RESEARCH ARTICLE

GCT-VAE-GAN: An Image Enhancement Network for Low-Light Cattle Farm Scenes by Integrating Fusion Gate Transformation Mechanism and Variational Autoencoder GAN

CHENGCHAO WANG[1], GUOHONG GAO [1], JIANPING WANG[1], YINGYING LV[1], QIAN LI[1], ZHIYU LI[1], XUEYAN ZHANG[1], AND HAOYU WU[2]

[1]Henan Institute of Science and Technology, Xinxiang 453003, China
[2]Wenzhou University, Wenzhou 325000, China

Corresponding author: Guohong Gao (gaoguohong@hist.edu.cn)

ABSTRACT In the context of cattle farm environments, intricate environmental interferences have presented challenges that impede seamless data acquisition. This paper introduces a novel approach, the integration of a fusion gate transformation mechanism and a variational autoencoder GAN, which we term GCT-VAE-GAN, aimed at enhancing low-light images from cattle farm settings. Initially, our approach involves the design of an encoding network tasked with augmenting the original images. Subsequently, we advance our methodology by formulating a generative network to effectively address the challenges of image diversification and poor image quality. Notably, the inclusion of an attention mechanism block within the FFN layer facilitates the fusion of these extracted features, resulting in the generation of high-quality images. Furthermore, to achieve proficient image discrimination, we implement a dual-discriminator structured discriminative network for the conclusive image discrimination task. The culmination of our approach involves the formulation of a comprehensive joint loss function, thereby constituting the core of the model's loss module. Moreover, through comparative experiments, we aptly demonstrate the remarkable superiority of the GCT-VAE-GAN approach. The conducted experiments reveal the model's consistent performance and resilience under varying illumination scenarios. The outcomes of our study underscore its significant relevance in elevating the quality of low-light images within cattle farm contexts. Furthermore, our approach exhibits the potential to enhance the efficacy of computer vision tasks, signifying a notable stride toward improved agricultural imaging techniques.

INDEX TERMS Low-light image enhancement, GAN, VAE, quality evaluation.

I. INTRODUCTION

In practical scenarios, adverse weather conditions, limited lighting, and equipment constraints frequently result in the capture of underexposed images. These conditions give rise to several challenges, including diminished detail clarity, reduced contrast, and color distortion, as substantiated in prior studies [1], [2]. Consequently, the implementation of techniques aimed at enhancing low-light images holds

The associate editor coordinating the review of this manuscript and approving it for publication was Mingbo Zhao .

substantial significance. These methodologies contribute to the enhancement of image quality, the accentuation of textural intricacies, and the overall improvement of image comprehensibility and usability. The realm of deep learning has paved the way for novel approaches to address the enhancement of low-light images, as evidenced in recent literature [3], [4]. Nevertheless, existing methodologies continue to confront noteworthy challenges, particularly in striking a balance between augmenting image details and preserving the authenticity of the original image. Another significant challenge lies in the acquisition of effective feature

representations from a limited dataset of low-light images [5]. The fusion gate transformation (GCT) mechanism emerges as a promising solution that adeptly fine-tunes network parameters by assimilating inherent illumination information from images. This augmentation enhances the network's adaptability across varying lighting conditions, thus strengthening its versatility and generalization capabilities, ultimately leading to improved enhancement outcomes [6]. However, it is important to note that the computational complexity of this method is relatively high and may necessitate increased computational resources. Simultaneously, the variational autoencoder (VAE) primarily serves as a conduit for acquiring latent representations of image data, encapsulating image information within a lower-dimensional latent space [8]. The VAE operational framework comprises an encoder that maps input images to the latent space and a decoder that reconstructs the latent representation into enhanced images. By effectively traversing the continuous latent representation space, VAE streamlines the image enhancement process, mitigating potential distortions and artifacts that may arise during enhancement [9]. Furthermore, VAE excels in maintaining image coherence and authenticity while acquiring compressed representations. It should be noted that VAE may introduce some blurriness or pseudo-pixels when generating images, particularly in cases with significant randomness. In parallel, the generative adversarial network (GAN) plays a pivotal role in generating lifelike enhanced images. GAN's architecture comprises a generator responsible for producing enriched images and a discriminator designed to differentiate between enhanced and authentic images, ultimately yielding appropriate outputs. Through an adversarial training framework, GAN systematically enhances image fidelity, progressively reducing the gap between enhanced and authentic images and significantly improving visual perceptual quality [10]. It is important to acknowledge that GAN are constrained by the quality of training data and may introduce unrealistic details when generating images. The main contributions of this paper are as follows:

1. This paper introduces a novel low-light image enhancement network that seamlessly integrates the GCT, VAE, and GAN. The encoding network is dedicated to augmenting input images, while the generative network extracts salient features from these enhanced images, orchestrating feature fusion and subsequently generating refined images. Concomitantly, the discriminative network critically assesses the synthesized images, yielding enhanced generated images that closely approximate the characteristics of the original input images.

2. A unified loss function has been developed, which entails the formulation of distinct loss functions for the three modules within the proposed network: the encoding network loss, the generative network loss, and the discriminative network loss. By synergistically accounting for diverse loss components, this approach effectively integrates information pertaining to image content, contrast, and illumination.

Consequently, the network is adeptly positioned to holistically enhance low-light images within cattle farm contexts.

3. Ablation experiments were conducted to validate the contributions of individual modules within the GCT-VAE-GAN framework, followed by a demonstration of the method's superiority through comparative experiments.

The structure of this paper is organized as follows: Section II comprehensively reviews pertinent research and the technical background; section III elaborates on our proposed cattle behavior classification network in detail; section IV presents the experimental outcomes and corresponding analyses; and ultimately, section V concludes the paper while also outlining potential avenues for future research. Grounded in our research, our objective is to provide an efficient solution for enhancing low-light images in cattle farm scenarios.

II. RELATE WORKS
A. COLOR AND CONTRAST ENHANCEMENT FOR LOW-LIGHT IMAGES

Liu et al. [11] developed a comprehensive framework tailored for enhancing degraded images to serve various purposes. Zhu et al. [12] introduced a cross-view capture network designed to accomplish the super-resolution of stereoscopic images by leveraging global context and local features extracted from two different views. Additionally, Zhu et al. [13] presented a lightweight single-image super-resolution network aimed at addressing the issue of high parameter count and computational demands often associated with deep learning-based super-resolution methods. Wang et al. [14] conducted a comprehensive review of prominent techniques developed in the last decades for low-light image enhancement. These methods were classified, including grayscale transformation, histogram equalization, retinex-based techniques, frequency domain methods, image fusion methods, haze removal models, and machine learning approaches. Zhang et al. [15] introduced an innovative underwater image enhancement methodology centered around retinex-inspired color correction and the preservation of fine details. Their strategy encompasses the correction of color bias induced by light scattering through color correction techniques. In a separate vein, Lv et al. [16] presented an end-to-end attention-guided strategy founded on a multi-branch convolutional neural network, precisely tailored to address issues associated with low-light image enhancement. Hao et al. [17] outlined a retinex-centered technique for the enhancement of low-light images. Their approach was geared towards the resolution of concerns linked to the inefficient decomposition of images and pronounced imaging noise. Zhang et al. [18] presented a robust and efficient method for the enhancement of underwater images named maximum likelihood low-light enhancement (MLLE). This approach facilitated localized adjustments in color and details based on principles of minimal color loss and guided fusion using maximum attenuation maps. Their methodology further incorporated

integral and squared integral images to gauge local block means and variances, fortified by a color balancing scheme aimed at harmonizing discrepancies within the a and b channels of the commission internationale deleclairage lab (CIELAB) color space. However, when we applied these methods to low-light environments on cattle farms, we found that they did not perform well, especially in terms of tone and detail contrast. Fu et al. [19] introduced a novel unsupervised low-light image enhancement network, termed low-light image enhancement generative adversarial network (LE-GAN), which featured an illumination-aware attention module designed to bolster feature extraction. Zhao et al. [20] proposed a novel approach, redefining retinex decomposition as a generation problem. Furthermore, Zhuang et al. [21] presented a retinex variational model inspired by the super laplacian reflection prior, tailored for enhancing underwater images. Hai et al. [22] introduced a retinex-based real-to-real network (R2RNet), meticulously designed to tackle a gamut of challenges inherent in low-light images. Li et al. [23] unveiled the progressive recurrent image enhancement network (PRIEN) for the enhancement of low-light images. The enhancement of low-light images relies more on the feature extraction of the optimized model to achieve image enhancement. In the low-light image tasks applied to cattle farms, the difficulty of feature extraction will be increased due to the inclusion of more detailed feature contrast. Our GCT-VAE-GAN method solves these problems by pre-processing the coding network and then feature extraction. Li et al. [24] proposed a zero-reference deep curve estimation technique, zero-reference deep curve estimation for low-light image enhancement (Zero-DCE), showcasing its accuracy in performing face detection under low-light conditions. Additionally, Wang et al. [25] presented the deep learning network (DLN) to address the challenge of subpar visual quality in photographs captured under low-light conditions. Jiang et al. [26] devised an innovative self-regularization strategy grounded in retinex theory to surmount color deviation issues in self-adjusting low-light image enhancement. This method showcased a capacity to enhance images across a broader spectrum of lighting conditions [27].

B. APPLY TEXTURE AND SATURATION ENHANCEMENT TO LOW LIGHT IMAGES

Lim et al. [28] introduced a low-light image enhancement technique based on the advantageous properties of the Laplacian pyramid in both image and feature spaces. Ma et al. [29] developed the contextual scene decomposition (CSD) Net network to address quality issues in low-light images. This network comprises a dual-stream estimation mechanism involving reflectance and illumination estimation networks, connected through contextual scene decomposition based on physical principles. Ren et al. [30] unveiled a pioneering enhancement framework that leverages camera response characteristics to mitigate distortions in low-light images.

Jiang et al. [31] introduced enlighten GAN, an efficient unsupervised generative adversarial network for training in the absence of low/normal light image pairs. Additionally, Li et al. [32] designed a low-light image enhancement network (LE-net), a convolutional neural network-based light enhancement network, to tackle challenges inherent in low-light image enhancement. These methods primarily focus on addressing color deviation and contrast issues in low-light images. However, they may not be ideal for feature extraction from cattle bodies and enhancing the contrast of cattle images in low-light conditions. Siriani et al. [33] integrated an improved YOLOv4 model with a bird tracking algorithm, incorporating a Kalman filter to track chicken movement within low-light images. Wang et al. [34] proposed the absorption light scattering model (ALSM) to decipher the intricacies of absorption light imaging in low-light scenarios. They obtained absorption light scattering images via ALSM under uniform illumination, effectively revealing contours and concealed details present in low-light images. Lore et al. [35] introduced a deep autoencoder-based approach for real-time on-board enhancement of brightness, contrast, and noise reduction in images. Li et al. [36] presented the contrastive learning-based unpaired image enhancement network (CLUIE-Net) network to address light absorption and scattering concerns in images, while also supporting a semi-supervised learning mode. Liu et al. [11] proposed a novel variational Retinex model that concurrently estimates smooth illumination components and reflective components displaying fine details. Li et al. [37] developed a lightweight and efficient brightness-aware pyramid network (LP Net) to effectively counter underexposure in low-light image enhancement. Zhang et al. [38] introduced an attenuation map-guided color correction strategy. They incorporated a weighted wavelet visual perception fusion strategy to seamlessly combine high-frequency and low-frequency image components at varying scales, resulting in high-quality images. Lu et al. [39] proposed a dual-branch exposure fusion network to tackle blind low-light image enhancement. Their strategy involved a generative fusion approach followed by adaptive attention unit-based fusion. Yang et al. [40] introduced the deep recursive band network (DRBN) as a solution for issues involving visibility degradation and color deviation in low-light image enhancement. Additionally, Li et al. [41] introduced a robust retinex model that introduces noise images to enhance low-light images while maintaining strong noise performance. Lee et al. [42] presented a method utilizing a luminance-channel prior for single-pass enhancement. Wu et al. [43] introduced an edge computing and multi-task-driven framework designed for rapid responsive image enhancement and object detection tasks.

C. SUMMARY

Conventional techniques frequently lean on manually devised rules and heuristics for image enhancement, which can often
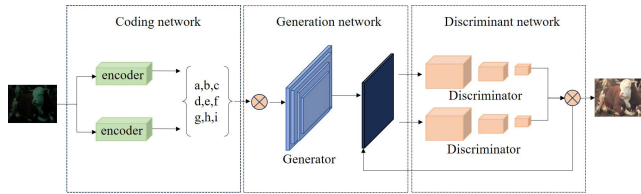
FIGURE 1. Overall architecture of the GCT-VAE-GAN method.

falter in capturing intricate image attributes. In contrast, the GCT-VAE-GAN framework excels in addressing content preservation, lighting adjustments, and detail enhancement across diverse image facets, leading to a more comprehensive and refined enhancement process. Deep learning methodologies possess the inherent capacity to autonomously derive features and patterns from data throughout the learning phase, facilitating heightened adaptability [44], [45]. However, when it comes to enhancing low-light cattle ranch images, the previously outlined model tends to exhibit limitations in terms of content retention and detail enhancement. In contrast, our proposed model proficiently learns features from data while accommodating varying lighting conditions, thereby ensuring a more adept adaptation to a wide array of lighting scenarios.

III. PROPOSED METHOD

Video frames captured from cattle farm videos serve as our input images. Initially, these images undergo processing via an encoding network to transform them into latent vectors. Subsequently, these vectors are fed into the generating network, which engages in feature extraction and fusion, thereby generating enhanced images to serve as input for the encoding network. In this process, the encoding network integrates both a global discriminator and a local discriminator. This enables a meticulous comparison between the enhanced image details and the original image, enabling the selection of high-fidelity images for retention while discarding those of inferior quality. The holistic structure of the GCT-VAE-GAN methodology is visually depicted in Figure 1.

A. DATA ACQUISITION

The study and associated experiments were conducted at a cattle farm situated in Yuanyang County. The experimental cohort consisted of 50 fully mature Simmental cattle. The cattle shed was bifurcated into two zones, separated by a cattle barrier, with feed troughs symmetrically positioned on either side of the walkway. Adjacent to the resting area, water troughs were conveniently situated for the cattle's accessibility. The vertical extent of the cattle shed measured 8 meters, while cameras were strategically mounted 2.8 meters above the ground at the extremities of the feed troughs. These cameras were precisely angled at 45 degrees downward to effectively monitor cattle activities. The doors flanking the resting area were connected to the cattle yard through barriers
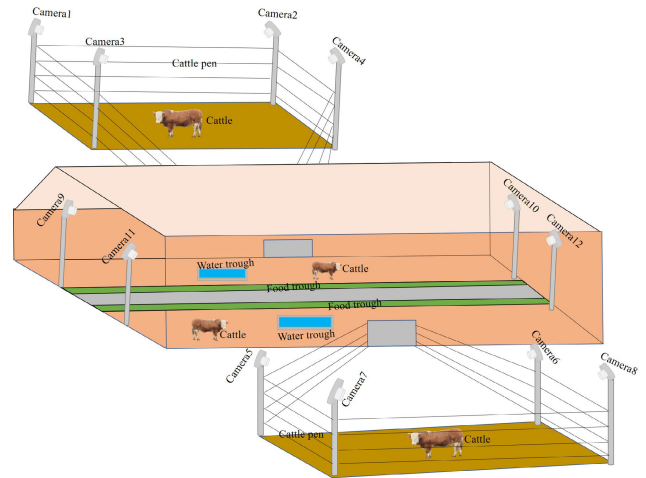

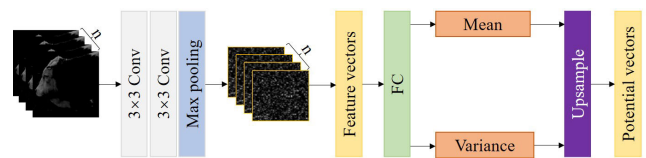**FIGURE 2.** Camera deployment diagram.

FIGURE 3. Encoding network diagram.

[46]. Enclosed by barriers, the cattle yard accommodated four cameras in each corner, positioned at an elevation of 3.5 meters above the ground. These cameras were similarly oriented at a 45-degree angle to efficiently surveil the ground. A diagram illustrates the specific deployment locations of these cameras. Captured video data was transmitted to the cloud with a resolution of 640×480 pixels. The experimental dataset was meticulously curated, encompassing images exhibiting diverse lighting conditions, as well as images with varying degrees of blurriness, sourced from the video data. Subsequent to frame extraction from the video content, a dataset housing 3000 images under both low-light and normal-light conditions was amassed. This dataset was thoughtfully partitioned into training, testing, and validation sets, adhering to an 8:1:1 ratio [47]. The camera deployment diagram is shown in Figure 2.

B. ENCODING NETWORK

Within our encoding network, the original image undergoes a transformation process, leading to the generation of n frame images. The dimensions of these images are subsequently diminished, and pivotal features are elicited through the application of two 3×3 convolutional layers, complemented by a subsequent max-pooling layer. These extracted features are then transformed into flattened feature vectors, which are further mapped onto mean and variance vectors via fully connected layers. This orchestrated mapping underpins the construction of the latent vector's distribution. Eventually, leveraging the mean and variance vectors, the latent vector is expansively upsampled, facilitating direct sampling from
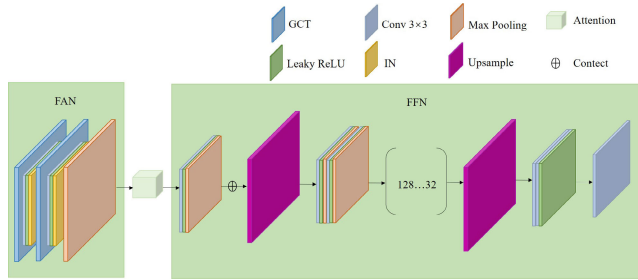
FIGURE 4. Generating network diagrams.

the latent space. This, in turn, empowers the generator to yield a spectrum of enhanced images exhibiting diversity. The diagram affords a visual elucidation of the encoding network's architecture. The coding network diagram is shown in Figure 3.

C. GENERATION NETWORK

The images procured from the encoding network serve as the input for subsequent stages. This entails subjecting these images to feature extraction via the feature extraction network (FEN). Further refinement occurs through the integration of the attention mechanism, channeling the extracted features into the feature fusion network (FFN), ultimately yielding enhanced generated images as the output. The architectural configuration of the generation network is effectively visualized through the accompanying diagram. In the progression, the enhanced images derived from the encoding network undergo initial processing via the GCT unit. This strategic step is pivotal in extracting global contextual information and modeling channel relationships. This facilitates the comprehensive aggregation and integration of global contextual insights. Subsequently, leveraging 3×3 convolutional layers, further feature extraction is executed, fortified by the leaky ReLU activation function to enhance the network's proficiency in feature extraction from images. To adeptly accommodate the multifaceted characteristics of barn image features, instance normalization (IN) layers are strategically embedded, effectively mitigating internal covariate shift. Culminating the process, the introduction of a max-pooling layer effectively trims computational complexity while upholding pivotal features, culminating in the desired output. As for the FFN, it is structurally comprised of convolutional residual blocks intricately linked by lateral skip connections. These blocks encompass a 3×3 convolutional layer, a leaky ReLU activation function, and a max-pooling layer. Operating in tandem, these residual blocks play a pivotal role in the fusion of image features extracted by the FEN, particularly at higher hierarchical levels. This orchestrated design proficiently prevents the unwarranted loss of valuable information during forward propagation. The architectural layout of the generation network is vividly portrayed through the illustrative diagram. The generating network diagrams is shown in Figure 4.
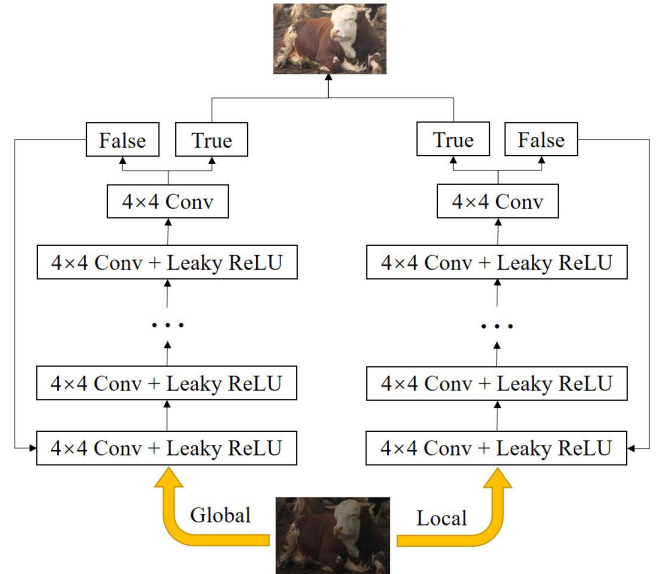

FIGURE 5. Discriminant network diagram.

D. DISCRIMINANT NETWORK

Given the intricate interplay of intricate details and textures in low-light images inherent to the complex cattle farm environment, a dual-discriminator architecture is embraced for the construction of the discriminative network. The overarching framework incorporates a global discriminator, which ingests images enhanced by the generation network. This facet is facilitated through the utilization of 4×4 convolutional layers coupled with residual blocks, underpinned by Leaky ReLU activation functions. This combination serves the purpose of heightening color intensity while concurrently enhancing overall brightness and contrast of the images. Following this enhancement, the global discriminator leverages subsequent 4×4 convolutional layers to discern the distinction between authentic and synthesized images. In instances where the input garners verification as authentic, an image is then produced; conversely, if it is deemed spurious, it is subjected to iterative re-evaluation. Parallelly, the local discriminator addresses the challenge of non-uniform local illumination in the improved images by randomly cropping four equivalently sized image patches as input. These input segments subsequently undergo evaluation through a series of residual blocks and convolutional layers, culminating in the derivation of real and synthetic values. In particular, the synthetic values undergo a re-evaluation step. The comprehensive structure of the discriminative network finds its articulate representation through the illustrative diagram. The discriminant network diagram is shown in Figure 5.

E. JOINT LOSS FUNCTION
1) ENCODING NETWORK LOSS
The enhanced images generated in the encoding network preserve the essential details of the input images, thus

necessitating the use of a reconstruction loss. We employ the Mean Squared Error (MSE) loss to quantify the pixel-wise differences between the original low-light images and their corresponding reconstructed images. The reconstruction loss for a single image pair (I, I') is defined as follows in (1).

$$L_{recon} = \frac{1}{N} \sum_{i=1}^{N} (I_i - I'_i)^2 \tag{1}$$

where N represents the total number of pixels in the image, I and I' denote the intensity values of the i pixel in the original image. The KL divergence loss is given by the following as shown in (2).

$$L_{KL} = \frac{1}{2} \sum_{i=1}^{N} \left(\mu_i^2 + \sigma_i^2 - \log\left(\sigma_i^2\right) - 1 \right) \tag{2}$$

where μ_i and σ_i respectively represent the mean and standard deviation of the i element in the encoded latent vector. The encoding loss is given by the following as shown in (3)

$$L_{cod} = L_{recon} + L_{KL} \tag{3}$$

2) GENERATING NETWORK LOSS

Our generation network's loss function is composed of adversarial loss and GCT loss, aimed at generating visually plausible images with focused spatial attention. The adversarial loss is intended to prevent the discriminative network from distinguishing between generated images and real images. The calculation of the adversarial loss is given by the following is shown in (4).

$$L_{adv} = -\mathbb{E}_{I_{real}} \left[\log \left(D\left(I_{real}\right) \right) \right] \\ - \mathbb{E}_{I_{enh}} \left[\log \left(1 - D\left(G\left(I_{enh}\right)\right) \right) \right] \tag{4}$$

where I_{real} represents the real low-light image, I_{enh} represents the enhanced image, G represents the generation network, and D represents the discriminative network. The GCT loss encourages the generation network to focus on important regions in the image, thus promoting spatial attention. The calculation method is provided by the following as shown in (5).

$$L_{GCT} = \frac{1}{N} \sum_{i=1}^{N} (A_i - A'_i)^2 \tag{5}$$

where A_i represents the attention map of the real image, A'_i represents the attention map of the enhanced image, and N represents the total number of elements in the attention map. The loss function for the generation network is given by the following is shown in (6).

$$L_{gen} = L_{adv} + L_{GCT} \tag{6}$$

3) DISCRIMINATE NETWORK LOSS

The discriminative network loss is comprised of global discriminator loss and local discriminator loss. The global discriminator loss consists of the logarithmic likelihood loss of real images and the logarithmic likelihood loss of

---

Algorithm 1 Low-Light Image Enhancement With GCT-VAE-GAN

Input: Low-light image I_{low};

Data preprocessing:
Perform data normalization on I_{low};

Encoding network:
Calculate global contrast map $M_{contrast}$ using I_{low};
Apply contrast enhancement to I_{low} based on $M_{contrast}$;
$\{z^i\} = E(I_{low}^i)$; //Initial enhancement of low light images

Generation network:
Encode I_{low} using VAE to obtain latent representation Z_{latent};
Generate enhanced latent representation $Z_{enhanced}$ by adjusting Z_{latent};
Decode $Z_{enhanced}$ to obtain $I_{VAE-enhanced}$; // Feature extraction and feature fusion
$\{I_{enhanced}^i\} = G(z^i)$; //Generate an enhanced image

Discriminant network:
Train a GAN model using I_{low} and $I_{VAE-enhanced}$;
Generate enhanced image $I_{GAN-enhanced}$ using the trained GAN model;
$\{D(I_{enhanced}^i)\}$; // Identify the generated image and output the enhanced image

Joint Loss Function:
Fuse $I_{GAN-enhanced}$ and $I_{VAE-enhanced}$ using a fusion method;
Obtain the final enhanced image $I_{enhanced}$;
$L_{reconstruction} = \frac{1}{N} \sum_i \|I_{enhanced}^i - I_{low}^i\|$;
$L_{adversarial} = -\frac{1}{N} \sum_i \log(D(I_{enhanced}^i))$;
$L_{feature} = \frac{1}{N} \sum_i \|F(E(I_{enhanced}^i)) - F(I_{enhanced}^i)\|$; //Calculate the loss of each part

return $I_{enhanced}$;
Output: Enhanced image $I_{enhanced}$;

---

generated images. Its objective is to distinguish between real and generated images. The global discriminator loss is defined by the following is shown in (7).

$$L_{D_{glo}} = -\mathbb{E}_{x_{real}} \left[\log D_{glo}\left(x_{real}\right) \right] \\ - \mathbb{E}_{x_g} \left[\log \left(1 - D_{glo}\left(x_g\right) \right) \right] \tag{7}$$

where D_{glo} represents the global discriminator, x_{real} represents real images, and x_g represents generated images. The task of the local discriminator is to determine the authenticity of the local regions in the images. We randomly select local blocks from real images and compare them with corresponding blocks from generated images. The calculation of the local discriminator loss is given by the following as

TABLE 1. Hyper parameter setting.

Parameters	Value	Parameters	Value
cuda	10.0.1	GCT weights	5
cudnn	10.0.1	VAE weights	3
pytorch	1.2.0	GAN weights	6
initial learning rate	0.001	final decay rate	0.95
termination learning rate	0.0001	number of epochs	150
regularization term	0.01	optimizer	adam
activation function	leaky ReLU	momentum factor	0.875

shown in (8).

$$L_{D_{loc}} = -\mathbb{E}_{x_{real,r}}\left[\log D_{loc}\left(x_{real}^r\right)\right]$$
$$- \mathbb{E}_{x_g,r}\left[\log\left(1 - D_{loc}\left(x_g^r\right)\right)\right] \quad (8)$$

where $x_{real,r}$ represents the local block sampled from the real image x_{real}^r, x_g, r represents the corresponding sampled local block from the generated image x_g^r, r is the index of the local block. The loss function for the discriminative network is given by the following is shown in (12).

$$L_{dis} = L_{D_{glo}} + L_{D_{loc}} \quad (9)$$

GCT-VAE-GAN algorithns are shown in Algorithm 1.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. EXPERIMENTAL SETTINGS

1) EXPERIMENTAL ENVIRONMENT AND PARAMETER SETTINGS

The experiments were conducted on the LINUX operating system with 16GB of memory. The GPU used was an NVIDIA GeForce RTX 3070 graphics card, while the CPU was an Intel Core i7. The network training was configured using an Intel(R) Core (TM) i7-10750H CPU @ 2.60GHz 2.59GHz processor. The model was built using the pytorch deep learning framework. The hyperparameters employed during the training phase are outlined in the table 1.

The experiment selected peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), naturalness image quality evaluator (NIQE), and level of enhancement (LOE) as evaluation metrics. PSNR is used to measure the difference between the enhanced and original images. In our task, a higher PSNR value indicates a smaller difference between the enhanced and original images, leading to better image quality. PSNR is calculated by computing the mean squared error (MSE) between the enhanced and original images and then converting it to a logarithmic scale in decibels. The calculation method is given by the following is shown in (10) (11).

$$\text{MSE} = \frac{1}{M \times N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(I\left(i,j\right) - I'\left(i,j\right)\right)^2 \quad (10)$$

$$\text{PSNR} = 10 \times \log_{10}\left(\frac{\text{MAX}^2}{\text{MSE}}\right) \quad (11)$$

where M and N represent image coefficients, I and I' respectively denote the original and enhanced images, and *MAX* is the maximum possible pixel value. SSIM is a

| (a) Input | (b) GCT-VAE-GAN | (c) w/o GCT | (d) w/o VAE | (e) w/o GAN |

FIGURE 6. Ablation experiments for each module of GCT-VAE-GAN.

TABLE 2. Comparison of performance under ablation for each module of GCT-VAE-GAN. In each case, the best result is shown in red, and the second best result is shown in blue.

Method	PSNR	SSIM
w/o GCT	16.38	0.498
w/o VAE	13.27	0.412
w/o GAN	14.97	0.453
GCT-VAE-GAN	17.59	0.523

TABLE 3. Experimental results for joint losses. In each case, the best result is shown in red, and the second best result is shown in blue.

Method	NIQE	LOE
L_{cod}	4.35	586
$L_{cod} + L_{gen}$	3.87	406
$L_{cod} + L_{dis}$	4.09	458
$L_{gen} + L_{dis}$	3.39	376
$L_{cod} + L_{gen} + L_{dis}$	3.15	282

| Input | $L_{cod} + L_{gen} + L_{dis}$ | $L_{gen} + L_{dis}$ | $L_{cod} + L_{gen}$ | $L_{cod} + L_{dis}$ | L_{cod} |

FIGURE 7. Experimental results for components of joint losses.

measure of structural similarity, which takes into account not only differences in pixel values between images but also considers their structural information. In our task, SSIM comprehensively reflects the degree of similarity between the enhanced and original images. The calculation method for SSIM is given by the following as shown in (12).

$$\text{SSIM}(x,y) = \frac{\left(2\mu_x\mu_y + C_1\right) \times \left(2\sigma_{xy} + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right) \times \left(\sigma_x^2 + \sigma_y^2 + C_2\right)} \quad (12)$$

where x represents the original image, y represents the enhanced image, μ_x and μ_y are the means of x and y respectively, σ_x and σ_y represent the standard deviations of x and y, σ_{xy} represents the covariance of x and y, C_1 and C_2 are two constants used to prevent division by zero. NIQE is used in our mission to assess the quality of low-light images of cattle farms, which are calculated as follows in (13).

$$\text{NIQE} = a \cdot \sigma_a + s \cdot \sigma_s + c \cdot \sigma_c + \mu_p \cdot \sigma_{\mu_p} + \mu \cdot \sigma_\mu + \sigma^2 \cdot \sigma_\sigma \quad (13)$$

where a stands represents coarseness, s represents structure, c represents contrast, σ represents standard deviation, μ represents mean, μ_p represents pixel mean.

B. ABLATION EXPERIMENT

We conducted ablation studies on each module of the GCT-VAE-GAN method. Specifically, we started by ablating

TABLE 4. Results of comparison experiments. In each case, the best result is shown in red, and the second best result is shown in blue.

Method	PSNR	SSIM	RT
RUAS	14.66	0.463	26.751
DLN [22]	14.78	0.447	12.214
CNN [46]	13.42	0.419	3.783
Retinex Net [12]	16.79	0.502	0.314
Lighten Net [23]	15.98	0.483	0.102
Enlighten GAN [27]	16.38	0.497	0.068
GCT-VAE-GAN	17.61	0.521	0.021

the GCT module. During training and testing on our self-constructed dataset, we removed the operations related to the GCT module and performed experiments solely with the VAE-GAN module to obtain experimental data and resultant images. Similarly, we then proceeded to individually remove the VAE and GAN modules. The results of the ablation experiments are presented in Table 2 and Figure 6.

We conducted experiments involving the joint loss function. We evaluated the image quality using individual components of the encoding network loss, generation network loss, and discriminative network loss. Through experimental results, we observed that employing the joint loss function yielded the best outcomes, resulting in higher visual quality of the images. The results of the joint loss experiment are shown in Table 3 and Figure 7.

C. COMPARISON EXPERIMENT

To assess the superior performance of the GCT-VAE-GAN method, a meticulously planned series of comparative experiments was conducted employing self-constructed datasets. Given the distinctive characteristics of the cattle farm environment, six distinct benchmark networks were chosen for in-depth comparison: CNN, DLN, RUAS, Retinex Net, Lighten Net, and Enlighten GAN. The selection of these benchmark networks was driven by their unique architectural designs and diverse feature extraction methodologies. Specifically, CNN is renowned for its versatility and applicability to image enhancement tasks across various domains. However, it may prove less effective in preserving image details and contrast under low-light conditions. In contrast, DLN employs an extensive array of modules and layers to enhance image quality, demonstrating superior performance in low-light scenarios. Nevertheless, the limitation of DLN lies in its demand for substantial training data, which can potentially lead to overfitting. RUAS emerges as a prominent method in image enhancement, distinguished by its application of residual connections and attention mechanisms, particularly suitable for enhancing aerial low-light images from various angles. However, the utilization of RUAS may entail higher costs and, for our specific tasks, serves more as a point of reference. Grounded in Retinex theory, Retinex Net concentrates on image illumination and contrast enhancement, making it well-suited for low-light image enhancement tasks. Yet, it exhibits sensitivity to image noise, and its performance may exhibit significant variability across different datasets. Lighten Net focuses exclusively on

(a) Inputs (b) RUAS (c) DLN (d) CNN

(e) Retinex Net (f) Lighten Net (g) Enlighten GAN (h) GCT-VAE-GAN

FIGURE 8. Comparative experiment on low light data in cattle farm.

TABLE 5. Experimental evaluation on two publicly available datasets, LOL [48] and SICE [49]. In each case, the best result is shown in red, and the second best result is shown in blue.

Method	LOL dataset [48]			SICE dataset [49]		
	PSNR	SSIM	RT	PSNR	SSIM	RT
RUAS	14.617	0.466	25.343	13.045	0.472	23.945
DLN [22]	13.718	0.421	11.367	13.121	0.438	10.268
CNN [46]	12.402	0.418	4.145	11.953	0.402	5.197
LIME	13.159	0.612	5.389	14.289	0.597	4.986
HE	14.273	0.479	2.876	15.109	0.443	3.021
LLNet	17.285	0.418	1.313	17.139	0.635	1.289
BIMEF	15.196	0.525	0.926	15.478	0.564	0.216
Retinex Net [12]	16.235	0.502	0.314	15.987	0.523	0.182
Lighten Net [23]	15.342	0.479	0.124	16.372	0.512	0.069
Enlighten GAN [27]	17.163	0.684	0.082	17.489	0.693	0.076
GCT-VAE-GAN	18.672	0.754	0.035	18.368	0.734	0.032

optimizing image quality in low-light image enhancement. However, it may grapple with color distortion in complex environments. On the other hand, Enlighten GAN places distinct emphasis on elevating the quality of low-light images by leveraging GAN frameworks to generate lifelike enhanced images. Nevertheless, this approach entails relatively high computational complexity. In addition to the aforementioned aspects, we consider elapsed time (RT) as a crucial factor in our evaluation, quantifying the speed of execution for each model. A detailed compilation of the results from these comprehensive analyses can be shown in Table 4, while a visual representation is shown in Figure 8.

To comprehensively assess the model's generalization performance, we conducted a series of experiments evaluating its generalization effects using two publicly available datasets. These datasets exhibit a wide coverage of diverse fields and represent various environments, extending beyond the specific scenarios encountered in cattle farming. To ensure a thorough comparison of our methods, we included several low-light image enhancement techniques, such as LIME, HE, LLNet, and BIMEF. LIME, grounded in Retinex theory, strives to enhance low-light images while preserving the natural details within the images. However, it exhibits relatively lower performance and imposes higher data quality

FIGURE 9. Comparative experiments on LOL and SICE datasets.

requirements. The HE method boasts a simple yet effective network structure, which successfully enhances the contrast of low-light images. However, it often introduces some noise, diminishing the natural feel of the images. LLNet specializes in enhancing images under low-light conditions, delivering commendable performance. Nevertheless, its relatively complex network structure results in high computational complexity, limiting its applicability. BIMEF introduces the theory of light and dark channel prior, facilitating the preservation of texture details in images. However, this method necessitates manual parameter adjustments, which detracts from its level of automation. Our thorough experimental evaluation encompassed these distinct low-light image enhancement models, and the detailed results are presented in Table 5 and Figure 9.

V. CONCLUSION

This paper introduces the GCT-VAE-GAN method, specifically designed for enhancing low-light images in cattle farm environments. Ablation experiments were meticulously conducted to validate the efficacy of each individual module, including GCT, VAE, and GAN. Notably, the GCT module emerged as a pivotal contributor, proficient at extracting global features and information. This strategic role helped prevent the introduction of unnatural artifacts and distortions in the generated images. Simultaneously, the VAE module prioritized the learning of latent space representations, substantially enriching the model's generative capacity and generalization through effective regularization. The outcome resulted in a perceptible proximity between the generated images and the distribution of authentic images. Complementing this, the GAN module, facilitated by dual discriminators, significantly enhanced the realism and naturalness of the generated images. The assessment of the GCT-VAE-GAN method's capabilities was rigorously established using a purpose-built dataset. The method's superiority was conclusively substantiated through meticulous comparisons with six existing techniques, including CNN, DLN, RUAS, Retinex Net, Lighten Net, and Enlighten GAN. In this evaluation, the GCT-VAE-GAN method demonstrated distinct advantages in effectively managing low-light conditions, adeptly preserving intricate details, and markedly improving overall image quality. Despite these achievements, it is

important to acknowledge that the utilization of cattle farm low-light image data introduced certain challenges. Notably, suboptimal outcomes were observed in the presence of factors such as cattle barriers and inclement weather conditions. Addressing these challenges remains the focal point of our future research endeavors as we aim to enhance the robustness and adaptability of our proposed solution.

REFERENCES

[1] Q. Zhang, Q. Yuan, M. Song, H. Yu, and L. Zhang, "Cooperated spectral low-rankness prior and deep spatial prior for HSI unsupervised denoising," *IEEE Trans. Image Process.*, vol. 31, pp. 6356–6368, 2022, doi: 10.1109/TIP.2022.3211471.

[2] A. Poursaberi, C. Bahr, A. Pluk, A. Van Nuffel, and D. Berckmans, "Real-time automatic lameness detection based on back posture extraction in dairy cattle: Shape analysis of cow with image processing techniques," *Comput. Electron. Agricult.*, vol. 74, no. 1, pp. 110–119, Oct. 2010, doi: 10.1016/j.compag.2010.07.004.

[3] K. Zhao, J. M. Bewley, D. He, and X. Jin, "Automatic lameness detection in dairy cattle based on leg swing analysis with an image processing technique," *Comput. Electron. Agricult.*, vol. 148, pp. 226–236, May 2018, doi: 10.1016/j.compag.2018.03.014.

[4] X. Song, T. Leroy, E. Vranken, W. Maertens, B. Sonck, and D. Berckmans, "Automatic detection of lameness in dairy cattle—Vision-based trackway analysis in cow's locomotion," *Comput. Electron. Agricult.*, vol. 64, no. 1, pp. 39–44, Nov. 2008, doi: 10.1016/j.compag.2008.05.016.

[5] C. Li, C. Guo, L. Han, J. Jiang, M.-M. Cheng, J. Gu, and C. C. Loy, "Low-light image and video enhancement using deep learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 12, pp. 9396–9416, Dec. 2022, doi: 10.1109/TPAMI.2021.3126387.

[6] W. Zhang, Z. Li, H.-H. Sun, Q. Zhang, P. Zhuang, and C. Li, "SSTNet: Spatial, spectral, and texture aware attention network using hyperspectral image for corn variety identification," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2022.3225215.

[7] T. Xu, W. Zhao, L. Cai, X. Shi, and X. Wang, "Lightweight saliency detection method for real-time localization of livestock meat bones," *Sci. Rep.*, vol. 13, Mar. 2023, Art. no. 4510, doi: 10.1038/s41598-023-31551-6.

[8] T. Xu, W. Zhao, X. Meng, Y. Ma, and J. Zhou, "Adversarial learning-based method for recognition of bionic and highly contextual underwater targets," *J. Electron. Imag.*, vol. 32, no. 2, Apr. 2023, Art. no. 023027, doi: 10.1117/1.JEI.32.2.023027.

[9] G. Kim, S. W. Park, and J. Kwon, "Pixel-wise Wasserstein autoencoder for highly generative dehazing," *IEEE Trans. Image Process.*, vol. 30, pp. 5452–5462, 2021, doi: 10.1109/TIP.2021.3084743.

[10] G.-D. Fan, B. Fan, M. Gan, G.-Y. Chen, and C. L. P. Chen, "Multiscale low-light image enhancement network with illumination constraint," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 11, pp. 7403–7417, Nov. 2022, doi: 10.1109/TCSVT.2022.3186880.

[11] Y. Liu, Z. Yan, J. Tan, and Y. Li, "Multi-purpose oriented single nighttime image haze removal based on unified variational retinex model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 4, pp. 1643–1657, Apr. 2023, doi: 10.1109/TCSVT.2022.3214430.

[12] X. Zhu, K. Guo, H. Fang, L. Chen, S. Ren, and B. Hu, "Cross view capture for stereo image super-resolution," *IEEE Trans. Multimedia*, vol. 24, pp. 3074–3086, 2022, doi: 10.1109/TMM.2021.3092571.

[13] X. Zhu, K. Guo, S. Ren, B. Hu, M. Hu, and H. Fang, "Lightweight image super-resolution with expectation-maximization attention mechanism," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 3, pp. 1273–1284, Mar. 2022, doi: 10.1109/TCSVT.2021.3078436.

[14] W. Wang, X. Wu, X. Yuan, and Z. Gao, "An experiment-based review of low-light image enhancement methods," *IEEE Access*, vol. 8, pp. 87884–87917, 2020, doi: 10.1109/ACCESS.2020.2992749.

[15] W. Zhang, L. Dong, and W. Xu, "Retinex-inspired color correction and detail preserved fusion for underwater image enhancement," *Comput. Electron. Agricult.*, vol. 192, Jan. 2022, Art. no. 106585, doi: 10.1016/j.compag.2021.106585.

[16] F. Lv, Y. Li, and F. Lu, "Attention guided low-light image enhancement with a large scale low-light simulation dataset," *Int. J. Comput. Vis.*, vol. 129, no. 7, pp. 2175–2193, Jul. 2021, doi: 10.1007/s11263-021-01466-8.

[17] S. Hao, X. Han, Y. Guo, X. Xu, and M. Wang, "Low-light image enhancement with semi-decoupled decomposition," *IEEE Trans. Multimedia*, vol. 22, no. 12, pp. 3025–3038, Dec. 2020, doi: 10.1109/TMM.2020.2969790.

[18] W. Zhang, P. Zhuang, H.-H. Sun, G. Li, S. Kwong, and C. Li, "Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement," *IEEE Trans. Image Process.*, vol. 31, pp. 3997–4010, 2022, doi: 10.1109/TIP.2022.3177129.

[19] Y. Fu, Y. Hong, L. Chen, and S. You, "LE-GAN: Unsupervised low-light image enhancement network using attention module and identity invariant loss," *Knowl.-Based Syst.*, vol. 240, Mar. 2022, Art. no. 108010, doi: 10.1016/j.knosys.2021.108010.

[20] Z. Zhao, B. Xiong, L. Wang, Q. Ou, L. Yu, and F. Kuang, "RetinexDIP: A unified deep framework for low-light image enhancement," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 3, pp. 1076–1088, Mar. 2022, doi: 10.1109/TCSVT.2021.3073371.

[21] P. Zhuang, J. Wu, F. Porikli, and C. Li, "Underwater image enhancement with hyper-Laplacian reflectance priors," *IEEE Trans. Image Process.*, vol. 31, pp. 5442–5455, 2022, doi: 10.1109/TIP.2022.3196546.

[22] J. Hai, Z. Xuan, R. Yang, Y. Hao, F. Zou, F. Lin, and S. Han, "R2RNet: Low-light image enhancement via real-low to real-normal network," *J. Vis. Commun. Image Represent.*, vol. 90, Feb. 2023, Art. no. 103712, doi: 10.1016/j.jvcir.2022.103712.

[23] J. Li, X. Feng, and Z. Hua, "Low-light image enhancement via progressive-recursive network," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 11, pp. 4227–4240, Nov. 2021, doi: 10.1109/TCSVT.2021.3049940.

[24] C. Li, C. Guo, and C. C. Loy, "Learning to enhance low-light image via zero-reference deep curve estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 8, pp. 4225–4238, Aug. 2022, doi: 10.1109/TPAMI.2021.3063604.

[25] L.-W. Wang, Z.-S. Liu, W.-C. Siu, and D. P. K. Lun, "Lightening network for low-light image enhancement," *IEEE Trans. Image Process.*, vol. 29, pp. 7984–7996, 2020, doi: 10.1109/TIP.2020.3008196.

[26] Z. Jiang, H. Li, L. Liu, A. Men, and H. Wang, "A switched view of retinex: Deep self-regularized low-light image enhancement," *Neurocomputing*, vol. 454, pp. 361–372, May 2021, doi: 10.1016/j.neucom.2021.05.025.

[27] Z. Zhao, Y. Zhou, B. Liu, J. He, J. Zhao, Y. Cai, J. Fan, X. Li, Z. Wang, Z. Lu, J. Wu, H. Qi, and Q. Dai, "Two-photon synthetic aperture microscopy for minimally invasive fast 3D imaging of native subcellular behaviors in deep tissue," *Cell*, vol. 186, no. 11, pp. 2475–2491.e22, May 2023, doi: 10.1016/j.cell.2023.04.016.

[28] S. Lim and W. Kim, "DSLR: Deep stacked Laplacian restorer for low-light image enhancement," *IEEE Trans. Multimedia*, vol. 23, pp. 4272–4284, 2021, doi: 10.1109/TMM.2020.3039361.

[29] L. Ma, R. Liu, J. Zhang, X. Fan, and Z. Luo, "Learning deep context-sensitive decomposition for low-light image enhancement," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 10, pp. 5666–5680, Oct. 2022, doi: 10.1109/TNNLS.2021.3071245.

[30] Y. Ren, Z. Ying, T. H. Li, and G. Li, "LECARM: Low-light image enhancement using the camera response model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 4, pp. 968–981, Apr. 2019, doi: 10.1109/TCSVT.2018.2828141.

[31] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "EnlightenGAN: Deep light enhancement without paired supervision," *IEEE Trans. Image Process.*, vol. 30, pp. 2340–2349, 2021, doi: 10.1109/TIP.2021.3051462.

[32] G. Li, Y. Yang, X. Qu, D. Cao, and K. Li, "A deep learning based image enhancement approach for autonomous driving at night," *Knowl.-Based Syst.*, vol. 213, Feb. 2021, Art. no. 106617, doi: 10.1016/j.knosys.2020.106617.

[33] A. L. R. Siriani, V. Kodaira, S. A. Mehdizadeh, I. de Alencar Nääs, D. J. de Moura, and D. F. Pereira, "Detection and tracking of chickens in low-light images using YOLO network and Kalman filter," *Neural Comput. Appl.*, vol. 34, no. 24, pp. 21987–21997, Dec. 2022, doi: 10.1007/s00521-022-07664-w.

[34] Y.-F. Wang, H.-M. Liu, and Z.-W. Fu, "Low-light image enhancement via the absorption light scattering model," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5679–5690, Nov. 2019, doi: 10.1109/TIP.2019.2922106.

[35] K. G. Lore, A. Akintayo, and S. Sarkar, "LLNet: A deep autoencoder approach to natural low-light image enhancement," *Pattern Recognit.*, vol. 61, pp. 650–662, Jan. 2017, doi: 10.1016/j.patcog.2016.06.008.

[36] K. Li, L. Wu, Q. Qi, W. Liu, X. Gao, L. Zhou, and D. Songing, "Beyond single reference for training: Underwater image enhancement via comparative learning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 6, pp. 2561–2576, Jun. 2023, doi: 10.1109/TCSVT.2022.3225376.

[37] J. Li, J. Li, F. Fang, F. Li, and G. Zhang, "Luminance-aware pyramid network for low-light image enhancement," *IEEE Trans. Multimedia*, vol. 23, pp. 3153–3165, 2021, doi: 10.1109/TMM.2020.3021243.

[38] W. Zhang, L. Zhou, P. Zhuang, G. Li, X. Pan, W. Zhao, and C. Li, "Underwater image enhancement via weighted wavelet visual perception fusion," *IEEE Trans. Circuits Syst. Video Technol.*, early access, Jul. 27, 2023, doi: 10.1109/TCSVT.2023.3299314.

[39] K. Lu and L. Zhang, "TBEFN: A two-branch exposure-fusion network for low-light image enhancement," *IEEE Trans. Multimedia*, vol. 23, pp. 4093–4105, 2021, doi: 10.1109/TMM.2020.3037526.

[40] W. Yang, S. Wang, Y. Fang, Y. Wang, and J. Liu, "Band representation-based semi-supervised low-light image enhancement: Bridging the gap between signal fidelity and perceptual quality," *IEEE Trans. Image Process.*, vol. 30, pp. 3461–3473, 2021, doi: 10.1109/TIP.2021.3062184.

[41] M. Li, J. Liu, W. Yang, X. Sun, and Z. Guo, "Structure-revealing low-light image enhancement via robust retinex model," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2828–2841, Jun. 2018, doi: 10.1109/TIP.2018.2810539.

[42] H. Lee, K. Sohn, and D. Min, "Unsupervised low-light image enhancement using bright channel prior," *IEEE Signal Process. Lett.*, vol. 27, pp. 251–255, 2020, doi: 10.1109/LSP.2020.2965824.

[43] Y. Wu, H. Guo, C. Chakraborty, M. Khosravi, S. Berretti, and S. Wan, "Edge computing driven low-light image dynamic enhancement for object detection," *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 5, pp. 3086–3098, Sep./Oct. 2022, doi: 10.1109/TNSE.2022.3151502.

[44] J. Wu, Y. Guo, C. Deng, A. Zhang, H. Qiao, Z. Lu, J. Xie, L. Fang, and Q. Dai, "An integrated imaging sensor for aberration-corrected 3D photography," *Nature*, vol. 612, no. 7938, pp. 62–71, Dec. 2022, doi: 10.1038/s41586-022-05306-8.

[45] J. Wu et al., "Iterative tomography with digital adaptive optics permits hour-long intravital observation of 3D subcellular dynamics at millisecond scale," *Cell*, vol. 184, no. 12, pp. 3318–3332, Jun. 2021, doi: 10.1016/j.cell.2021.04.029.

[46] G. Gao, C. Wang, J. Wang, Y. Lv, Q. Li, Y. Ma, X. Zhang, Z. Li, and G. Chen, "CNN-Bi-LSTM: A complex environment-oriented cattle behavior classification network based on the fusion of CNN and bi-LSTM," *Sensors*, vol. 23, no. 18, p. 7714, Sep. 2023, doi: 10.3390/s23187714.

[47] G. Gao, C. Wang, J. Wang, Y. Lv, Q. Li, X. Zhang, Z. Li, and G. Chen, "UD-YOLOv5s: Recognition of cattle regurgitation behavior based on upper and lower jaw skeleton feature extraction," *J. Electron. Imag.*, vol. 32, no. 4, Aug. 2023, Art. no. 043036, doi: 10.1117/1.JEI.32.4.043036.

[48] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," 2018, *arXiv:1808.04560*.

[49] J. Cai, S. Gu, and L. Zhang, "Learning a deep single image contrast enhancer from multi-exposure images," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 2049–2062, Apr. 2018, doi: 10.1109/TIP.2018.2794218.

CHENGCHAO WANG was born in Anyang, Henan, China, in 1997. He is currently pursuing the master's degree with the School of Information Engineering, Henan Institute of Science and Technology. His research interests include agricultural engineering and information technology.
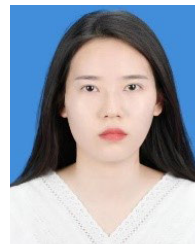
GUOHONG GAO is currently a Professor with the School of Information Engineering, Henan University of Science and Technology. His research interest includes wireless sensor networks.

JIANPING WANG was born in Baoji, Shaanxi, China, in 1981. He received the M.S. degree from the Department of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China, in 2011. He is currently pursuing the Ph.D. degree with the Wuhan University of Technology. He is also an Associate Professor with the School of Information Engineer, Henan Institute of Science and Technology. His research interests include underwater acoustic sensor networks, software-defined networks, and wireless sensor networks.

YINGYING LV is currently a Teacher with the School of Information Engineering, Henan Institute of Science and Technology. Her main teaching direction is computer science and technology.

QIAN LI is currently a Teacher with the School of Information Engineering, Henan Institute of Science and Technology. Her main research interest includes big data.

ZHIYU LI was born in Xinxiang, Henan, in 1999. He is currently pursuing the master's degree in agricultural engineering and information technology with the Henan Institute of Science and Technology.

XUEYAN ZHANG was born in Luoyang, Henan, China, in 1995. She is currently pursuing the master's degree with the Henan Institute of Science and Technology. Her current research direction is agricultural engineering and information technology.

HAOYU WU was born in Xinxiang, Henan, China, in 2003. He is currently pursuing the bachelor's degree with Wenzhou University. His areas of expertise include artificial intelligence and computer science and technology.

• • •