery, extrabed activity time, postoperative complications, residual stones, and time spent in hospital for the two groups of patients were observed and recorded during the surgery.

The total serum bilirubin, direct bilirubin, alanine aminotransferase, and alkaline phosphatase were measured with the Romer combined biochemical analyzer (Roche Diagnostics (Shanghai) Co., Ltd. COBAS INTEGRA 800) and compared.

2.5. Statistical Analysis. In this study, SPSS 25.0 was applied to calculate the data. The count data was expressed in the form of percentage, and the χ^2 test was used. The measurement data was expressed in the form of $x \pm s$, and the t-test was used. $P < 0.05$ was considered that the difference was statistically significant.

3. Results

3.1. Image Segmentation Results. Figure 5 shows the captured images of the gallbladder stones and choledocholithiasis under laparoscopy and duodenoscope. Figures 5(a)–5(d) are the initial images under the endoscopy, and Figures 5(a1)–5(d1) are the images under the M-Unet segmentation. Figure 5(a) is a mirror image of the complete gallbladder of the abdominal cavity. It showed that the diameter of the gallbladder was slightly increased, but there was no inflammatory response on the surface. Figure 5(a1) shows the shape of the gallbladder segmented by M-Unet in the original image which performed. Figure 5(b) is an image of the bile duct after the surface mucosal tissue had been removed under laparoscopic surgery, and the shape of the bile duct can be clearly observed where the yellow arrow pointed. Figure 5(b1) shows the morphological performance of the bile duct after M-Unet segmentation. Figure 5(c) shows the state of stones in the bile duct observed under the duodenum meridian, as pointed by the yellow arrow. It can be concluded that the surface of the stones was smooth, and the overall colour was black. Compared with the diameter of the bile duct, the diameter of the stones was relatively small. Figure 5(c1) shows the mor-

phological performance of bile duct stones obtained by M-Unet segmentation. Observation suggested that only the segmentation was the part shown by the plain film. Figure 5(d) shows a stone blocked in the bile duct orifice under a duodenoscope (pointed by the yellow arrow). It can be observed that the bile duct orifice had been cracked due to the blockage of the stone, and inflammation was visible. It also indirectly showed that the diameter of the stones in the figure was larger than the diameter of the bile duct, and the stones had to be crushed before taking them. Figure 5(d1) shows the stone part of the bile duct orifice segmented by M-Unet. After comparative analysis, the segmentation was relatively complete, and there was basically no missing segmentation. After calculation, the precision, recall, and the mean intersection ratio of the M-Unet-based segmentation algorithm were 94.56%, 96.56%, and 98.92%, respectively. The results were quite good and can be used in this study.

3.2. Comparison on General Data of Patients. According to the research statistics, Figure 6 shows the general treatment status of the two groups of patients in this study, including gender distribution, age distribution, average course of disease, and distribution of the number of stones. In terms of gender distribution, the proportion of men in the AI group was 48.72% (38/78) and the proportion of women was 51.28% (40/78); the proportions of men and women in the conventional group were 46.36% (51/110) and 53.64% (59/110), respectively (Figure 6(a)). In terms of age, the patients were divided into five stages: 25–35, 36–45, 46–55, 56–65, and >66 years old. The proportions of patients in five stages in the AI group were 8.97% (7/78), 19.23% (15/78), 29.49% (23/78), 15.38% (12/78), and 26.92% (21/78), respectively; while those in the conventional group were 7.27% (8/110), 21.82% (24/110), 24.55% (27/110), 18.18% (20/110), and 28.18% (31/110), respectively (Figure 6(b)). The average course of disease for patients in the AI group was 3.37 ± 0.78 months, and that in the conventional group was 3.44 ± 0.12 months (Figure 6(c)). In terms of the distribution of the number of stones, the proportion of patients with single stones in the AI group was 21.79% (17/78), the proportion of multiple stones was 78.21% (61/78); the proportions of patients with single stones and multiple stones in the conventional group accounted for 23.64% (26/110) and 76.36% (84/110), respectively (Figure 6(d)). After comparative analysis, the two groups were basically the same in gender distribution, age distribution, course of disease, and distribution of the number of stones, so there was no obvious statistical difference ($P > 0.05$).

3.3. Comparison on Observation Indicators for Patients in Two Groups. The statistics results of some of observation indicators for the two groups of patients are given in Table 1, including operation time, postoperative drainage tube placement time, time required for recovery of gastrointestinal function, time required to get out of bed, and time spent in hospital. The observation and comparison revealed that the operation time (2.74 ± 0.45 hours), postoperative drainage tube placement time (4.31 ± 1.15 days), time required for recovery of gastrointestinal function
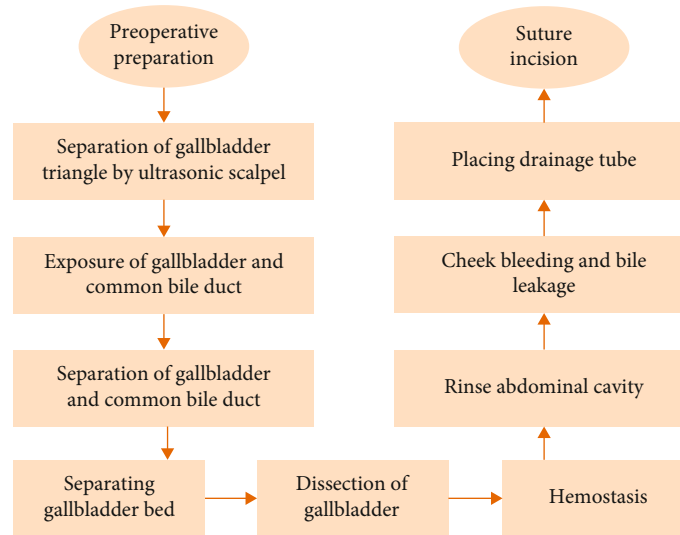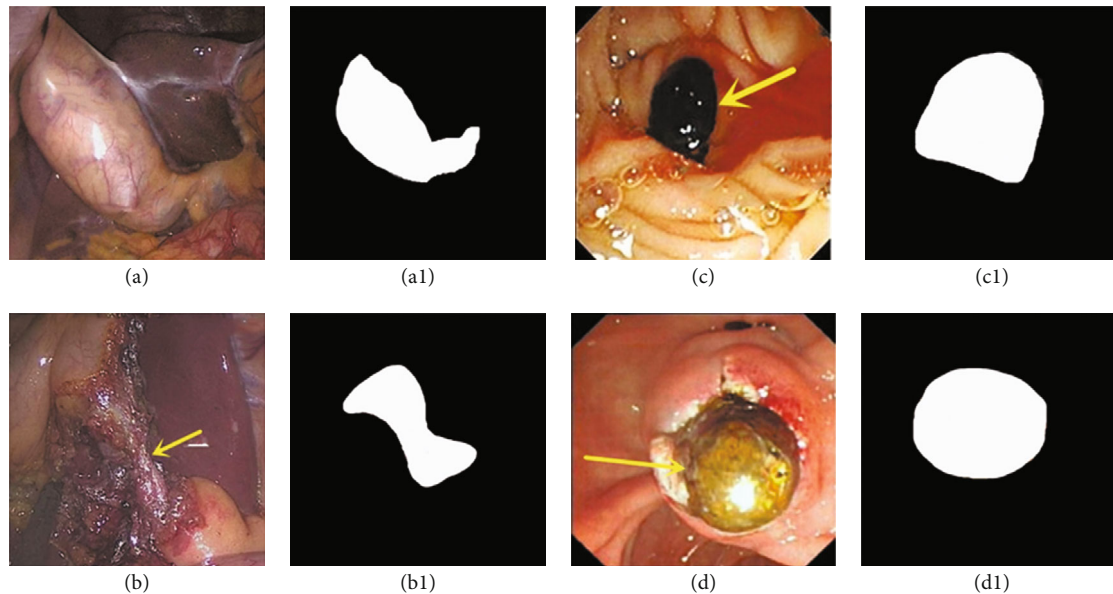
FIGURE 4: Laparoscopic cholecystectomy process.

FIGURE 5: The captured images of the gallbladder stones and choledocholithiasis under laparoscopy and duodenoscope. Note: (a) and (a1) show the images of the gallbladder; (b) and (b1) show the images of the bile duct; (c) and (d) show the images of the common bile duct; and (c1) and (d1) show the stone in the common bile duct. The yellow arrows on images (b), (c), and (d) referred to the cystic duct, stones in the common bile duct, and stones in the common bile duct, respectively.

(1.74 ± 0.54 days), time required to get out of bed (1.14 ± 0.55 hours), and time spent in hospital (9.94 ± 1.45 days) for patients in the AI group were shorter than those in the conventional group, showing obvious statistical differences ($P < 0.05$). Figure 7 shows the intraoperative blood loss and postoperative stone residual rate of the two groups of patients. Figure 7(a) shows the intraoperative blood loss of the two groups of patients. The intraoperative blood loss for patients in the AI group was 79.74 ± 6.45 mL, and that in the conventional group was 150.67 ± 15.71 mL. Figure 7(b) shows the residual status of stones in the two groups of patients. The number of patients with residual stones in the AI group was 0, and there were 12 patients with residual stones in the conventional group, accounting for 10.91%. After comparative analysis, it can be found that the intraoperative blood loss and postoperative stone residual rate of patients in the AI group were lower compared with those of the conventional group, and the differences were statistically notable ($P < 0.05$).

3.4. Comparison on Postoperative Complications. Based on the statistics of the results of this study, the postoperative complications of the two groups of patients were mostly biliary fistula, biliary hemorrhage, bile reflux esophagitis, acute pancreatitis, and incision infection (Figure 8). In the AI group, the numbers of patients with biliary fistula, biliary
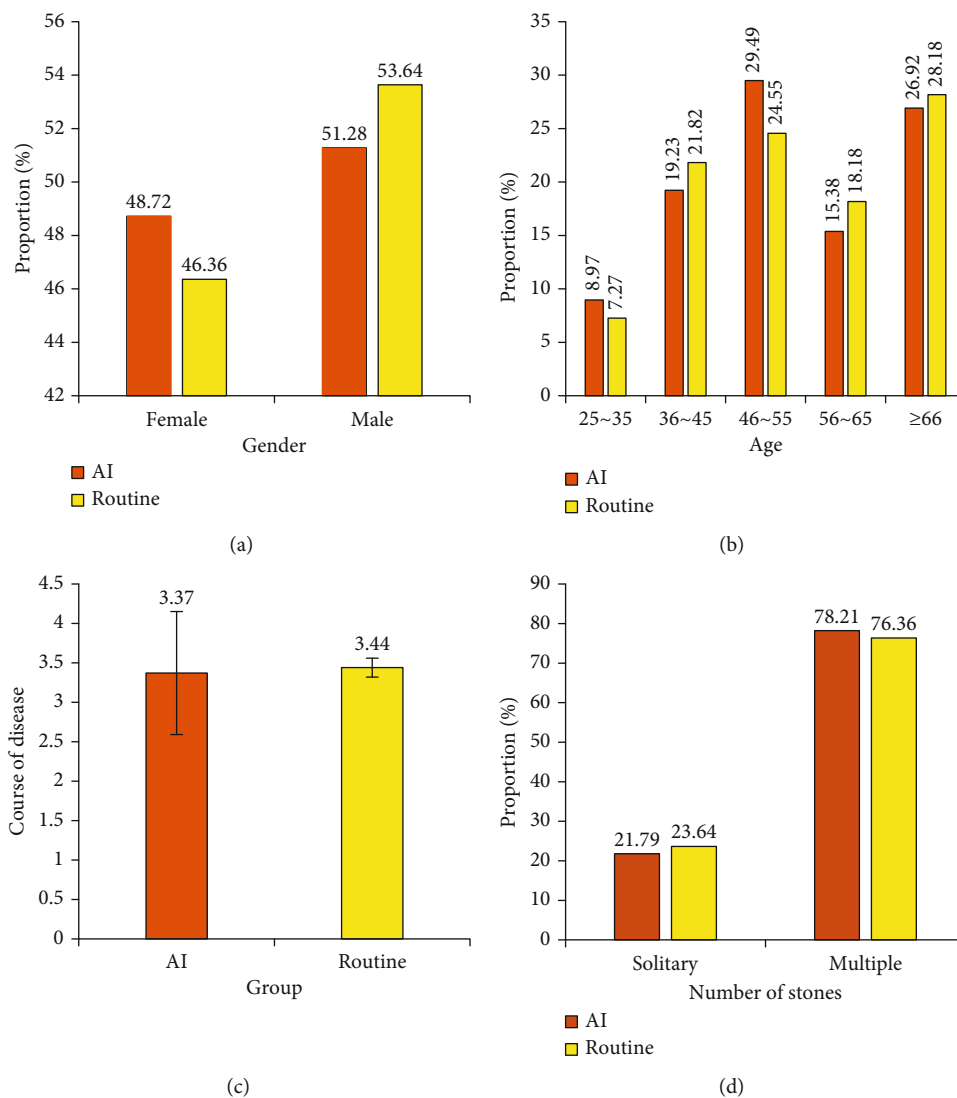
(a)

(b)

(c)

(d)

FIGURE 6: Comparison on general data of patients in two group. Note: (a)–(d) illustrate the comparison on gender, age, course of disease, and the number of stones, respectively.

TABLE 1: Comparison on observation indicators for patients in the two groups.

Observation indicators	AI group ($n = 78$)	Conventional group ($n = 110$)
Operation time (h)	2.74 ± 0.45	3.21 ± 0.32
Postoperative drainage tube placement time (d)	4.31 ± 1.15	12.14 ± 2.98
Time required for recovery of gastrointestinal function (d)	1.74 ± 0.54	2.89 ± 0.67
Time required to get out of bed (h)	1.14 ± 0.55	2.09 ± 0.87
Time spent in hospital (d)	9.94 ± 1.45	14.14 ± 1.15

hemorrhage, bile reflux esophagitis, acute pancreatitis, and incision infections were 0, 2, 1, 5, and 2, respectively. The total incidence of complications in the AI group was 10.26%. The numbers of patients with biliary fistula, biliary hemorrhage, bile reflux esophagitis, acute pancreatitis, and incision infections were 2, 4, 5, 9, and 6, respectively. After statistics and calculation analysis, the total incidence of complications in conventional group patients was 23.64%.

Therefore, it was clear that the incidence of complications in the AI group was lower obviously than that in the conventional group, showing statistical difference ($P < 0.05$).

3.5. Comparison on Postoperative Gallbladder Function. Figure 9 shows the comparison on average values of serum total bilirubin, direct bilirubin, alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, and
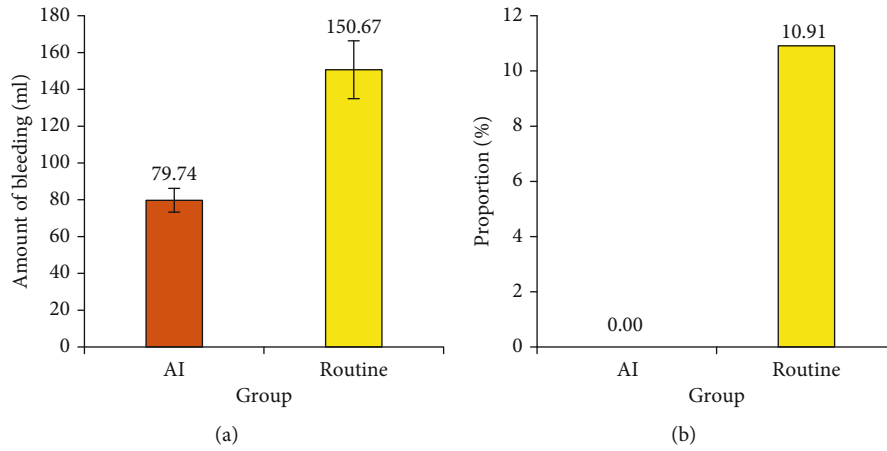
FIGURE 7: Comparison on intraoperative blood loss and postoperative stone residual rate of patients. Note: (a) shows the comparison on intraoperative blood loss; (b) shows the comparison on postoperative stone residual rate.
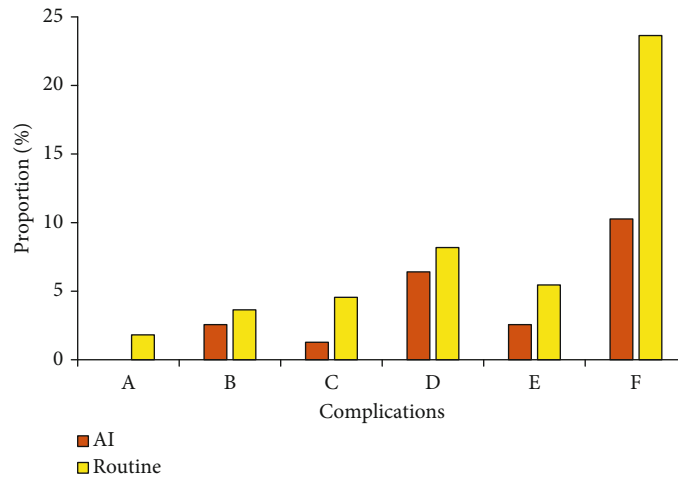

FIGURE 8: Comparison on postoperative complications. Note: A~F in the figure referred to biliary fistula, biliary hemorrhage, bile reflux esophagitis, acute pancreatitis, incision infection, and total incidence, respectively.
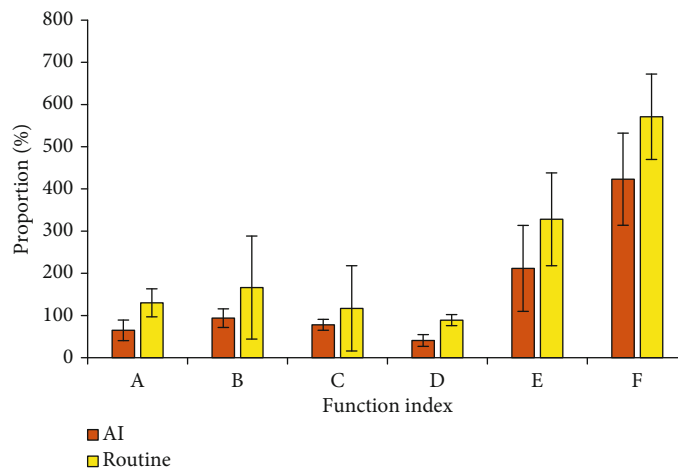

FIGURE 9: Comparison on postoperative gallbladder function indicators. Note: A~F in the figure referred to serum total bilirubin, direct bilirubin, alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, and glutamyl transpeptidase, respectively.

glutamyl transpeptidase in the two groups of patients. After analysis, it was found that the serum total bilirubin (64.99 ± 24.45 μmol/L), direct bilirubin (93.79 ± 22.11 μmol/L), alanine aminotransferase (77.99 ± 13.01 U/L), aspartate aminotransferase (40.98 ± 13.9 U/L), alkaline phosphatase (211.77 ± 101.91 U/L), and glutamyl transpeptidase (423.12 ± 109.21 U/L) were lower remarkably than those in the conventional group, which were 30.13 ± 33.25 μmol/L, 166.11 ± 122.09 μmol/L, 116.89 ± 100.99 U/L, 89.08 ± 13.01 U/L, 328.09 ± 110.01 U/L, and 570.99 ± 101.21 U/L, respectively. The differences in above indicators between patients in the two groups were statistically great ($P < 0.05$).

4. Discussion

Cholelithiasis refers to the stone-like disease that occurs in the gallbladder and bile duct. It is not only a disease with a high incidence and a high incidence in China but also has a high incidence in the world. According to relevant statistical reports, about 10%~15% of European adult population suffer from gallstone disease, and the incidence of gallbladder stones combined with choledocholithiasis in patients with gallstone disease is 10%~20% [21, 22]. There are also related statistics suggesting that in the United States, there are about 2,000 new cases of gallstones appearing every year. Studies have shown that 18%~33% of symptomatic cholelithiasis is related to acute biliary pancreatitis [23]. The surgical treatment of biliary stones has undergone many improvements and is still being improved, mainly to provide patients with more efficient and convenient treatment methods. The surgical treatment of gallbladder stones has evolved from traditional open surgery to today's minimally invasive surgical treatment under the guidance of endoscopy. Such series of changes have greatly reduced the pain of the patient during the treatment process and can further reduce the time required for the prognosis of the patients.

In this study, the visual sensing technology based on the CNN system was applied to further optimize the surgical treatment of cholelithiasis guided by endoscopy and analyze its application effects through clinical experiments. Firstly, the improved CNN algorithm was adopted to optimize the visual sensing technology, which was tested and verified then. The results showed that the precision of the M-Unet-based segmentation algorithm reached 94.56%, the recall reached 96.56%, and the mean intersection ratio was 98.92%. Such results are basically consistent with the trend of a study on the multitarget detection and recognition system in the intelligent visual sensor network (in which the precision, accuracy, and mean intersection ratio based on the segmentation under intelligent algorithms were 96.63%, 56.92%, and 93.32%, respectively) [24]. A study on the application effects and functions of industrial visual perception technology in smart cities found that the improved CNN algorithm was superior to other algorithms in image and video processing [25]. Secondly, this technology was combined with endoscopic technology and applied to the surgical treatment of patients with gallbladder stones combined with choledocholithiasis. It was found in this study that in the AI group, the operation time (2.74 ± 0.45 h), postoperative drainage tube placement time (4.31 ± 1.15 d), time required for recovery of gastrointestinal function (1.74 ± 0.54 d), time required to get out of bed (1.14 ± 0.55 h), and time spent in hospital (9.94 ± 1.45 d) were all shorter compared with those in the conventional group, which were 3.21 ± 0.32 h, 12.14 ± 2.98 d, 2.89 ± 0.67 d, 2.09 ± 0.87 h, and 14.14 ± 1.15 h, showing statistical differences ($P < 0.05$); the intraoperative blood loss (79.74 ± 6.45 mL) and residual status of stones (0%) in the AI group were much lower than those in the conventional group ($P < 0.05$); in addition, the incidence of complications (10.26%) during the treatment was lower than that (23.64%) in the conventional group ($P < 0.05$). What is more, the therapeutic effect of the biliary function index of the AI group was also much better than that of the conventional group. The above results are consistent with most other complication rates after treatment with laparoscopy combined with duodenoscope for treatment of gallbladder stones combined with common bile duct stones [4, 5, 26–28], providing support for laparoscopy combined with duodenoscope in the treatment of gallstones. Of course, AI-assisted visual sensing technology shows a wide range of applications, not just for assisted endoscopic treatment of gallbladder stones, which has been illustrated in a number of studies on the optimization of colonoscopy bowel preparation and the establishment and preliminary verification of the artificial intelligence-assisted colorectal polyp identification system [29–32]. AI-assisted visual sensing technology can also be used to prevent and control children's myopia. A systematic study on the use of multipurpose 3D visual sensing technology to prevent and control children's myopia illustrates this research [33]. AI-assisted visual sensing technology can also be applied to the identification and detection of drugs, which is specifically embodied in a research on intelligent and visual sensing technology and the application of this technology in the automatic visual detection of drugs [34]. Both the research results of this study and the related research results of predecessors reflect that it is natural and inevitable that intelligent and visual sensing technology is popularized in all aspects of people's lives.

5. Conclusions

In this study, the visual sensing technology based on the CNN system was applied to further optimize the treatment of cholelithiasis guided by endoscopy and analyze its application effects through clinical experiments. The conclusions were summarized as follows. Firstly, CNN algorithm showed good effects in image segmentation processing. Secondly, based on the CNN algorithm, AI-assisted visual sensing technology can reasonably process the target image and has a good optimization effect. Thirdly, applying the visual sensing technology based on CNN algorithm to endoscopic technology for the treatment of gallbladder stones combined with choledocholithiasis can effectively improve the treatment effect. However, the representativeness of the research results was not up to standard due to the insufficient number of samples taken in this study, so further research was

needed. This also gave certain empirical hints, which would be continuously summarized and corrected in subsequent research. However, it was proved in this study that visual sensing technology based on intelligent algorithms showed a good future in the medical field.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] A. di Ciaula, D. Q. Wang, and P. Portincasa, "An update on the pathogenesis of cholesterol gallstone disease," *Current Opinion in Gastroenterology*, vol. 34, no. 2, pp. 71–80, 2018.

[2] D. M. Shabanzadeh, L. T. Sørensen, and T. Jørgensen, "Determinants for gallstone formation - a new data cohort study and a systematic review with meta-analysis," *Scandinavian Journal of Gastroenterology*, vol. 51, no. 10, pp. 1239–1248, 2016.

[3] F. Lv, S. Zhang, M. Ji, Y. Wang, P. Li, and W. Han, "Single-stage management with combined tri-endoscopic approach for concomitant cholecystolithiasis and choledocholithiasis," *Surgical Endoscopy*, vol. 30, no. 12, pp. 5615–5620, 2016.

[4] L. Hu, Y. Chai, X. Yang, Z. Wu, H. Sun, and Z. Wang, "Duodenoscope combined with laparoscopy in treatment of biliary stones for a patient with situs inversus totalis: a case report," *Medicine (Baltimore)*, vol. 98, no. 7, article 14272, 2019.

[5] X. Jiang, G. Yang, K. Wang, W. Bi, D. Shang, and G. Zhang, "Clinical efficacy analysis of the combination of the laparoscope and preoperative or intraoperative duodenoscope in the treatment of cholecystolithiasis with choledocholithiasis: a retrospective study," *Journal of Laparoendoscopic & Advanced Surgical Techniques. Part A*, vol. 29, no. 12, pp. 1539–1543, 2019.

[6] Y. Lyu, Y. Cheng, T. Li, B. Cheng, and X. Jin, "Laparoscopic common bile duct exploration plus cholecystectomy versus endoscopic retrograde cholangiopancreatography plus laparoscopic cholecystectomy for cholecystocholedocholithiasis: a meta-analysis," *Surgical Endoscopy*, vol. 33, no. 10, pp. 3275–3286, 2019.

[7] L. Pan, M. Chen, L. Ji et al., "The safety and efficacy of laparoscopic common bile duct exploration combined with cholecystectomy for the management of cholecystocholedocholithiasis: an up-to-date meta-analysis," *Annals of Surgery*, vol. 268, no. 2, pp. 247–253, 2018.

[8] M. Sága, V. Bulej, N. Čuboňova, I. Kuric, I. Virgala, and M. Eberth, "Case study: performance analysis and development of robotized screwing application with integrated vision sensing system for automotive industry," *International Journal of Advanced Robotic Systems*, vol. 17, no. 3, p. 2, 2020.

[9] J. S. Chou and C. H. Liu, "Automated sensing system for real-time recognition of trucks in river dredging areas using computer vision and convolutional deep learning," *Sensors (Basel)*, vol. 21, no. 2, p. 555, 2021.

[10] Z. Lv, L. Qiao, A. K. Singh, and Q. Wang, "Fine-grained visual computing based on deep learning," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 17, no. 1s, pp. 1–19, 2021.

[11] Z. Lv, B. Hu, and H. Lv, "Infrastructure monitoring and operation for smart cities based on IoT system," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 1957–1962, 2020.

[12] T. Kattenborn, J. Eichel, and F. E. Fassnacht, "Convolutional neural networks enable efficient, accurate and fine-grained segmentation of plant species and communities from high-resolution UAV imagery," *Scientific reports*, vol. 9, no. 1, article 17656, 2019.

[13] S. Wang, J. Zhou, T. Lei et al., "Estimating land surface temperature from satellite passive microwave observations with the traditional neural network, deep belief network, and convolutional neural network," *Remote Sensing*, vol. 12, no. 17, pp. 2691–2691, 2020.

[14] R. B. Hegde, K. Prasad, H. Hebbar, and B. M. K. Singh, "Feature extraction using traditional image processing and convolutional neural network methods to classify white blood cells: a study," *Australasian Physical & Engineering Sciences in Medicine*, vol. 42, no. 2, pp. 627–638, 2019.

[15] E. Kumar, "An efficient image classification of malaria parasite using convolutional neural network and ADAM optimizer," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 2, pp. 3376–3384, 2021.

[16] K. Nishimura, J. Hayashida, C. Wang, D. F. E. Ker, and R. Bise, "Weakly-supervised cell tracking via backward-and-forward propagation," in *Computer Vision – ECCV 2020: 16th European Conference*, vol. 12357, pp. 104–121, Glasgow, UK, 2020.

[17] Y. Chen, S. Hu, H. Mao, W. Deng, and X. Gao, "Application of the best evacuation model of deep learning in the design of public structures," *Image and Vision Computing*, vol. 102, article 103975, 2020.

[18] A. A. Lestari and A. Munir, "Climate forecasting uses backpropagation algorithm artificial neural network model for agricultural planning in Gowa regency," in *IOP Conference Series: Earth and Environmental Science*, vol. 1no. 473, pp. 12–124, South Sulawesi, Indonesia, 2020.

[19] S. Kuanar, D. Mahapatra, M. Bilas, and K. R. Rao, "Multi-path dilated convolution network for haze and glow removal in nighttime images," *The Visual Computer*, no. 3, pp. 1–14, 2021.

[20] Y. Wu, L. Lin, J. Wang, and S. Wu, "Application of semantic segmentation based on convolutional neural network in medical images," *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*, vol. 3, no. 37, pp. 533–540, 2020.

[21] the SAGES Guidelines Committee, V. K. Narula, E. C. Fung, D. W. Overby, W. Richardson, and D. Stefanidis, "Clinical spotlight review for the management of choledocholithiasis," *Surgical Endoscopy*, vol. 34, no. 4, pp. 1482–1491, 2020.

[22] F. C. Schacher, S. M. Giongo, F. J. P. Teixeira, and Â. Z. Mattos, "Endoscopic retrograde cholangiopancreatography versus surgery for choledocholithiasis - a meta-analysis," *Annals of Hepatology*, vol. 18, no. 4, pp. 595–600, 2019.

[23] J. Kaneko, K. Kawata, S. Watanabe et al., "Clinical characteristics and risk factors for stent-stone complex formation following biliary plastic stent placement in patients with common bile duct stones," *Journal of Hepato-Biliary-Pancreatic Sciences*, vol. 25, no. 10, pp. 448–454, 2018.

[24] S. A. Roseline, S. Geetha, S. Kadry, and Y. Nam, "Intelligent vision-based malware detection and classification using deep

random forest paradigm," *IEEE Access*, vol. 8, no. 8, pp. 206303–206324, 2020.

[25] Z. Lv and D. Chen, "Industrial visual perception technology in smart city," *Image and Vision Computing*, vol. 105, article 104070, 2021.

[26] M. Gonzalez-Urquijo, A. A. Baca-Arzaga, E. Flores-Villalba, and M. Rodarte-Shade, "Laparoscopy-assisted transgastric endoscopic retrograde cholangiopancreatography for choledocholithiasis after Roux-en-Y gastric bypass: a case report," *Annals of Medicine and Surgery (Lond)*, vol. 44, pp. 46–50, 2019.

[27] A. M. da Ponte-Neto, W. M. Bernardo, L. M. Coutinho et al., "Comparison between enteroscopy-based and laparoscopy-assisted ERCP for accessing the biliary tree in patients with Roux-en-Y gastric bypass: systematic review and meta-analysis," *Obesity surgery*, vol. 28, no. 12, pp. 4064–4076, 2018.

[28] L. Zheng, S. Huang, F. Liu, and J. Yang, "Clinical efficacy of duodenoscopy combined with laparoscopy in the treatment of patients with severe acute pancreatitis and pancreatic pseudocyst, and the effects on IL-6 and CRP," *Experimental and Therapeutic Medicine*, vol. 21, no. 1, p. 55, 2021.

[29] M. F. Byrne, N. Chapados, F. Soudan et al., "Su1614 artificial intelligence (AI) in endoscopy–deep learning for optical biopsy of colorectal polyps in real-time on unaltered endoscopic videos," *Gastrointestinal Endoscopy*, vol. 85, no. 5, pp. -AB364–AB365, 2017.

[30] U. Deding, J. Herp, A. L. Havshoei et al., "Colon capsule endoscopy versus Ct colonography after incomplete colonoscopy. application of artificial intelligence algorithms to identify complete colonic investigations," *United European Gastroenterology Journal*, vol. 8, no. 7, pp. 782–789, 2020.

[31] M. F. Byrne, N. Chapados, F. Soudan et al., "Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model," *Gut*, vol. 68, no. 1, pp. 94–100, 2019.

[32] K. M. Konstantinidis, P. Hirides, S. Hirides, P. Chrysocheris, and M. Georgiou, "Cholecystectomy using a novel single-site (®) robotic platform: early experience from 45 consecutive cases," *Surgical Endoscopy*, vol. 26, no. 9, pp. 2687–2694, 2012.

[33] K. Sharma and M. Giannakos, "Sensing technologies and child-computer child-computer interaction: opportunities, challenges and ethical considerations," *International Journal of Child-Computer Interaction*, vol. 30, article 100331, 2021.

[34] L. Kelly, S. Sachan, L. Ni, F. Almaghrabi, R. Allmendinger, and Y.-W. Chen, "Explainable artificial intelligence for digital forensics: opportunities, challenges and a drug testing case study," *AJN, American Journal of Nursing*, vol. 9, no. 5, pp. 3–5, 2020.