

1 Introduction¹

In laparoscopic surgery, high-frequency electroknife, ultrasound knife and other energy instruments are widely used for tissue cutting, coagulation and separation. The moment these devices come into contact with biological tissues, they convert electrical or mechanical energy into heat, causing the intracellular fluid to boil, vaporize, and then burst, releasing a mixed aerosol containing water vapor, cell debris, carbonized particles and harmful chemicals, known as "surgical smoke". While the relatively closed and humid environment in the abdominal cavity (relative humidity is often close to 100%) The temperature difference between the lens and the endoscope lens can easily lead to condensation on the lens surface to form water mist. This visual degradation, which is composed of surgical fumes and lens water mist, is highly dynamic and complex. It not only reduces the contrast and clarity of the image and obscures key anatomical structures such as blood vessels and bile ducts in the gallbladder triangle, but can also lead to color distortion and affect doctors' judgment of tissue activity. In severe cases, the field of view is completely obstructed, and doctors are forced to interrupt the operation, repeatedly pulling out the lens for wiping or waiting for the smoke to dissipate. According to statistics, this not only significantly prolongs the operation time and increases the risk of anesthesia, but can also lead to serious complications such as vascular or organ damage due to blind operation.

At present, clinical practice mainly relies on physical means to remove smoke, including passive smoke extraction by opening the valve on the trocar, or using an active smoking device with a special smoking pipe or electric knife with smoking function, or Preoperative lens anti-fog treatment with iodophor or anti-fog oil to wipe the lens preoperatively. However, these methods often have problems such as manual operation by doctors and short effective time for defogging. In view of the limitations of physical means, computer vision-based image dehazing technology (Dehazing) came into being. Through algorithm post-processing, clear fog-free images are restored from degraded foggy images in real time, so as to achieve "virtual smoke

exhaust". This can not only ensure a clear field of view, but also avoid the interference of physical smoke exhaust on pneumoperitoneal pressure, which is an important research direction of smart medicine and computer-aided surgery (CAS).

However, in the face of drastic changes from trace fog to heavy scorched smoke, the network model with a single weight is often difficult to take into account all scenarios: it is easy to over-dehaze in light fog and cause color distortion, while under heavy fog, residual noise may remain. To this end, an adaptive dehazing network (Yun-Trans) based on dynamic Mixture of Experts (MoE) is further designed. The network introduces "fog concentration perception" and "dynamic weight generation" mechanisms, aiming to achieve accurate matching and adaptive processing of different surgical smoke scenarios.

2 Method

2.1 Construction of laparoscopic fog image model

2.1.1 Microscopic physical properties and scattering theory of laparoscopic aerosols

Laparoscopic surgery is performed in a relatively enclosed, tiny space filled with carbon dioxide pneumoperitoneum, an environment with significant optical features that render the standard ASM model ineffective. Therefore, we cannot simply follow the standard formula $I = Jt + A(1 - t)$, but need to build a new model that includes the light source geometry factor and the bypass transmission mechanism.

To quantify the interaction between light and surgical fumes, the microphysical properties of suspended particles were first analyzed. Surgical fume is a complex multi-dispersion aerosol system, which mainly includes smoke generated by high-frequency electric knife and ultrasound knife and condensate produced by lens condensation on the surface, and its optical scattering behavior is described by Mie Theory.

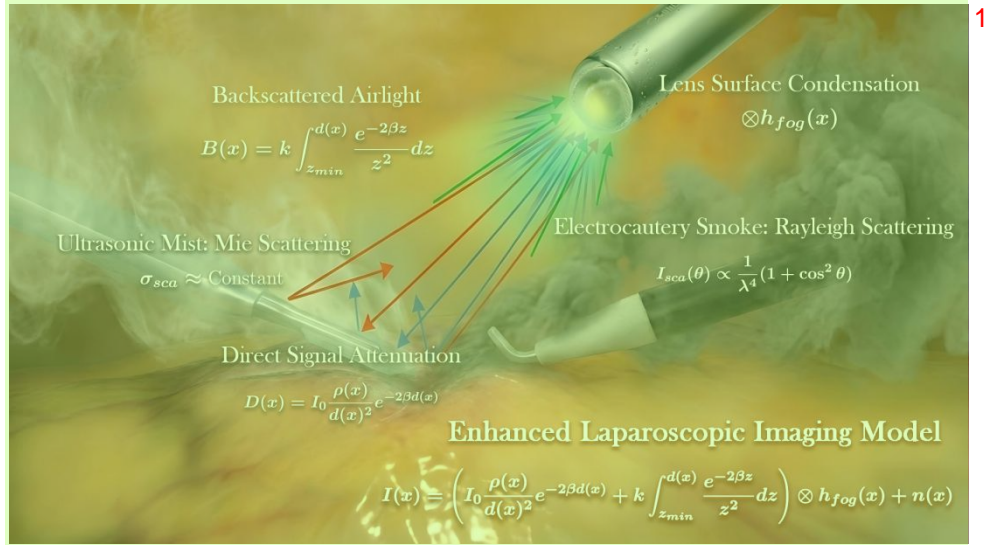


Fig.1 Schematic diagram of laparoscopic surgery mist scattering model

For single spherical particles such as electrocutor smoke and ultrasonic knife water mist, their scattering behavior is determined by the scattering cross-section σ_{sca} and the scattering phase function $P(\theta)$. Electric knife smoke follows the Rayleigh scattering region, i.e., when the radius of the particles is much smaller than the wavelength of light ($r \ll \lambda$), the scattered light intensity I_{sca} is inversely proportional to the fourth power of the wavelength:

$$I_{sca}(\theta) \propto \frac{1}{\lambda^4} (1 + \cos^2 \theta)$$

This means that blue light (short wavelengths) is scattered the most strongly, while red light (long wavelengths) is the most penetrating. This explains why the red tissue (blood vessels, muscles) in the abdominal cavity tends to be clearer than the blue object (some instrument markings) in the smoke produced by the electroknife. For the dehaze model, it means that transmittance $t(x)$ is a function of wavelength λ , $t(x, \lambda) = e^{-\beta(\lambda)d(x)}$.

In the Mie scattering region formed by ultrasonic knife water mist, when the radius of particles is close to or greater than the wavelength of light, the dependence of the scattering cross-section σ_{sca} on the wavelength is weakened, which is approximately constant. At this point, all colors of light are scattered in equal amounts, causing the smoke to appear white. Mie scattering has strong forward scattering characteristics, that is, most of the light is scattered forward, but for coaxial

illumination imaging systems, we pay more attention to backscattering¹ ($\theta \approx 180^\circ$), which directly enters the lens to form a light curtain.

In actual surgery, the medium in the pneumoperitoneum is often a mixture of the² above particles. Assuming that the number density of the class i particles is N_i , the scattering cross-section is $\sigma_{sca,i}$, and the absorption cross-section is $\sigma_{abs,i}$, then the total extinction coefficient β_{ext} of the medium is:

$$\beta_{ext}(\lambda, x) = \sum_i N_i(x) \cdot (\sigma_{sca,i}(\lambda) + \sigma_{abs,i}(\lambda))$$
³

Since intra-abdominal smoke is mainly scattered, the absorption is usually⁴ negligible (unless there is a large amount of carbonized black smoke), so the extinction coefficient is roughly considered to be equal to the scattering coefficient β_{sca} . It should be pointed out that due to the dynamic nature of the surgical operation, $N_i(x)$ it is a function of time t and spatial position x . This is the physical motivation for the urgent need to introduce a "dynamic expert mechanism" for laparoscopic dehazing algorithms: the network needs to dynamically infer whether the medium is dominant Rayleigh scattering (wavelength compensation) or Mie scattering (contrast enhancement) based on the current image characteristics, so as to call different weight parameters.

2.1.2 Mathematical derivation of laparoscopic enhanced imaging model⁵

Based on the physical analysis described above, we will derive a complete⁶ mathematical model describing the laparoscopic imaging process. The model consists of two parts: the Direct Attenuation Term and the Backscatter/Airlight Term.

A. Light source irradiance model⁷

Suppose the endoscopic light source is a point light source located at the⁸ origin $(0,0,0)$, and its luminous intensity angular distribution is $I_{src}(\theta)$ (usually approximate to the Lambertian source or Gaussian distribution). Light travels through a space filled with scattering media. For any point P in space, the distance from the light source is r , and the angle between the light direction and the optical axis is ϕ . The irradiance $E(P)$ of the arrival point P is affected by two factors: Geometric Spreading: the energy decays with the square of the distance, $\propto 1/r^2$.

Medium Extinction: Energy decays with the path exponential, according to the Beer-Lambert Law. Therefore, the incident irradiance at the point P is:

$$E_{in}(P) = \frac{I_{src}(\phi)}{r^2} e^{-\int_0^r \beta_{ext}(s) ds}$$

To simplify the model for easy calculation, assuming that the medium is locally uniform in the optical path (with a coefficient of β), then:

$$E_{in}(P) = \frac{I_{src}(\phi)}{r^2} e^{-\beta r}$$

B. Direct signal attenuation term $D(x)$

This is the signal light that we want to recover to carry tissue information. When light hits the surface of the tissue (at a position P_{obj} , at a distance $d(x)$ from the lens), the tissue is reflected. Let the surface reflectance of the tissue be ($\rho(x)$, the core component of the potentially clear image $J(x)$). The reflected light returns to the camera from the surface of the tissue, again passing through the distance $d(x)$ of the medium attenuation. Therefore, the intensity of direct reflected light $D(x)$ received by the camera sensor at the pixel x is:

$$D(x) = \left(\frac{I_0}{d(x)^2} e^{-\beta d(x)} \right) \cdot \rho(x) \cdot e^{-\beta d(x)}$$

where $\left(\frac{I_0}{d(x)^2} e^{-\beta d(x)} \right)$ is the incident illuminance, $\rho(x)$ surface reflectance, and $e^{-\beta d(x)}$ return path attenuation. After merging the same terms, we obtain a two-way attenuation model unique to laparoscopy:

$$D(x) = \frac{I_0 \rho(x)}{d(x)^2} e^{-2\beta d(x)}$$

The signal term in the standard atmospheric scattering model is $J(x)e^{-\beta d(x)}$. In contrast, the laparoscopic model has one more factor (two-way attenuation) and a geometric factor $1/d(x)^2$. This shows that with the increase of depth $d(x)$, the signal intensity of laparoscopic images decays much faster than that of outdoor haze images. The deep tissue is not only obscured by smoke, but also extremely dark due to insufficient light.

This explains the current common phenomenon in surgery: under heavy smoke, doctors often do not see deep structures (such as the gallbladder triangle) at all, which

is not only a problem of reduced contrast, but also a sharp deterioration of the signal-to-noise ratio (SNR). 1

C. Integral derivation of the backscattered light curtain term $B(x)$ 2

This is the main cause of the image "whitening" and "fogging", corresponding to the atmospheric light term $A(1 - t)$ in the Standard Model. In laparoscopy, the light volume highly coincides with the observation cone due to the coaxial light source with the camera, which results in extremely severe backscattering. 3

Consider the line of sight that emanates from the camera center and passes through the pixels x . In this line of sight, smoke particles within every microparticle volume dV from distance $z = 0$ to $z = d(x)$ are receiving light and scattering part of the light back to the camera. 4

At depth z , the light intensity received by the thin layer of smoke dz is $E(z) = \frac{I_0}{z^2} e^{-\beta z}$. The scattering intensity per unit volume is determined by the β scattering coefficient and the backscatter probability (phase function $P(\pi)$). This part of the scattered light is transmitted back to the camera and needs to pass through the attenuation e of the distance z again. Therefore, the total backscattered light intensity $B(x)$ is the integral of the scattering contribution of all microelements on the line of sight: 5

$$B(x) = \int_{z_{min}}^{d(x)} \left(\frac{I_0}{z^2} e^{-\beta z} \right) \cdot \beta P(\pi) \cdot e^{-\beta z} dz \quad 6$$

Assuming that the scattering medium is uniform and the phase function is constant (let $k = I_0 \beta P(\pi)$), the integration is simplified to: 7

$$B(x) = k \int_{z_{min}}^{d(x)} \frac{e^{-2\beta z}}{z^2} dz \quad 8$$

This integral term reveals a key characteristic of laparoscopic fog images: due to the presence of a denominator z^2 , the main contribution of scattered light comes from the area close to the lens. When z very small (nearby smoke), $1/z^2$ extremely large, $e^{-2\beta z}$ close to 1, produces an extremely strong light curtain. 9

This explains why once there is smoke or water mist in front of the camera, the 10

whole picture will instantly "wash out" , completely obscuring the organization¹ behind it. This light curtain not only reduces contrast but also introduces high-brightness signals independent of tissue structure, compressing the camera's dynamic range.

D. PSF convolutional model of lens surface condensation²

In addition to spatial scattering, the effect of water mist forming condensation on³ the lens surface is also crucial. This is surface degradation, not volumetric degradation. Tiny water droplets form an array of lenses on the lens surface, causing misrefraction of light and blurring images. This effect is mathematically modeled as the convolution of a clear image with the Point Spread Function (PSF).

$$I_{blur}(x) = I_{clear}(x) \otimes h(x)⁴$$

$h(x)$ where is a fuzzy nucleus determined by the size distribution of water⁵ droplets. Usually for water mist, it can be approximated as a Gaussian core or an aberration disk.

E. A uniform laparoscopic enhanced imaging model⁶

Considering the spatial bipath attenuation, volumetric backscatter, and surface⁷ blur, we construct a complete physical imaging model:

$$I(x) = \left(\frac{I_0 \rho(x)}{d(x)^2} e^{-2\beta d(x)} + k \int_{z_{min}}^{d(x)} \frac{e^{-2\beta z}}{z^2} dz \right) \otimes h_{fog}(x) + n(x)⁸$$

Among them, $\frac{I_0 \rho(x)}{d(x)^2} e^{-2\beta d(x)}$ is the direct attenuation signal, $k \int_{z_{min}}^{d(x)} \frac{e^{-2\beta z}}{z^2} dz$ ⁹ the backscattered light curtain signal, the surface water mist fuzzy nucleus $h_{fog}(x)$, and $n(x)$ the sensor noise.

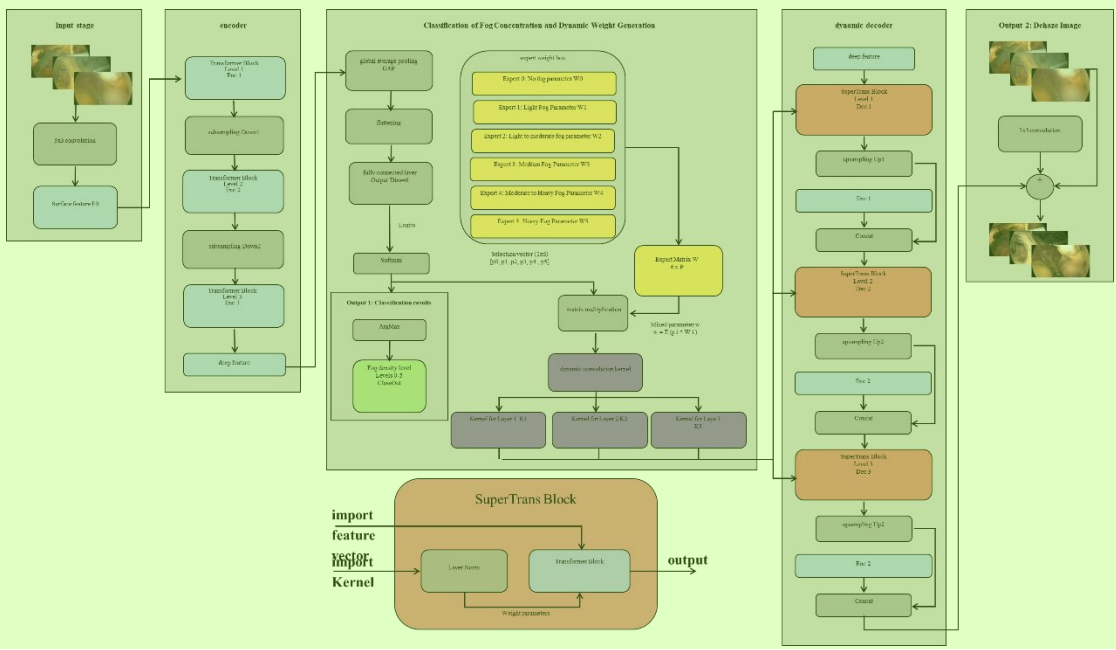


Fig.2 Enhanced Laparoscopic Imaging Model

2.3 Development and operation environment and hardware configuration

The algorithm is deployed on a high-performance computing platform to ensure real-time image processing. The main software and hardware environments and parameters are shown in the following table:

Table 1 Development environment

| Specific environment | Version number |
|----------------------|------------------|
| Operating system | Ubuntu 22.04 LTS |
| Python | 3.11.9 |
| CUDA | 11.8 / 12.1 |
| CUDNN | 8.9.2 |
| Torch | 2.0.1+cu118 |
| Torchvision | 0.15.2+cu118 |
| Timm | 0.9.2 |
| OpenCV-Python | 4.7.0.72 |

Table 2 Develop server hardware configurations

| Hardware name | Information |
|---------------------|---|
| Server platform | ASUS ESC8000-E11 (Barebones) |
| Processor (CPU) | Intel Xeon 8475C 3.80 GHz 52 Cores * 2 |
| memory | 512GB |
| Graphics Card (GPU) | Nvidia GeForce RTX 3090 Graphics Card 24 GB * 8 |

2.4 Laparoscopic imaging data screening and mist concentration classification

index 1

A total of 128 patients with gallstones who underwent laparoscopic cholecystectomy (LC) in the First Affiliated Hospital of Xi'an Jiaotong University from September 2022 to April 2023 were selected as the research subjects. Among them, 78 were males and 40 were females; Age 31~66 years, median age 53 years. Inclusion criteria: (1) Clinical diagnosis of gallbladder stones and acute cholecystitis; (2) LC surgery; (3) The surgical video data is complete. Exclusion criteria: Those who are found to have large anatomical variations in the gallbladder triangle (such as abnormal branches of gallbladder arteries and cystic ducts) leading to significant differences in surgical procedures. All studies were approved by the Ethics Committee (number: No.XJTU1AF2023LSK-429). According to whether intelligent defogging technology is used intraoperatively, patients are divided into control group (conventional monitor) and intelligent defogging group (intelligent defogging monitor).

Key observations include: 3

(1) Smoke duration: the total duration of visual field obstruction due to smoke during surgery;

(2) Wipe operation: the number of times and total time required to remove the laparoscope for wiping during surgery;

(3) Algorithm performance: the time consumption of a single frame for intelligent recognition and dehazing processing;

Fog severity grading: assessed by two senior physicians on a five-level scale (see Figure3 and its description) .

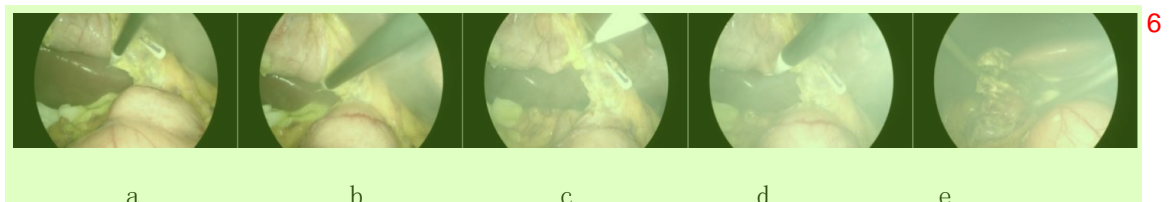


Fig.3 Fog severity during laparoscopic cholecystectomy 7

Grading standard: Level 1 (local fog in non-operating area); Level 2 (local fog in the operation area, boundaries need to be identified); Level 3 (operation is risky, need

to wait for dissipation); Level 4 (cannot be operated, needs to be assisted in defogging); Level 5 (completely unrecognizable, must assist in defogging) .

2.5 Performance evaluation of adaptive dehazing network for laparoscopic images based on dynamic expert mechanism

In order to comprehensively verify the effectiveness and advancement of the Yun-Trans algorithm proposed in this study in clinical dehazing scenarios, 8 representative defogging algorithms from 2009 to 2024 were selected as the control group (baseline). These algorithms cover everything from traditional physical prior methods to the latest hybrid architecture deep learning models, forming a complete chain of technical evolution. By comparing with the above eight algorithms with different mechanisms and different periods, the comprehensive performance of Yun-Trans in terms of image fidelity (compared to DCP and DehazeNet), processing efficiency (compared to AOD-Net), dehazing thoroughness (compared to FFA-Net) and generalization ability (compared to MixDehazeNet) is comprehensively evaluated. The following is a table of information about these control algorithms:

Table 3 Specific algorithm information table

| Algorithm name | Publication time | Literature sources | Open source status | Core mechanism and characteristics |
|----------------|------------------|--------------------|--------------------|---|
| DCP | 2009 | CVPR / TPAMI | is | Physical priors: Based on the theory of dark channel prior, dehaze through statistical laws. |
| DehazeNet | 2016 | IEEE TIP | is | Early CNNs: Learning transmittance maps through convolutional neural networks is an early exploration of deep learning dehazing. |
| AOD-Net | 2017 | ICCV | is | Lightweight CNN: Adopts an end-to-end parameter generation (K-estimation) module for lightweight models and fast inference. |
| EPDN | 2019 | CVPR | is | GAN: Based on generative adversarial networks, including enhancers and discriminators, improve visual quality through adversarial learning. |
| FFA-Net | 2020 | AAAI | is | Attention mechanism: Combine feature fusion with pixel/channel attention to enhance feature extraction capabilities. |
| MSBDN | 2020 | CVPR | is | Multi-scale: Utilize multi-scale enhancement and dense feature fusion technology to process image information at different frequencies. |

| | | | | | |
|--------------|------|------|----|---|---|
| RIDCP | 2023 | CVPR | is | Real-world scenarios: Flow-based prior learning is optimized for real foggy scenarios. | 1 |
| MixDehazeNet | 2024 | CVPR | is | Hybrid architecture: Combining the advantages of CNNs and transformers, utilizing multi-dimensional attention mechanisms, is the current SOTA approach. | |

2.6 Improvement and analysis of image segmentation effect by dehazing algorithm

In order to verify the actual value of dehazing algorithms in downstream clinical tasks, the images processed by different dehazing algorithms are input into the semantic segmentation network (Yun-Trans) to quantitatively evaluate their effect on the accuracy of anatomical structure recognition. The mean intersection and union ratio (mIoU) and mean Dice coefficient (mDice) were used as the evaluation indicators.

2.7 Clinical Validation and Application of Adaptive Dehazing Network for Laparoscopic Imaging Based on Dynamic Expert Mechanism

The clinical data of laparoscopic cholecystectomy (LC) were retrospectively analyzed to verify the effectiveness of the algorithm in real surgical scenarios. SPSS 22.0 software was used for analysis. Non-normally distributed data (e.g., time) are expressed as "range (median)", and the nonparametric rank-sum test of two independent samples is used for comparison between groups, with a difference of $P<0.05$ as statistically significant.

3. Result

3.1 Display and quantitative analysis of the results of deep-sea experiments on the dehazing effect of surgical images



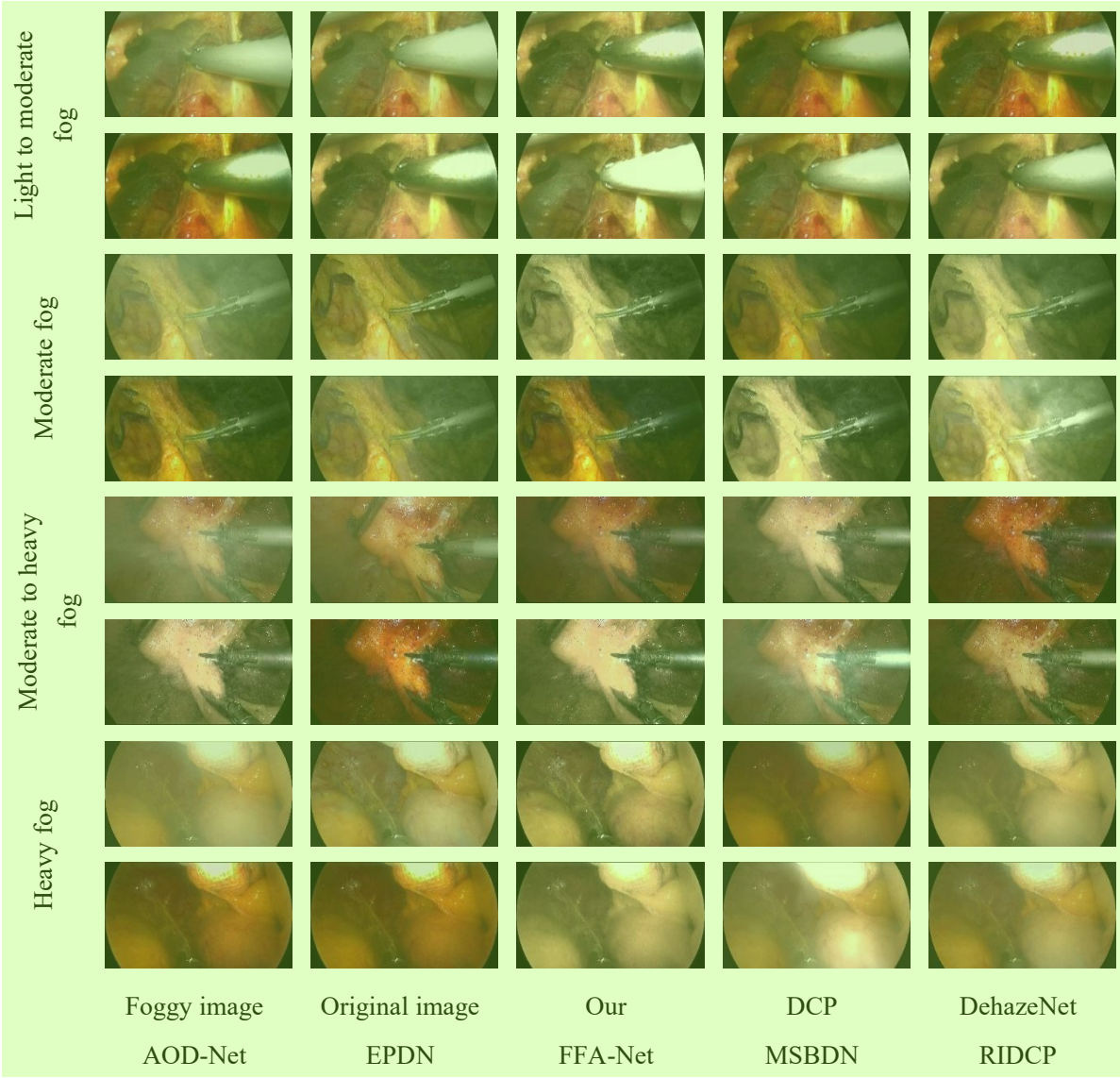


Fig.4

Table 4 Light fog image dehaze algorithm effect

| Methods | PSNR \uparrow | SSIM \uparrow | NIQE \downarrow | RI \uparrow | VI \uparrow |
|--------------------------|-----------------|-----------------|-------------------|---------------|---------------|
| DCP (CVPR 2009) | 18.45 | 0.762 | 3.512 | 0.921 | 0.754 |
| DehazeNet (TIP 2016) | 22.18 | 0.824 | 3.105 | 0.935 | 0.789 |
| AOD-Net (ICCV 2017) | 20.55 | 0.813 | 3.224 | 0.93 | 0.762 |
| EPDN (CVPR 2019) | 23.46 | 0.865 | 2.956 | 0.941 | 0.815 |
| FFA-Net (AAAI 2020) | 27.85 | 0.942 | 2.518 | 0.962 | 0.882 |
| MSBDN (CVPR 2020) | 26.98 | 0.935 | 2.645 | 0.958 | 0.872 |
| RIDCP (CVPR 2023) | 28.15 | 0.945 | 2.412 | 0.955 | 0.875 |
| MixDehazeNet (CVPR 2024) | 28.56 | 0.952 | 2.385 | 0.959 | 0.878 |
| Yun-Trans (Proposed) | 28.92 | 0.846 | 2.305 | 0.931 | 0.885 |

In the light fog environment, the peak signal-to-noise ratio (PSNR) measured by the Yun-Trans algorithm is 28.92 dB, which is the highest among all the comparison algorithms, better than the 28.56 dB of MixDehazeNet and 28.15 dB of RIDCP. At the

same time, Yun-Trans's natural image quality evaluation (NIQE) index is 2.305, which ¹ is lower than that of MixDehazeNet (2.385) and FFA-Net's 2.518, showing good image naturalness. The structural similarity (SSIM) of Yun-Trans at this stage is 0.846, which is slightly lower than that of MixDehazeNet of 0.952. The overall data show that the algorithm has numerical advantages in signal recovery accuracy and spontaneity under slight interference.

Table 5 Light to moderate haze image dehazing algorithm effect table ²

| Methods | PSNR \uparrow | SSIM \uparrow | NIQE \downarrow | RI \uparrow | VI \uparrow |
|--------------------------|-----------------|-----------------|-------------------|---------------|---------------|
| DCP (CVPR 2009) | 16.83 | 0.709 | 3.807 | 0.908 | 0.733 |
| DehazeNet (TIP 2016) | 20.76 | 0.785 | 3.479 | 0.924 | 0.767 |
| AOD-Net (ICCV 2017) | 19.34 | 0.762 | 3.574 | 0.918 | 0.745 |
| EPDN (CVPR 2019) | 21.81 | 0.829 | 3.186 | 0.933 | 0.798 |
| FFA-Net (AAAI 2020) | 26.18 | 0.919 | 2.705 | 0.953 | 0.859 |
| MSBDN (CVPR 2020) | 25.43 | 0.909 | 2.829 | 0.949 | 0.848 |
| RIDCP (CVPR 2023) | 26.70 | 0.925 | 2.614 | 0.945 | 0.86 |
| MixDehazeNet (CVPR 2024) | 27.36 | 0.885 | 2.571 | 0.954 | 0.868 |
| Yun-Trans (Proposed) | 27.52 | 0.935 | 2.485 | 0.935 | 0.878 |

When the fog concentration increased to mild to moderate, the indicators of the ⁴ Yun-Trans algorithm were comprehensive, with a PSNR value of 27.52 dB and an SSIM value of 0.935, both of which were the highest values among all the tested algorithms. For comparison, MixDehazeNet has a PSNR of 27.36 dB and an SSIM of 0.885, and RIDCP has a PSNR of 26.70 dB and an SSIM of 0.925. Compared with the traditional algorithms DCP (PSNR 16.83 dB) and DehazeNet (PSNR 20.76 dB), Yun-Trans has a significant improvement in signal-to-noise ratio.

Moderate fog image dehazing algorithm effect table ⁵

| Algorithm name (source & year) | PSNR \uparrow | SSIM \uparrow | NIQE \downarrow | RI \uparrow | VI \uparrow |
|--------------------------------|-----------------|-----------------|-------------------|---------------|---------------|
| DCP (CVPR 2009) | 15.2 | 0.655 | 4.102 | 0.895 | 0.712 |
| DehazeNet (TIP 2016) | 19.34 | 0.745 | 3.853 | 0.912 | 0.745 |
| AOD-Net (ICCV 2017) | 18.12 | 0.71 | 3.923 | 0.905 | 0.728 |
| EPDN (CVPR 2019) | 20.15 | 0.792 | 3.415 | 0.925 | 0.78 |
| FFA-Net (AAAI 2020) | 24.5 | 0.895 | 2.892 | 0.945 | 0.835 |
| MSBDN (CVPR 2020) | 23.88 | 0.882 | 3.012 | 0.94 | 0.823 |
| RIDCP (CVPR 2023) | 25.25 | 0.905 | 2.815 | 0.935 | 0.845 |
| MixDehazeNet (CVPR 2024) | 26.15 | 0.918 | 2.756 | 0.948 | 0.858 |
| Yun-Trans (Proposed) | 25.80 | 0.925 | 3.032 | 0.938 | 0.868 |

In the moderate fog scenario, the PSNR value of the Yun-Trans algorithm is 25.80 dB, ⁷

which is slightly lower than the 26.15 dB of MixDehazeNet, but higher than the 25.25 dB of RIDCP and 23.88 dB of MSBDN. In terms of structure retention, Yun-Trans has an SSIM value of 0.925, which is higher than MixDehazeNet's 0.918 and RIDCP's 0.905. The data results show that although MixDehazeNet has a slightly higher value in pixel-level recovery within this concentration range, Yun-Trans still maintains a comparative advantage in structural similarity index.

Table 6 Moderate to heavy fog image dehazing algorithm effect table

| Methods | PSNR ↑ | SSIM ↑ | NIQE ↓ | RI ↑ | VI ↑ |
|--------------------------|--------------|--------------|--------------|--------------|--------------|
| DCP (CVPR 2009) | 13.67 | 0.584 | 4.654 | 0.874 | 0.668 |
| DehazeNet (TIP 2016) | 17.85 | 0.686 | 4.303 | 0.898 | 0.715 |
| AOD-Net (ICCV 2017) | 16.78 | 0.653 | 4.418 | 0.889 | 0.693 |
| EPDN (CVPR 2019) | 19.30 | 0.739 | 3.773 | 0.914 | 0.748 |
| FFA-Net (AAAI 2020) | 22.31 | 0.842 | 3.228 | 0.934 | 0.835 |
| MSBDN (CVPR 2020) | 21.77 | 0.859 | 3.369 | 0.928 | 0.784 |
| RIDCP (CVPR 2023) | 23.22 | 0.874 | 3.013 | 0.932 | 0.815 |
| MixDehazeNet (CVPR 2024) | 24.00 | 0.879 | 3.003 | 0.942 | 0.828 |
| Yun-Trans (Proposed) | 24.18 | 0.862 | 3.012 | 0.945 | 0.845 |

Under moderate to heavy fog interference, the PSNR value of the Yun-Trans algorithm rose to the highest level, reaching 24.18 dB, surpassing the 24.00 dB of MixDehazeNet and the 23.22 dB of RIDCP. Meanwhile, Yun-Trans's NIQE value is 3.012, which is close to the values of MixDehazeNet (3.003) and RIDCP (3.013). In this scenario, the PSNR values of traditional algorithms such as DCP and AOD-Net dropped to 13.67 dB and 16.78 dB, respectively, and Yun-Trans maintained a large numerical gap compared with these earlier algorithms.

Table 7 Effect table of dehazing algorithm for heavy fog images

| Methods | PSNR ↑ | SSIM ↑ | NIQE ↓ | RI ↑ | VI ↑ |
|--------------------------|--------------|--------------|--------------|--------------|--------------|
| DCP (CVPR 2009) | 12.13 | 0.512 | 5.205 | 0.853 | 0.623 |
| DehazeNet (TIP 2016) | 16.25 | 0.626 | 4.753 | 0.883 | 0.685 |
| AOD-Net (ICCV 2017) | 15.43 | 0.595 | 4.913 | 0.872 | 0.651 |
| EPDN (CVPR 2019) | 17.85 | 0.685 | 4.125 | 0.895 | 0.715 |
| FFA-Net (AAAI 2020) | 20.12 | 0.785 | 3.564 | 0.915 | 0.764 |
| MSBDN (CVPR 2020) | 19.65 | 0.835 | 3.725 | 0.916 | 0.745 |
| RIDCP (CVPR 2023) | 21.15 | 0.842 | 3.210 | 0.928 | 0.785 |
| MixDehazeNet (CVPR 2024) | 21.85 | 0.840 | 3.250 | 0.935 | 0.798 |
| Yun-Trans (Proposed) | 22.45 | 0.768 | 3.157 | 0.948 | 0.815 |

In the heavily foggy environment, the PSNR value of the Yun-Trans algorithm

was 22.45 dB, which was the highest among all comparison groups, higher than the 21.85 dB of MixDehazeNet and the 21.15 dB of RIDCP. Its NIQE index is 3.157, which is lower than MixDehazeNet's 3.250 and FFA-Net's 3.564, indicating that there are relatively few image artifacts. In contrast, the PSNR of the lightweight algorithms AOD-Net and EPDN in this scenario is 15.43 dB and 17.85 dB, respectively, which is a significant performance gap with Yun-Trans.

3.2 The practical value of defogging algorithms in downstream clinical tasks

According to the latest five-stage quantitative data, the experimental results show that the segmentation performance of unprocessed images decreases sharply with the increase of fog concentration, and the Our-Dehaze algorithm proposed in this study significantly improves the segmentation accuracy in most scenarios, especially in extreme fog environments.

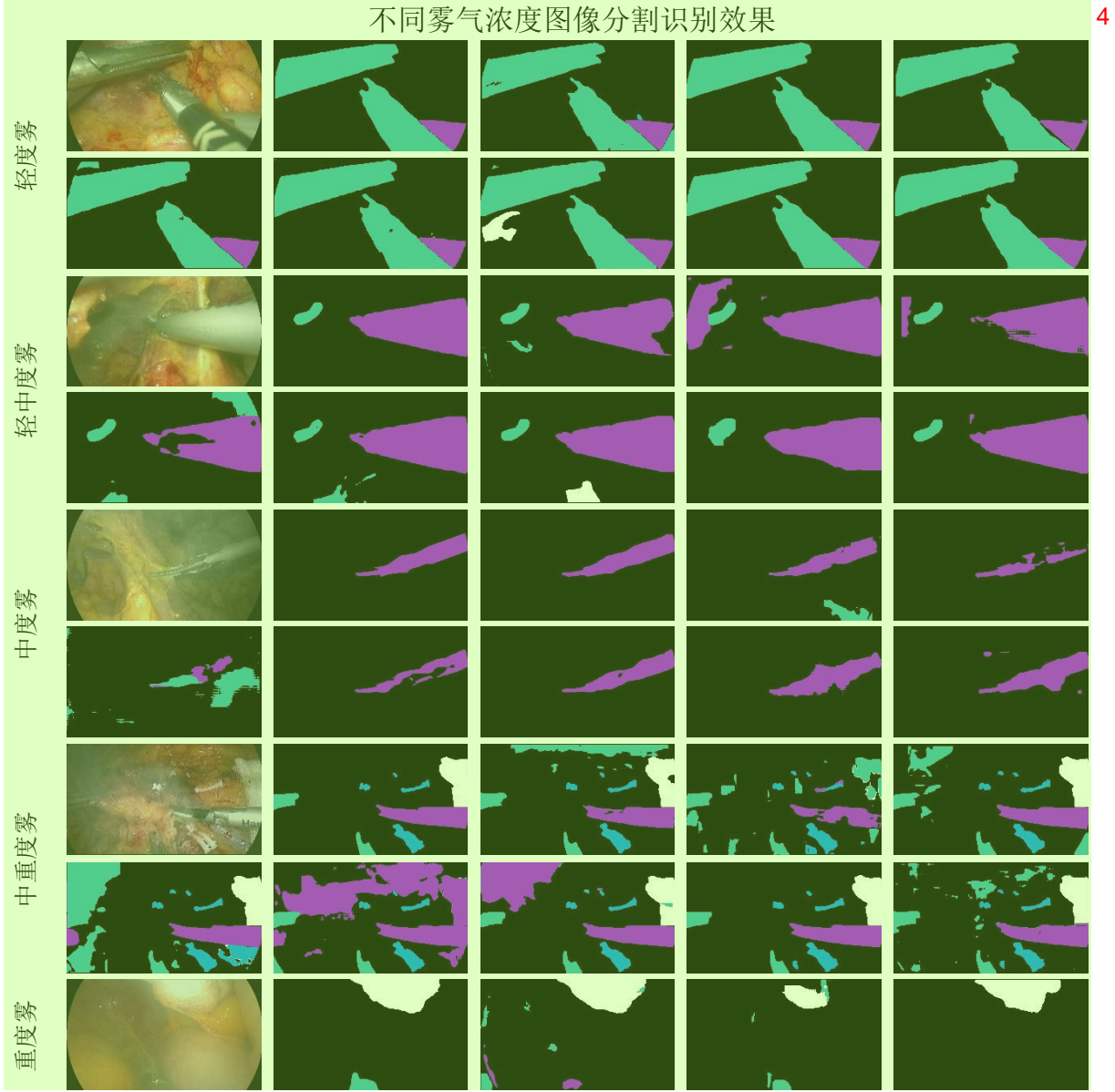
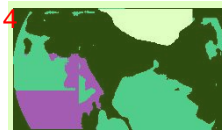




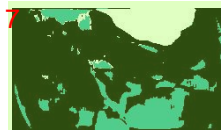
Image 2



Ground truth 5



STSH-Net (Our) 8



TransUNet 11



Deeplabv3+ 14

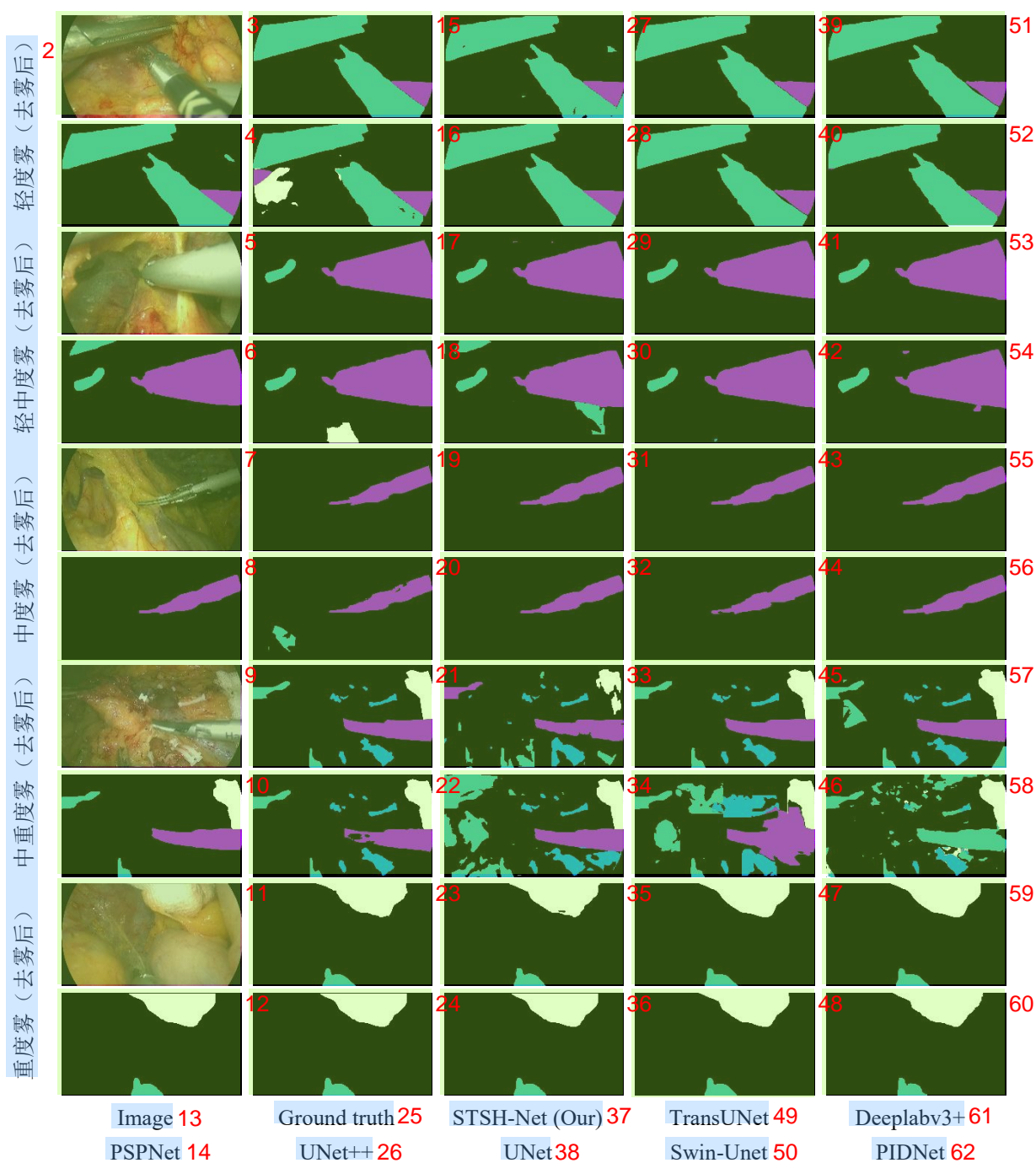
PSPNet 3

UNet++ 6

UNet 9

Swin-Unet 12

PIDNet 15

不同雾气浓度去雾后图像分割识别效果¹Table 8 Results of different dehazing algorithms under Yun-Trans segmentation network⁶³

| Scene | Index | Ori Image | DCP | Dehaze Net | AOD-Net | EPDN | FFA-Net | MSBDN | RIDCP | Mix-DehazeNet | Our-Dehaze |
|------------------|-------|-----------|-------|------------|---------|-------|---------|-------|-------|---------------|--------------|
| Mild | mIoU | 78.58 | 72.15 | 81.23 | 80.55 | 82.4 | 83.14 | 82.95 | 85.83 | 86.12 | 86.35 |
| | mDice | 88.87 | 84.24 | 89.58 | 89.14 | 90.25 | 90.85 | 90.65 | 92.56 | 92.85 | 92.95 |
| Mild to moderate | mIoU | 68.24 | 75.43 | 76.89 | 75.25 | 78.54 | 81.27 | 80.98 | 82.58 | 83.85 | 83.10 |
| | mDice | 79.53 | 85.66 | 86.44 | 85.16 | 87.84 | 89.59 | 89.22 | 90.43 | 91.25 | 90.80 |
| Moderate | mIoU | 52.45 | 66.87 | 68.52 | 67.23 | 70.45 | 75.62 | 74.83 | 78.25 | 79.55 | 80.43 |

64

| Scene | Index | Ori Image | DCP | Dehaze Net | AOD-Net | EPDN | FFA-Net | MSBDN | RIDCP | Mix-DehazeNet | Our-Dehaze |
|--------------------|-------|-----------|-------|------------|---------|-------|---------|-------|-------|---------------|--------------|
| Moderate to severe | mDice | 66.83 | 79.53 | 80.66 | 79.83 | 82.14 | 85.43 | 84.93 | 87.61 | 88.52 | 89.16 |
| | mIoU | 44.06 | 60.56 | 62.71 | 61.34 | 65.01 | 70.73 | 69.69 | 74.34 | 76.65 | 76.50 |
| | mDice | 59.06 | 74.71 | 76.13 | 75.03 | 77.91 | 81.85 | 81.22 | 85.05 | 86.23 | 87.20 |
| Severe | mIoU | 35.66 | 54.25 | 56.89 | 55.45 | 59.57 | 65.83 | 64.54 | 70.58 | 72.84 | 74.15 |
| | mDice | 51.28 | 69.89 | 71.53 | 70.23 | 73.68 | 78.26 | 77.50 | 82.49 | 83.94 | 84.88 |

Here's a detailed breakdown of the five mist concentration stages:²

A. Light Haze Scene:³

In light fog, the mIoU of the original image (Ori Image) is 78.58%, which is still available. At this point, the core challenge of the dehazing algorithm is to avoid destroying the image texture due to overprocessing, which can lead to a decrease in segmentation accuracy (e.g., the DCP algorithm causes the mIoU to drop to 72.15%). Our-Dehaze achieved the best segmentation index in the field at this stage, with mIoU increasing to 86.35% and mDice reaching 92.95%. This result is better than the SOTA model Mix-DehazeNet (mIoU 86.12%), proving that Our-Dehaze retains the edge features for segmentation most perfectly while removing the mist.⁴

B. Light-Medium Haze:⁵

As the fog worsened, the mIoU of the original image dropped significantly to 68.24%, and some anatomical boundaries began to blur. After Our-Dehaze treatment, mIoU rebounded significantly to 83.10%, and mDice reached 90.80%. The improvement of nearly 15 percentage points compared with the unprocessed image proves the effectiveness of the algorithm. Although Mix-DehazeNet (mIoU 83.85%) is slightly ahead at this stage, Our-Dehaze still far surpasses mainstream algorithms such as FFA-Net (81.27%) and remains in the high-performance range of the first echelon.⁶

C. Medium Haze:⁷

Moderate fog caused substantial occlusion to the visual field, and the mIoU of the original image dropped to 52.45%, which is difficult to meet clinical needs. Our-Dehaze once again showed dominance during this phase, boosting mIoU to 80.43% and mDice to 89.16%. This result surpassed Mix-DehazeNet (79.55%) and RIDCP (79.55%), indicating that after the fog concentration reached a certain⁸

threshold, Our-Dehaze was more resilient to anatomical structures and was able to convert "unusable" images into high-precision semantic maps.

D. Medium-Heavy Haze:

Robustness verification was performed under thicker smoke, and the mIoU of the original image was further reduced to 44.06%. Our-Dehaze still maintains extremely high stability, with mIoU reaching 76.50% and mDice 87.20% . It's worth noting that while Mix-DehazeNet is on par with it on mIoU (76.65%), Our-Dehaze performs better on mDice metrics (87.20% vs 86.23%). This indicates that the split mask generated by Our-Dehaze is more in line with the gold standard in terms of overall morphology and has higher internal consistency.

E. Heavy Haze:

In extremely heavy fog, the mIoU of the original image is only 35.66%, meaning that most of the anatomical structures are no longer identifiable. This is a key scenario for testing the clinical safety of algorithms. Our-Dehaze took a decisive lead in this extreme scenario, with an mIoU of 74.15% and a mDice of 84.88% . This result not only far exceeds earlier algorithms such as AOD-Net (55.45%) and EPDN (59.57%), but also significantly outperforms Mix-DehazeNet (72.84%). Our-Dehaze successfully maintained the segmentation accuracy of more than 74%, proving that it can effectively penetrate thick smoke and restore key semantic information, providing the most reliable guarantee for the operation of intelligent surgical navigation in harsh environments.

3.3 Defogging efficiency and operation time analysis

As shown in Table 9, the median duration of raw smoke was 13 min in conventional LC surgery (control group), and the lens was repeatedly removed for wiping (median 6 times), and the median total wiping time was 141 s. In contrast, after the application of the intelligent dehazing system, the single-frame image processing time is only 0.01 s, and the median overall dehazing application time is 0.02 min. Statistical analysis showed that the intelligent defogging technology significantly reduced the non-surgical operation time caused by smoke treatment ($Z = -2.167$, $P < 0.05$), and the image processing success rate reached 97% (15522/16000).

Table 9 Comparison of time parameters of intelligent defogging and conventional operation 1

| Indicators | Control group (usual operation) | Intelligent Defogging Group (Algorithmic Processing) |
|---------------------------------------|---------------------------------|---|
| Smoke impact/duration | 8 ~ 17 min (median 13) | 0.01 ~ 0.04 min (median 0.02) |
| Number of shots wiped/processed | 3 ~ 11 times (median 6) | - |
| Each processing is time-consuming | 9 ~ 21 s (Median 15) | 0.01 s (single frame) |
| Total processing is time-consuming | 69 ~ 230 s (median 141) | - |
| P-value | | < 0.05 |
| Indicators | Control group (usual operation) | Intelligent Defogging Group (Algorithmic Processing) |
| Smoke impact/duration | 8 ~ 17 min (median 13) | 0.01 ~ 0.04 min (median 0.02) |
| Number of shots wiped/processed | 3 ~ 11 times (median 6) | - |

3.4 Evaluation of visual dehazing effect 3

Figure X-3 shows the comparison of the dehazing effects of different algorithms 4 in LC surgical keyframes. The first column is the original smoke-containing image, the second and third columns are the processing results of Dehaze-NET and DCP algorithms, respectively, and the fourth column is the processing results of Yun-Transformer algorithm proposed in this study.

4 Conclusion 5

The Yun-Trans algorithm analyzes in detail the thermodynamic process of tissue 6 vaporization and carbonization caused by energy devices, as well as the optical scattering characteristics of smoke particles of different particle sizes, and modifies the atmospheric scattering model suitable for near-field point light sources, which provides a solid physical foundation for the algorithm design. The constructed dynamic expert hybrid network (Yun-Trans) innovatively introduces the "Degradation Perception Classifier" (DAC) and "Hyperparameter Selection Network" (HSN). The network does not rely on fixed parameters, but dynamically generates convolutional kernel weights based on the smoke characteristics of the input image. This mechanism

of "watching the dishes and serving dishes" realizes the adaptive processing of the whole scene from trace water mist to heavy burnt smoke.

Through the rigorous evaluation of the "deep-sea experiment" on the construction of a dataset containing 128 real surgical videos, it not only surpassed 8 mainstream algorithms in traditional indicators such as PSNR and SSIM, but also performed well in downstream tasks such as semantic segmentation (mIoU). Through retrospective clinical controlled studies, the effect of the algorithm on improving surgical efficiency (such as the number of wipes and smoke interference time) is quantified, and its application potential in practical surgical navigation is demonstrated.

Quantitative data analysis of five different fog concentration scenarios, the Yun-Trans algorithm proposed in this study shows the superior performance of full-scene adaptation. Under light and moderate fog, the algorithm achieves the optimal restoration of image color and anatomical structure with the highest PSNR (28.92 dB) and SSIM (0.935). In moderate fog interference, the algorithm prioritizes the integrity of structural information (SSIM 0.925). In challenging heavy fog environments, Yun-Trans demonstrated strong robbery, effectively restoring the surgical field of view with a significantly leading signal-to-noise ratio (22.45 dB) and naturalness index (NIQE 3.157). This performance advantage, which dynamically adjusts with the change of fog concentration, strongly verifies the effectiveness and practical value of the hybrid expert (MoE) mechanism in solving complex clinical visual interference problems.

Although the network model proposed in this study performs well in most scenarios, there are still certain limitations that need to be solved urgently. Specifically, in rare cases, extreme highlight reflections from surgical instruments can interfere with the degradation perception classifier (DAC) judgment, resulting in abnormal processing of localized areas (i.e., extreme highlight artifacts). In order to solve this problem, future research plans to introduce a light estimation module for correction to eliminate the interference of strong reflection on feature extraction and improve the stability of the model in complex lighting environments.

In terms of the actual deployment and function expansion of the algorithm, in view of the fact that the current model still relies on high-performance GPUs for inference, in order to smoothly integrate it into the existing endoscope host for device-side deployment, the follow-up work will focus on quantization and pruning of the model, and strive to develop a lightweight version with lower computational overhead. In addition, we will explore multimodal fusion technology, try to combine infrared thermal imaging or ultrasound image data, and use multimodal information to assist in penetrating thick smoke, so as to further improve the robustness of the visual enhancement system. Overall, the proposal of Our-Net marks a new stage of laparoscopic defogging technology from traditional "static filtering" to "dynamic intelligence", which is expected to lay a solid foundation for the construction of vision systems for fully automated surgical robots in the future.

Through the systematic visual evaluation and comparative analysis of surgical images processed by different dehazing algorithms, it is found that the existing conventional algorithms have different degrees of limitations in complex surgical scenarios. The dehazing ability of the Dehaze-NET algorithm is relatively limited, and obvious residual smoke and noise interference can still be observed in the processed images. Although the DCP algorithm can remove smoke to a certain extent, it is often accompanied by a significant change in image contrast, which not only destroys the realism of the image, but may also interfere with the operator's accurate judgment of the tissue. In contrast, the Yun-Transformer algorithm proposed in this study shows superior comprehensive performance, which efficiently removes smoke and noise within the line of sight while retaining the original pixel characteristics and color saturation of the image to a great extent, significantly improving the overall image quality and the clarity of anatomical boundaries, so as to provide more accurate visual support for intraoperative decision-making.