Tech Science Press

Deep-Learning-Empowered 3D Reconstruction for Dehazed Images in IoT-Enhanced Smart Cities

Jing Zhang[1,2], Xin Qi[3,*], San Hlaing Myint[3] and Zheng Wen[4]

[1]Department of Computer Science and Technology, Xi'an University of Science and Technology, Xi'an, 710054, China
[2]Tamoritsusho Co., Ltd., Tokyo, 110-0005, Japan
[3]Global Information and Telecommunication Institute, Waseda University, Tokyo, 169-8050, Japan
[4]School of Fundamental Science and Engineering, Waseda University, Tokyo, 169-8050, Japan
*Corresponding Author: Xin Qi. Email: samqixin@aoni.waseda.jp
Received: 29 January 2021; Accepted: 01 March 2021

Abstract: With increasingly more smart cameras deployed in infrastructure and commercial buildings, 3D reconstruction can quickly obtain cities' information and improve the efficiency of government services. Images collected in outdoor hazy environments are prone to color distortion and low contrast; thus, the desired visual effect cannot be achieved and the difficulty of target detection is increased. Artificial intelligence (AI) solutions provide great help for dehazy images, which can automatically identify patterns or monitor the environment. Therefore, we propose a 3D reconstruction method of dehazed images for smart cities based on deep learning. First, we propose a fine transmission image deep convolutional regression network (FT-DCRN) dehazing algorithm that uses fine transmission image and atmospheric light value to compute dehazed image. The DCRN is used to obtain the coarse transmission image, which can not only expand the receptive field of the network but also retain the features to maintain the nonlinearity of the overall network. The fine transmission image is obtained by refining the coarse transmission image using a guided filter. The atmospheric light value is estimated according to the position and brightness of the pixels in the original hazy image. Second, we use the dehazed images generated by the FT-DCRN dehazing algorithm for 3D reconstruction. An advanced relaxed iterative fine matching based on the structure from motion (ARI-SFM) algorithm is proposed. The ARI-SFM algorithm, which obtains the fine matching corner pairs and reduces the number of iterations, establishes an accurate one-to-one matching corner relationship. The experimental results show that our FT-DCRN dehazing algorithm improves the accuracy compared to other representative algorithms. In addition, the ARI-SFM algorithm guarantees the precision and improves the efficiency.

Keywords: 3D reconstruction; dehazed image; deep learning; fine transmission image; structure from motion algorithm

1 Introduction

Artificial intelligence (AI) has recently become very popular, and a wide range of applications use this technique [1]. There are many smart systems based on deep learning for social services, such as smart cities and smart transportation. Smart cities can use AI technology such as machine learning, deep learning and computer vision to save money and improve the quality of life of residents [2]. Through using AI or machine learning technology to perform intelligent image processing, we can obtain important geographical region data. Such real-time data can be continuously monitored through AI technology, which would further develop cities' governance and planning [3].

In hazy environments, the reflected light of an object is attenuated before it reaches the camera or monitoring equipment, resulting in the degradation of the quality of an outdoor image [4]. Therefore, obtaining dehazed images in smart cities is an important problem to be solved by AI technology. In recent years, with the rapid development of AI technology, some dehazing algorithms based on deep learning have been proposed. Tang et al. [5] used the random forest algorithm to remove haze. Although the accuracy of the transmission image was improved, the texture features in the image were not used, which has certain limitations; and the effect of dehazing is not ideal. Cai et al. [6] adopted the convolutional neural network to learn the features of hazy image to estimate the transmission image. The convolutional neural network only uses a single scale for feature extraction, which makes it prone to color distortion, detail loss and excessive dehazing for many specific scenes. Li et al. [7] proposed AOD-Net based on a convolutional neural network to dehaze images. To avoid using additional methods to estimate atmospheric light via mathematical transformation, the network structure of this algorithm is relatively simple.

We propose a fine transmission image deep convolutional regression network (FT-DCRN) dehazing algorithm that uses fine transmission image [8] and atmospheric light value [9] to compute dehazed image. First, this paper proposes a deep convolutional regression network (DCRN) to obtain the coarse transmission image. The DCRN can not only expand the receptive field of the network, but it can also retain the features to maintain the nonlinearity of the overall network [10]. Second, the fine transmission image is obtained by refining the coarse transmission image using a guided filter [11]. The guided filter is used to optimize the coarse transmission image to improve the accuracy of dehazed images. Furthermore, the atmospheric light value is estimated according to the position and brightness of the pixels in the original hazy image. According to the obtained fine transmission image and atmospheric light value, the dehazed image is inverted using the atmospheric physical scattering model [12].

With increasingly more smart cameras deployed in infrastructure and commercial buildings, 3D reconstruction can quickly obtain information on cities and geographical regions [13]. It is important to solve the image matching problem using structure from motion (SFM) 3D reconstruction algorithms [14]. Feature detection and feature matching are subordinate image matching problems. Hossain et al. [15] proposed a CADT corner detection algorithm, which effectively reduces the positioning error and improves the average repeatability. Zhang et al. [16] proposed a Harris SIFT algorithm including illumination compensation, which not only improves the matching accuracy but also improves the real-time performance of the algorithm. However, the above image matching algorithms are not universal and cannot accurately extract image feature points under special lighting conditions. Zhou et al. [17] proposed a registration algorithm based on geometric invariance and local similar features, but it relies heavily on rough matching, and the correct matching points are eliminated. To solve the problems of the above algorithms, an

advanced relaxed iterative fine matching based on the SFM (ARI-SFM) algorithm is proposed. The ARI-SFM algorithm, which obtains the fine matching corner pairs and reduces the number of iterations, establishes an accurate one-to-one corner matching relationship.

The contributions of this paper are listed as follows:

(a) We use a deep learning algorithm for dehazed images in smart cities. The FT-DCRN dehazing algorithm is proposed, which uses fine transmission image and atmospheric light value to compute dehazed image. First, this paper proposes a DCRN algorithm to obtain the coarse transmission image. The DCRN can not only expand the receptive field of the network, but can also retain the features to maintain the nonlinearity of the overall network. Second, the fine transmission image is obtained by refining the coarse transmission image using a guided filter. The guided filter is used to optimize the coarse transmission image to improve the accuracy of dehazed images.

(b) We perform 3D reconstruction using the dehazed images generated from the FT-DCRN algorithm. The ARI-SFM algorithm is proposed, which can obtain fine matching corner pairs and reduce the number of iterations. Compared with other representative algorithms, the ARI-SFM algorithm establishes an accurate one-to-one corner matching relationship, which guarantees the precision and improves the efficiency.

2 Our Approach

2.1 FT-DCRN Dehazing Algorithm

The purpose of a dehazing algorithm is to restore a sharp image from a blurred image caused by haze. Deep learning algorithms can provide great help for dehazy images [18], which can automatically identify patters or monitor the environment. In this paper, we propose a FT-DCRN dehazing algorithm to obtain useful images for 3D reconstruction. The steps of the FT-DCRN dehazing algorithm are as follows:

Step 1: Obtain the coarse transmission image. Input the hazy images, and the coarse transmission image is obtained by using the DCRN dehazing algorithm.

Step 2: Compute the fine transmission image. The fine transmission image is obtained by refining the coarse transmission image using a guided filter.

Step 3: Estimate the atmospheric light value. The atmospheric light value [19] is estimated according to the position and brightness of the pixels in the original hazy image.

Step 4: Compute the dehazed image. According to the obtained fine transmission image and atmospheric light value, the dehazed image is inverted using the atmospheric physical scattering model.

2.1.1 DCRN Dehazing Algorithm

To obtain a coarse transmission image, this paper proposes a DCRN dehazing algorithm, and the overall network structure is shown in Fig. 1. The DCRN is an end-to-end network based on a convolutional neural network [20] that inputs hazy images and outputs corresponding coarse transmission images.

The DCRN is similar to the encoder–decoder network. The core unit of the encoder network is the convolutional unit (Conv), which is mainly composed of a convolutional layer [21], an ReLU, a pooling layer and a batch normalization (BN) layer [22]. The core unit of the decoder network is the deconvolutional unit (DeConv), which is mainly composed of a deconvolutional

layer, a BN layer, an ReLU and a convolutional layer. The fully connected (FC) layer [23] is replaced by the convolutional layer.
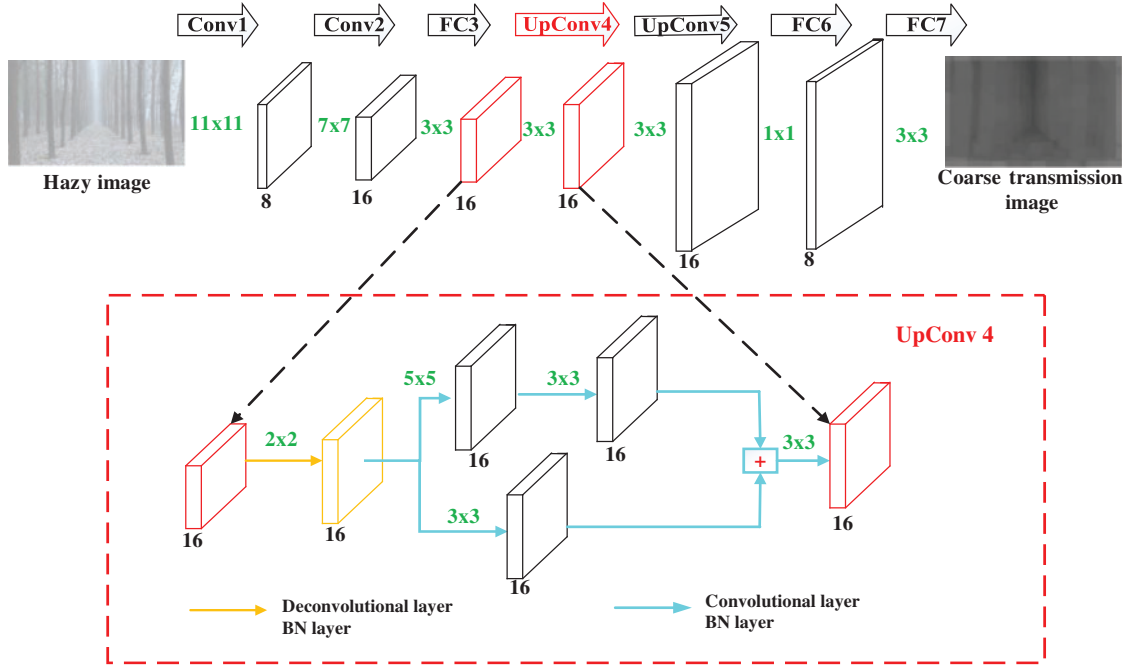
Figure 1: Overall network structure of DCRN dehazing algorithm

The features of the overall network are extracted by the encoder network. The decoder network is used to ensure the size of the output transmission image and retain the features to maintain the nonlinearity of the overall network [24]. Our DCRN can not only expand the receptive field of the network, but it can also ensure that the overall network has a certain nonlinear learning ability.

The main characteristic of an end-to-end network is that the input and output of the network are identical in size. However, due to the use of two pooling layers [25] in the encoder network, the feature set is smaller, and the original image information is lost.

To solve the problem of information loss from the original image, this paper uses a decon-volutional layer [26] to replace the upsampling layer, which can not only increase the size of the feature set, but can also produce a dense feature set with a larger spatial structure, the "Upconv4" is shown in the red box of Fig. 1. The Upconv4 contains the deconvolutional layer, convolutional layer and BN layer. The deconvolutional layer is often used in the densest mapping estimation problem. Furthermore, the cross-convolutional layer used in the DeConv unit can not only provide the DCRN with the multiscale feature learning ability, but can also avoid the vanishing gradient problem in the backpropagation process. Therefore, the DCRN can estimate the coarse transmission image more accurately.

2.1.2 Fine Transmission Image

In this paper, a guided filter is used to optimize the coarse transmission image to improve the accuracy of dehazed images [27]. The guided filter can be defined as:

$$q_{(x,y)} = a_k M_{(x,y)} + b_k \ ((x,y) \in \omega_k) \tag{1}$$

where q is the output image, which is the fine transmission image obtained after optimization. a_k and b_k are the coefficients of the window ω_k.

$$E(a_k, b_k) = \Sigma \left(\left(a_k M_{(x,y)} + b_k - p_{(x,y)} \right)^2 + \varepsilon a_k^2 \right) \tag{2}$$

where ε is the regularization parameter. A linear regression was used to obtain the following results:

$$a_k = \frac{\Sigma_{(x,y)\in\omega_k} M_{(x,y)}/\omega - p_{(x,y)} - u_k \overline{p}_k}{\sigma_k^2 + \varepsilon} \tag{3}$$

$$b_k = \overline{p}_k - a_k u_k \tag{4}$$

where u_k and σ_k^2 are the average and variance of image M in current window ω_k, respectively. $\overline{p}_k$ is the average value of p in window ω_k.

2.1.3 Atmospheric Light Value

When estimating the atmospheric light value, He et al. [28] selected the pixels with the top one percent brightnesses in the hazy image and then calculated the average brightness of these pixels as the atmospheric light value. This method is more effective in most cases, but when a large white area appears in the image, the method will not accurately estimate the atmospheric light value, which will lead to image color distortion.

To solve the above problems, this paper uses the method of combining the pixel position and brightness to estimate the atmospheric light value. The relative height of each pixel is defined as $H(x,y)$, and the brightness value is $V(x,y)$. The probability of a pixel being located in the white area is defined as follows:

$$P(x,y) = H(x,y) \ V(x,y) \tag{5}$$

The process determines the pixels with the probability value $P(x,y)$ of being among the top one percent and uses the average brightness value of these pixels as the atmospheric light value.

2.2 ARI-SFM Algorithm

It is important to solve the image matching problem in the 3D SFM reconstruction algorithm. In the image matching process, the coarse matching relationship between corners is established by using the zero mean normalized cross-correlation method [29]. This method only builds a one-to-many set of matching corner pairs, so there are many unclear and incorrect matching pairs. We propose an ARI-SFM algorithm to guarantee the precision and improve the efficiency.

2.2.1 ARI Algorithm

To establish an accurate one-to-one corner matching relationship, an advanced relaxed iterative (ARI) algorithm is proposed, which obtains fine matching corner pairs and reduces the number of iterations. The flowchart of the ARI algorithm is shown in Fig. 2.
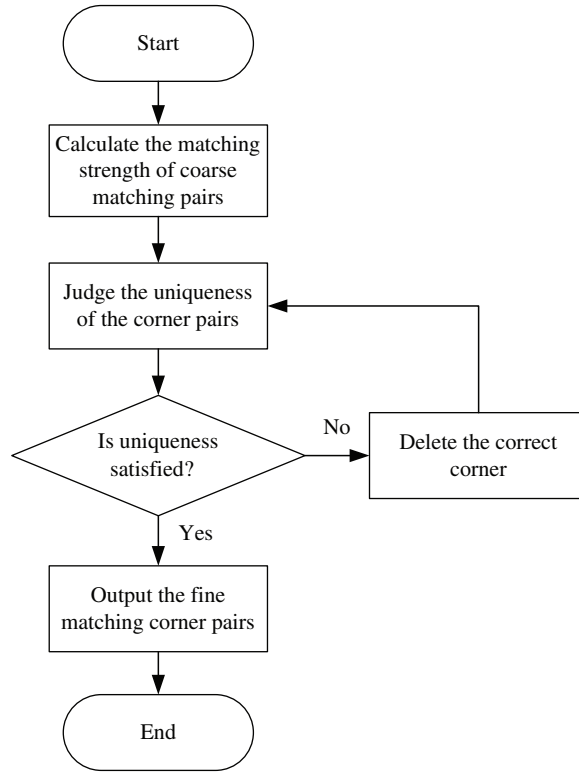
Figure 2: Flowchart of the ARI algorithm

The steps of the ARI algorithm are as follows:

Step 1: Calculate the matching strength of coarse matching pairs. The matching strength is used as the indicator for fine matching corner selection [30].

Step 2: Judge the uniqueness of the corner pairs. The matching pairs are sorted according to the matching strength from large to small. We select the corner pairs $S_M(p_{1i}, q_{2j})$ and $S'_M(p_{1i}, q_{2j})$ with the largest and second largest matching strengths, respectively. We calculate $S_P(p_{1i}, q_{2j})$ and use it to measure the uniqueness of the corner matching.

$$S_P(p_{1i}, q_{2j}) = 1 - \frac{S'_M(p_{1i}, q_{2j})}{S_M(p_{1i}, q_{2j})} \tag{6}$$

The value range of S_P is 0~1. According to $S_M(p_{1i}, q_{2j})$ and $S_P(p_{1i}, q_{2j})$, all matching pairs in the set are sorted. If all corners are in the top 60% of the two items, the corner pairs are accurate matching pairs. Then, proceed to Step 4; otherwise, proceed to Step 3.

Step 3: Delete the correct corner after fine matching, and return to Step 1.

Step 4: Output the fine matching corner pairs.

2.2.2 Calculation of Matching Strength

The initial matching corner pairs are represented as (p_{1i}, q_{2j}), where p_{1i} is the corner of image I_1 and q_{2j} is the corner of image I_2. $N(p_{1i})$ and $N(q_{2j})$ are neighborhoods with point p_{1i} and point q_{2j} as centers, respectively, and R as the radius. If (p_{1k}, q_{2f}) is the correct matching pair, there

must be more correct matching pairs (p_{1k}, q_{2f}) in its neighborhoods $N(p_{1i})$ and $N(q_{2j})$, where $p_{1k} \in N(p_{1i})$ and $q_{2f} \in N(q_{2j})$ satisfy the conditions for calculating the matching strength.

Condition 1: Only the matching corner pair (p_{1k}, q_{2f}) affects the matching corner pair (p_{1i}, q_{2j}).

Condition 2: The angle between $\overline{p_{1i}p_{1k}}$ and $\overline{q_{2j}q_{2f}}$ is less than 90 degrees.

If the matching corner pairs (p_{1i}, q_{2j}) and (p_{1k}, q_{2f}) satisfy the above two conditions, the matching strength is calculated using formula (7):

$$S_M = \sum_{p_{1k} \in N(q_{1i})} \max_{q_{2f} \in N(q_{2j})} \left[\frac{Similar_{ij} Similar_{kf} \Delta(p_{1i}, p_{1k}, q_{2j}, q_{2f})}{1 + dist(p_{1i}, p_{1k}, q_{2j}, q_{2f})} \right] \tag{7}$$

$Similar_{ij}$ and $Similar_{kf}$ are the gray cross-correlation values of the matching corner pairs (p_{1i}, q_{2j}) and (p_{1k}, q_{2f}), respectively. $dist$ is the average distance of the corner pairs. The expression of δ is shown in formula (8):

$$\delta(p_{1i}, p_{1k}, q_{2j}, q_{2f}) = \begin{cases} e^{-r/\varepsilon_r}, & r < \varepsilon_r \\ 0, & others \end{cases} \tag{8}$$

r is the relative distance deviation of the corner pair. The similarity contribution δ of (p_{1k}, q_{2f}) to (p_{1i}, q_{2j}) is a power function with a negative exponent and relative distance deviation r. δ is a monotonically decreasing function of r. When r is very large, the matching corner pair (p_{1k}, q_{2f}) is ignored.

3 Experiments

In this paper, synthetic hazy images and real hazy images are used to train and test the performance of the FT-DCRN dehazing algorithm. First, we adopt the Make3D dataset [31] (http://make3d.cs.cornell.edu/data.html) to synthesize hazy images using the atmospheric scattering model. We selected 900 pairs of hazy and sharp images as training samples. Second, we take 1200 real hazy images of outdoor scenes, such as those of buildings, gardens, and parking areas, to analyze the results of the FT-DCRN dehazing algorithm. The FT-DCRN dehazing algorithm runs on a GeForce RTX 2080Ti GPU and executes using Python.

To verify the efficiency and accuracy of the ARI-SFM algorithm, the algorithm is implemented on an experimental platform with 64-bit Windows 10, an Intel(R) Core(TM) i5-10210U@1.60 GHZ CPU, and 8.00 GB of memory; the development platform is MATLAB R2018b.

3.1 Generation of Synthetic Hazy Images

Given a random value for the transmission image $t(x, y) \in [0, 1]$ and the atmospheric light value $A \in [200, 255]$, the synthetic hazy image $I(x)$ is generated by formula (9).

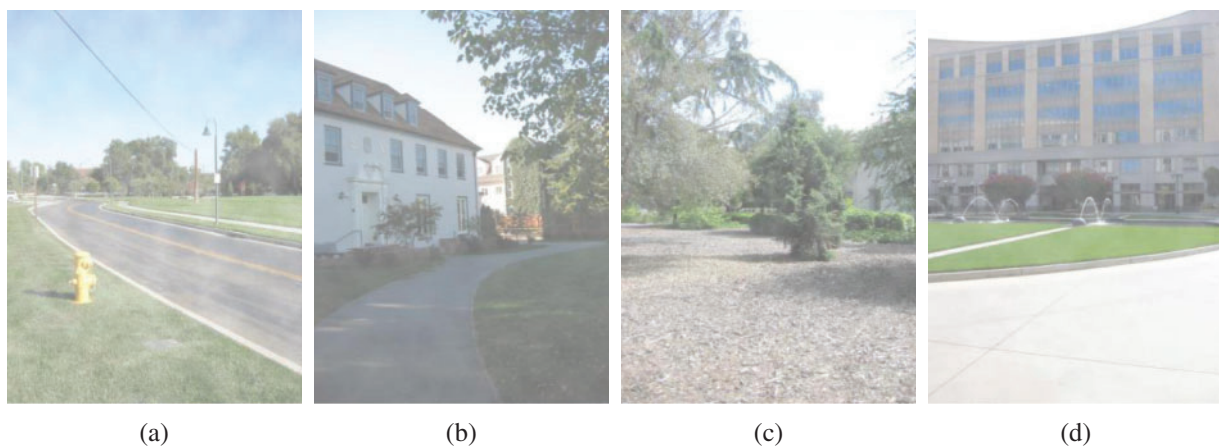$$I(x, y) = J(x, y) t(x, y) + A(1 - t(x, y)) \tag{9}$$

where $J(x, y)$ represents the original sharp image.

Fig. 3 shows the original sharp images, including images of a road, house, tree, and fountain, from the Make3D dataset. Fig. 4 shows the corresponding synthetic hazy image of the sharp images in the Make3D dataset.

Figure 3: Original sharp images from the Make3D dataset. (a) Road. (b) House. (c) Tree. (d) Fountain

Figure 4: Corresponding synthetic hazy images from the Make3D dataset. (a) Road. (b) House. (c) Tree. (d) Fountain

3.2 Experiment of FT-DCRN Dehazing Algorithm
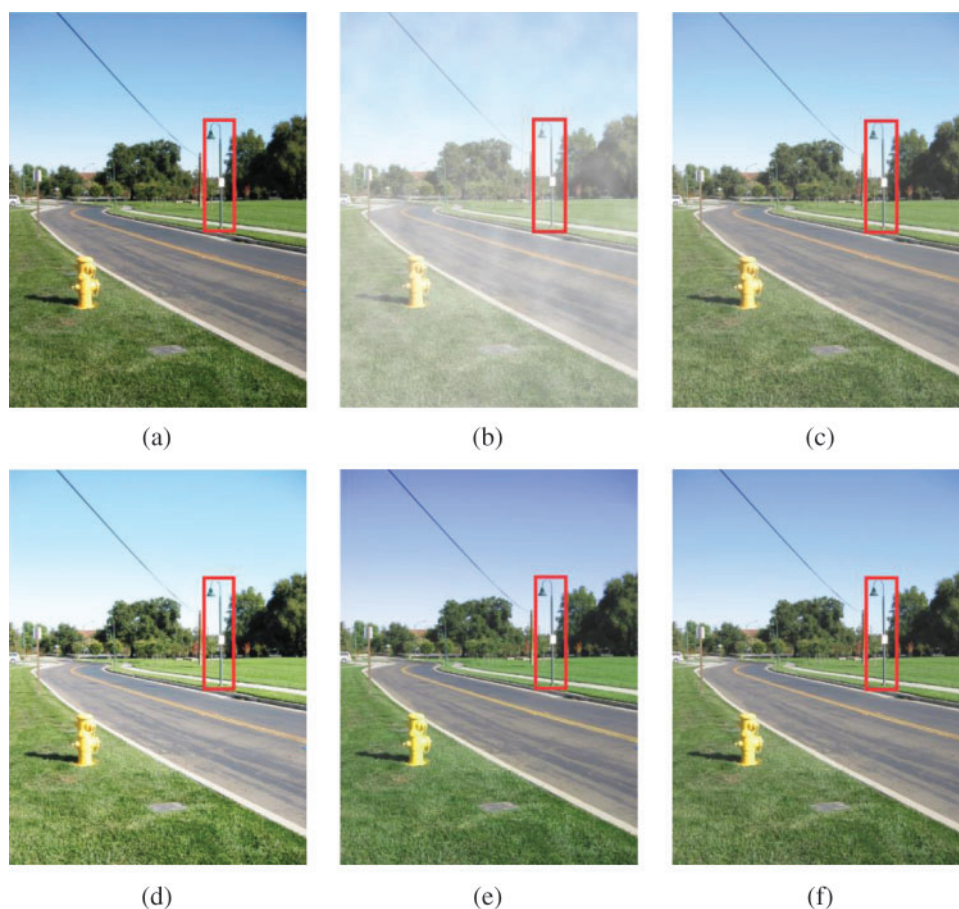
3.2.1 Qualitative Evaluation

(1) Results of Synthetic Hazy Images

To verify the effect of the FT-DCRN dehazing algorithm on the synthetic hazy images, the results of the algorithm are compared with some representative algorithms. Because different deep learning algorithms have their own advantages,we adopt the Tang's algorithm [5], Cai's algorithm [6] and Li's algorithm [7], which are described in the introduction. Figs. 5–8 show the comparison results of the dehazing of the road, house, tree and fountain images in the Make3D dataset.

Figs. 5–8 show that Tang's algorithm does not consider the texture features in the images, and the results of the algorithm have unclear boundaries. Cai's algorithm has color oversaturation in dehazed images, resulting in large areas of color distortion. Li's algorithm contains detailed

information, but the overall colors of the images are not close to the normal visual effect. Our approach is superior to other methods. The boundaries of the dehazed images are clear, and the textures are close to those of the real images.

Figure 5: Comparison results of the dehazing of the road images. (a) Sharp image. (b) Synthetic hazy image. (c) Tang's algorithm. (d) Cai's algorithm. (e) Li's algorithm. (f) Our approach
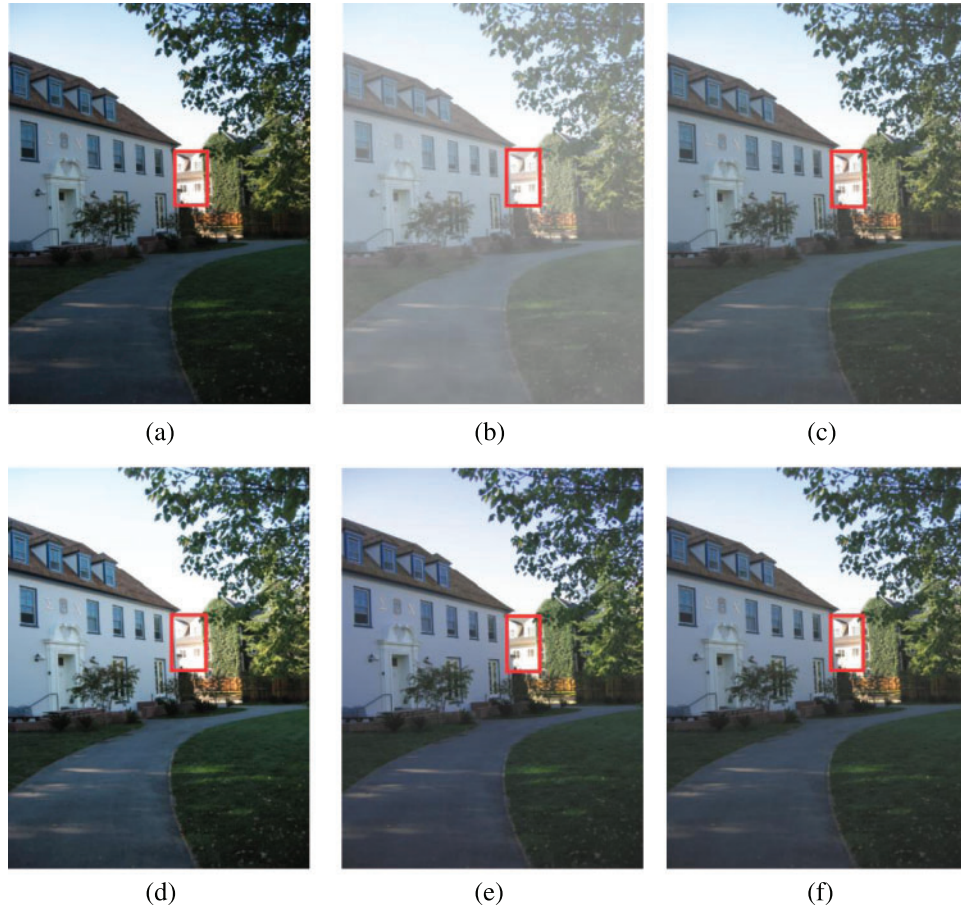
(2) Results of Real Hazy Image

To verify the effect of the FT-DCRN dehazing algorithm on real hazy images, we analyze 1200 real hazy images of outdoor scenes, such as building, garden and parking area. We compare the results of the FT-DCRN with those of Tang's algorithm, Cai's algorithm and Li's algorithm. Fig. 9 show the comparison results of the dehazing of the images of building.

Fig. 9 show that Tang's algorithm results in unclear boundaries for the buildings. Cai's algorithm easily produces color distortion, which makes the scene of the buildings look unreal. Li's algorithm changes the color of the white areas. Our approach has clear boundaries and textures, and the overall colors of the images are close to the normal visual effect.

3.2.2 Quantitative Evaluation

To perform the quantitative evaluation, synthetic hazy images and real hazy images are selected. We adopt the structural similarity (SSIM) [32], peak signal-to-noise ratio (PSNR) [33] and information entropy (IE) [34] to evaluate the effect of the FT-DCRN dehazing algorithm. The SSIM is an indicator of the similarity of two images. When two images are the same, the SSIM is equal to 1. The PSNR is a statistical indicator that is based on the gray values of image pixels. The higher the PSNR is, the better the image restoration. The IE is a statistical measure of features that reflects the average amount of information in the image. The larger the entropy is, the clearer the image. The experimental results for the synthetic hazy image are shown in Tabs. 1–4.

Figure 6: Comparison results of the dehazing of the house images. (a) Sharp image. (b) Synthetic hazy image. (c) Tang's algorithm. (d) Cai's algorithm. (e) Li's algorithm. (f) Our approach

The results of Tang's algorithm provided relatively low values for each indicator. Tang's algorithm did not use the texture features in the image, which creates certain limitations on the dehazing effect. Cai's algorithm and Li's algorithm have significantly higher SSIM and PSNR values than those of Tang's algorithm. However, Cai's algorithm and Li's algorithm do not result in normal visual color effects in Figs. 5–8. The IE of our approach is higher than those of other algorithms, which reflects that the dehazed image retains more detail and texture information.
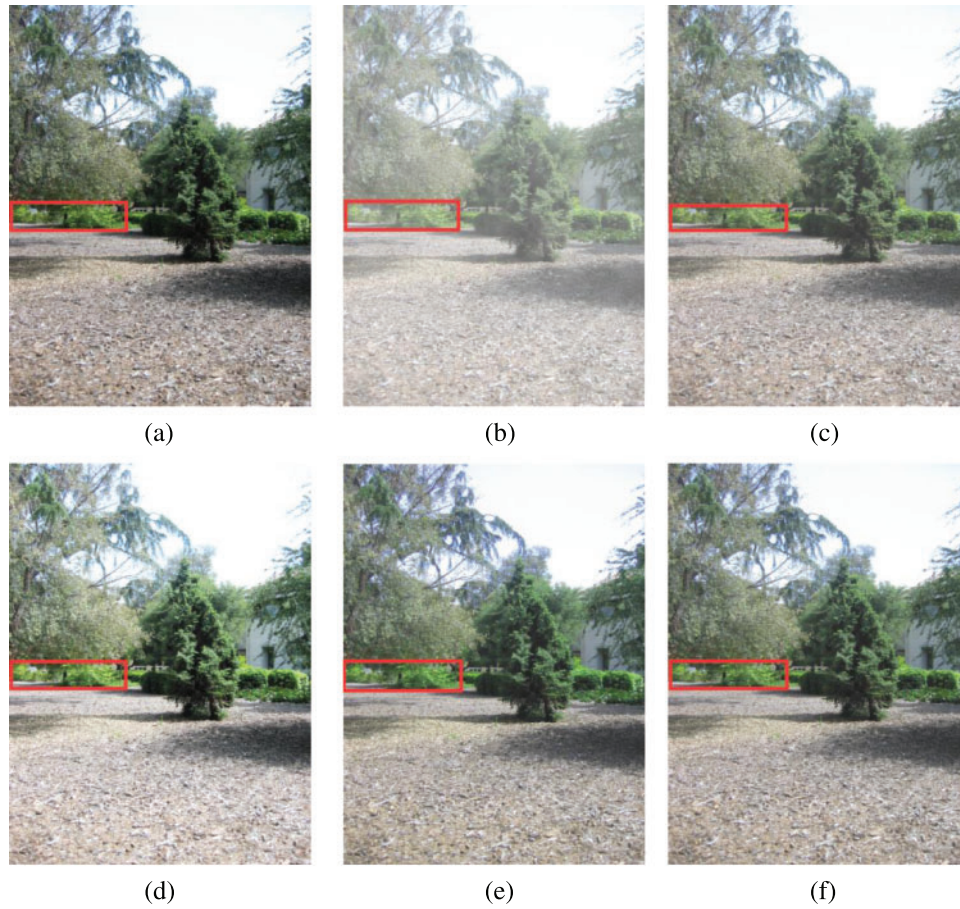
Figure 7: Comparison results of the dehazing of the tree images. (a) Sharp image. (b) Synthetic hazy image. (c) Tang's algorithm. (d) Cai's algorithm. (e) Li's algorithm. (f) Our approach

3.3 Experiment of ARI-SFM Algorithm

3.3.1 Results of 3D Reconstruction

Figs. 10a–10f includes six images of a building taken from different perspectives. The images of the building from different perspectives are taken from a real scene.

After using the ARI-SFM algorithm, the one-to-one relationship between corners is determined, and one-to-many relationship almost does not exist. In this experiment, we selected 6 images shown in Fig. 10 for 3D reconstruction. The final experimental results are shown in Fig. 11. Fig. 11 is the point cloud of 3D reconstruction of building. The Fig. 11 shows that the ARI-SFM algorithm can accurately reconstruct the 3D building, and the signs on the building are clearly visible.

3.3.2 Performance of 3D Reconstruction

To verify the matching efficiency of the ARI-SFM algorithm, the results of the algorithm are compared with some representative algorithms. Because different image matching algorithms have their own advantages, we adopt the Hossain's algorithm [15], Zhang's algorithm [16] and Zhou's algorithm [17], which are described in the introduction. We analyze the comparison of the matching results of the building, as shown in Tab. 5.

(a)   (b)   (c)

(d)   (e)   (f)
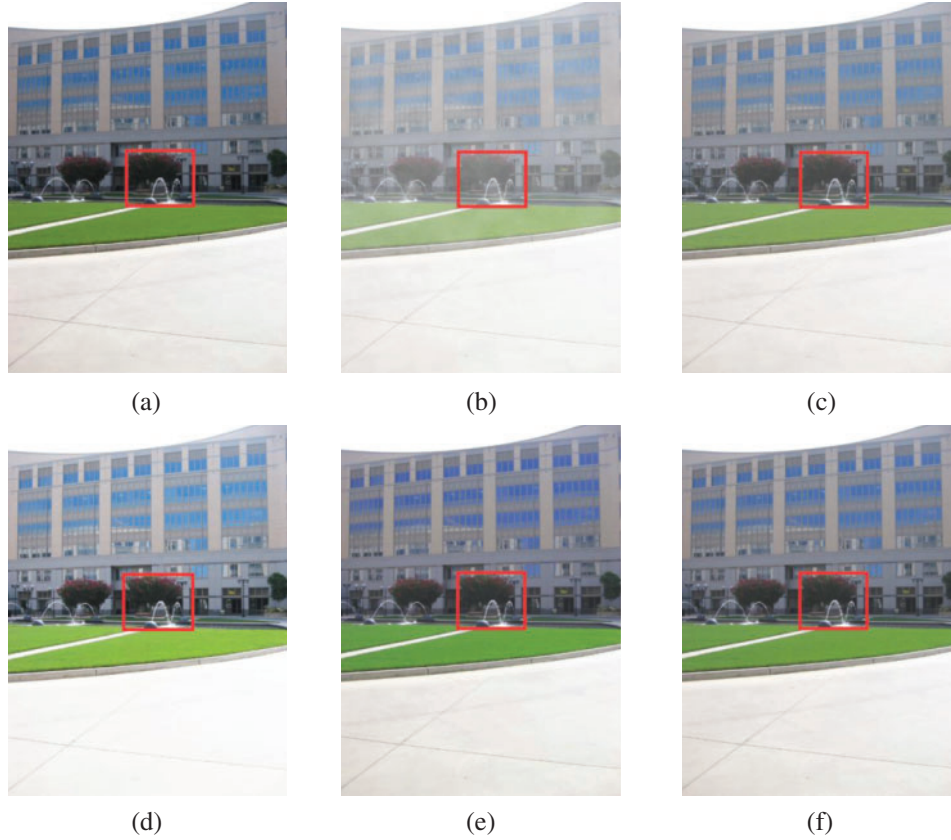
Figure 8: Comparison results of the dehazing of the fountain images. (a) Sharp image. (b) Synthetic hazy image. (c) Tang's algorithm. (d) Cai's algorithm. (e) Li's algorithm. (f) Our approach

Tab. 5 show that our approach have higher matching accuracy and cost less match time, which indicates that we guarantee the precision and improve the efficiency compared with other algorithms. Our approach determine the one-to-one relationship between corners and almost does not exist one-to-many relationship, which obtains fine matching corner pairs and reduces the number of iterations.
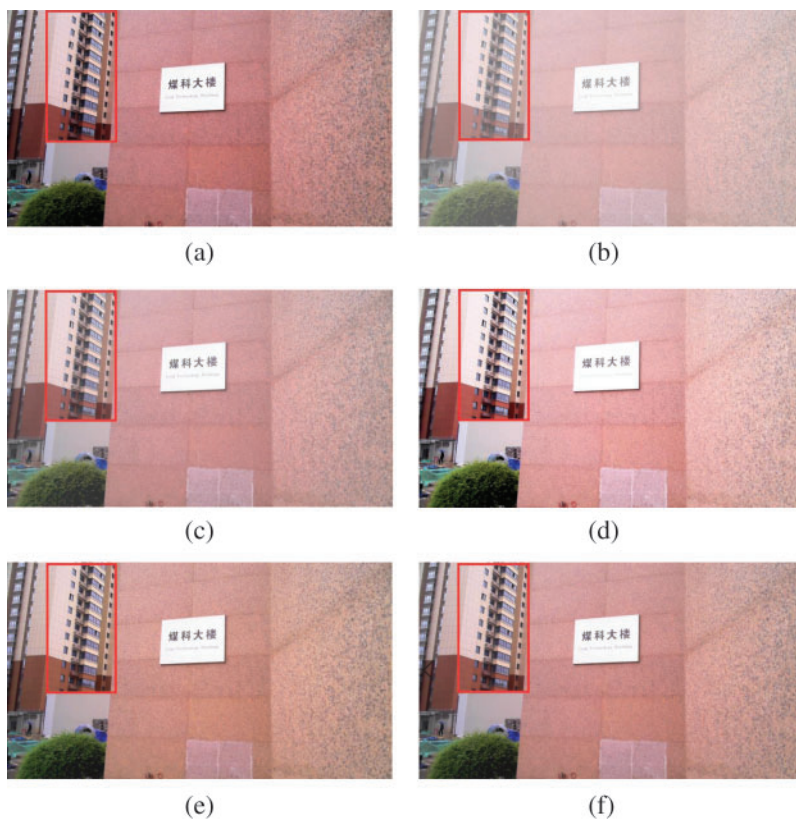
Figure 9: Comparison results of the dehazing of the building images. (a) Sharp image. (b) Real hazy image. (c) Tang's algorithm. (d) Cai's algorithm. (e) Li's algorithm. (f) Our approach

Table 1: Experimental results for the synthetic hazy road images

Indicator	Tang's algorithm	Cai's algorithm	Li's algorithm	Our approach
SSIM	0.8322	0.8522	0.9023	0.9128
PSNR	24.2644	26.5940	27.1845	27.8906
IE	7.2173	7.2802	7.3429	7.4018

**Table 2:** Experimental results for the synthetic hazy house images

Indicator	Tang's algorithm	Cai's algorithm	Li's algorithm	Our approach
SSIM	0.8452	0.8756	0.9048	0.9230
PSNR	24.3325	26.7622	27.9630	28.2588
IE	7.0518	7.2090	7.3281	7.4979

**Table 3:** Experimental results for the synthetic hazy tree images

Indicator	Tang's algorithm	Cai's algorithm	Li's algorithm	Our approach
SSIM	0.8472	0.8822	0.9123	0.9278
PSNR	25.2511	26.0771	26.6549	27.3971
IE	7.2811	7.3573	7.4412	7.5134

Table 4: Experimental results for the synthetic hazy fountain images

Indicator	Tang's algorithm	Cai's algorithm	Li's algorithm	Our approach
SSIM	0.8485	0.8863	0.9194	0.9283
PSNR	25.0496	26.2763	27.2319	27.4703
IE	7.0188	7.1977	7.3857	7.4603
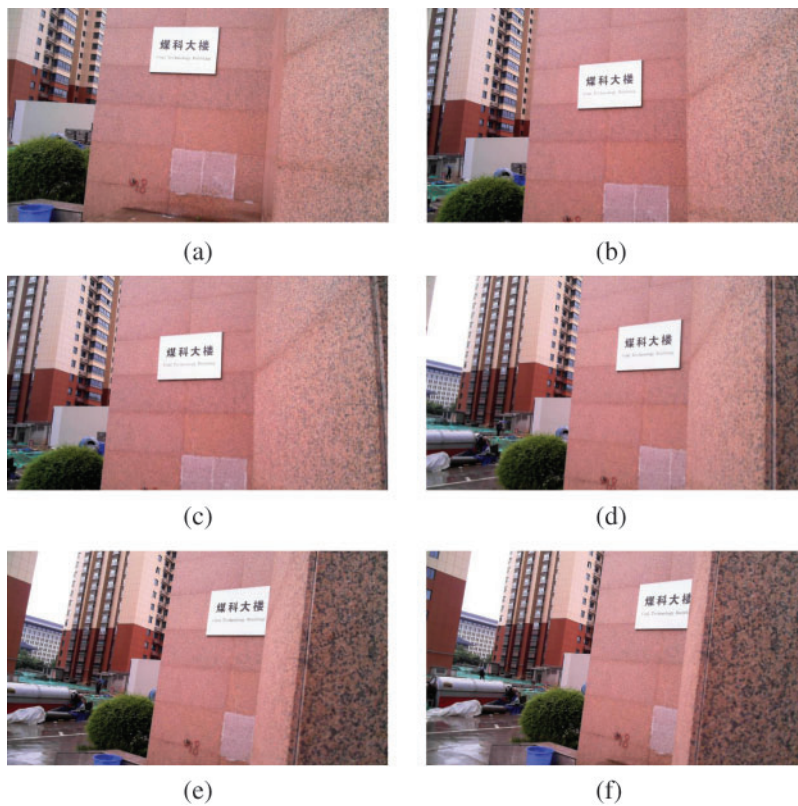
Figure 10: Six images of a building taken from different perspectives. (a) First perspective. (b) Second perspective. (c) Third perspective. (d) Fourth perspective. (e) Fifth perspective. (f) Sixth perspective
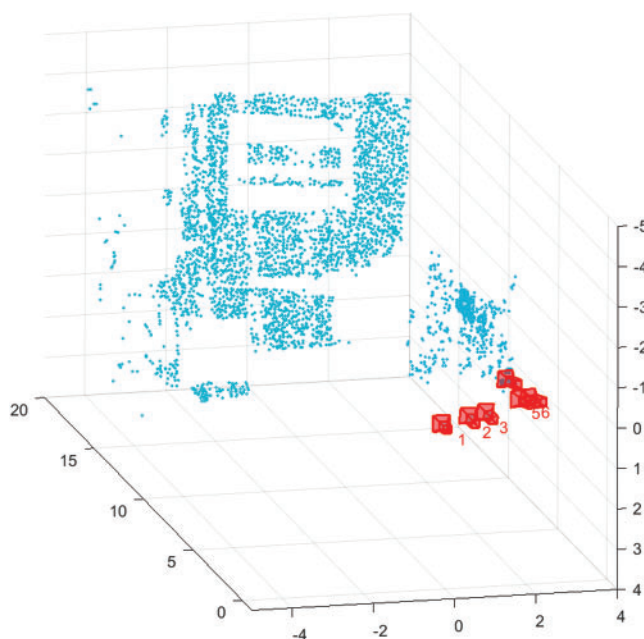
Figure 11: Point cloud of 3D reconstruction of building

Table 5: Comparison of the matching results of the building

Building	Hossain's algorithm	Zhang's algorithm	Zhou's algorithm	Our approach
Matching accuracy (%)	83	85	86	90
Match time (s)	15.69	14.77	13.56	10.83

4 Conclusion

AI solutions can provide great help for dehazing images, which can automatically identify patterns or monitor the environment. Therefore, we propose a 3D reconstruction method for dehazed images for smart cities based on deep learning. First, we propose an FT-DCRN dehazing algorithm that uses fine transmission images and atmospheric light values to compute dehazed images. The DCRN is used to obtain the coarse transmission image, which can not only expand the receptive field of the network, but can also retain the features to maintain the nonlinearity of the overall network. The fine transmission image is obtained by refining the coarse transmission image using a guided filter. The atmospheric light value is estimated according to the position and brightness of the pixels in the original hazy image. Second, we use the dehazed images generated by the FT-DCRN dehazing algorithm for 3D reconstruction. The ARI-SFM algorithm, which obtains the fine matching corner pairs and reduces the number of iterations, establishes an accurate one-to-one matching corner relationship. The experimental results show that our FT-DCRN dehazing algorithm improves the accuracy compared to other representative algorithms. In addition, the ARI-SFM algorithm guarantees the precision and improves the efficiency.

Developing AI systems supporting smart cities requires considerable data. Through the acquisition of effective information, smart cities can truly become sustainable developments. By 2021,

one billion smart cameras will be deployed in infrastructure and commercial buildings. The large amount of raw data collected is far beyond the scope that can be viewed, processed or analyzed manually. Through the machine learning training process, images can be analyzed for city planning and development. AI algorithms have become the developmental trend and key point of smart cities [35]; therefore, how to manage deep learning algorithms, data, software, hardware and services will become another problem in the future.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Z. Guo, K. Yu, Y. Li, G. Srivastava and J. C.-W. Lin, "Deep learning-embedded social internet of things for ambiguity-aware social recommendations," *IEEE Transactions on Network Science and Engineering*, pp. 1, 2021. https://doi.org/10.1109/TNSE.2021.3049262.

[2] K. Yu, L. Lin, M. Alazab, L. Tan and B. Gu, "Deep learning-based traffic safety solution for a mixture of autonomous and manual vehicles in a 5G-enabled intelligent transportation system," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2020. https://doi.org/10.1109/TITS.2020.3042504.

[3] L. Zhao, K. Yang, Z. Tan, X. Li, S. Sharma *et al.*, "A novel cost optimization strategy for SDN-enabled UAV-assisted vehicular computation offloading," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2020. https://doi.org/10.1109/TITS.2020.3024186.

[4] M. Y. Ju, Z. F. Gu and D. Y. Zhang, "Single image haze removal based on the improved atmospheric scattering model," *Neurocomputing*, vol. 260, no. 18, pp. 180–191, 2017.

[5] K. T. Tang, J. C. Yang and J. Wang, "Investigating haze-relevant features in a learning framework for image dehazing," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, Columbus, IEEE, pp. 2995–3002, 2014.

[6] B. Cai, X. M. Xu, K. Jia, C. Qing and D. C. Tao, "DehazeNet: An end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187–5198, 2016.

[7] B. Y. Li, X. L. Peng, Z. Y. Wang, J. Z. Xu, D. Feng *et al.,* "AOD-Net: All-in-one dehazing network," in *Int. Conf. on Computer Vision*, Venice, IEEE, pp. 4770–4778, 2017.

[8] Q. B. Wu, J. G. Zhang, W. Q. Ren and W. M. Zuo, "Accurate transmission estimation for removing haze and noise from a single image," *IEEE Transactions on Image Processing*, vol. 29, pp. 2583–2597, 2020.

[9] D. Park, D. K. Han and H. Ko, "Nighttime image dehazing using local atmospheric selection rule and weighted entropy for visible-light systems," *Optical Engineering*, vol. 56, no. 5, pp. 050501, 2017.

[10] M. Demant, P. Virtue, A. Kovvali, S. X. Yu and S. Rein, "Learning quality rating of as-cut mc-si wafers via convolutional regression networks," *IEEE Journal of Photovoltaics*, vol. 9, no. 4, pp. 1064–1072, 2019.

[11] Z. B. Wang, S. Wang and Y. Zhu, "Multi-focus image fusion based on the improved PCNN and guided filter," *Neural Processing Letters*, vol. 45, no. 1, pp. 75–94, 2017.

[12] G. L. Bi, J. Y. Ren, T. J. Fu, T. Nie, C. Z. Chen *et al.,* "Image dehazing based on accurate estimation of transmission in the atmospheric scattering model," *IEEE Photonics Journal*, vol. 9, no. 4, pp. 7802918, 2017.

[13] J. Zhang, K. P. Yu, Z. Wen and X. Qin, "3D reconstruction for motion blurred images using deep learning-based intelligent systems," *Computers, Materials & Continua*, vol. 66, no. 2, pp. 2087–2104, 2021.

[14] E. Casella, A. Collin, D. Harris, S. Ferse, S. Bejarano *et al.,* "Mapping coral reefs using consumer-grade drones and structure from motion photogrammetry techniques," *Coral Reefs*, vol. 36, no. 1, pp. 269–275, 2017.

[15] M. A. Hossain and A. K. Tushar, "Chord angle deviation using tangent (CADT), an efficient and robust contour-based corner detector," in *Proc. of IEEE Int. Conf. on Imaging, Vision & Pattern Recognition*, Dhaka, Bangladesh, IEEE, pp. 1–6, 2017.

[16] H. L. Zhang, Z. Gao, X. J. Zhang and K. F. Shi, "Image matching method combining hybrid simulated annealing and antlion optimizer," *Computer Science*, vol. 46, no. 6, pp. 328–333, 2019.

[17] W. Zhou, A. N. Bowen, M. Zhao and S. Pan, "Heterologous remote sensing image registration algorithm based on geometric invariance and local similarity features," *Infrared Technology*, vol. 41, no. 6, pp. 561–571, 2019.

[18] C. H. Yeh, C. H. Huang and L. W. Kang, "Multi-scale deep residual learning-based single image haze removal via image decomposition," *IEEE Transactions on Image Processing*, vol. 29, pp. 3153–3167, 2019.

[19] S. C. Raikwar and S. Tapaswi, "Lower bound on transmission using non-linear bounding function in single image dehazing," *IEEE Transactions on Image Processing*, vol. 29, pp. 4832–4847, 2020.

[20] H. Zhu, R. Vial and S. Lu, "TORNADO: A spatio-temporal convolutional regression network for video action proposal," in *2017 IEEE Int. Conf. on Computer Vision*, Venice, IEEE, pp. 5814–5822, 2017.

[21] L. Zhao, Y. Liu, A. Al-Dubai, Z. Tan, G. Min *et al.,* "A novel generation adversarial network-based vehicle trajectory prediction method for intelligent vehicular networks," *IEEE Internet of Things Journal*, vol. 8, no. 3, pp. 2066–2077, 2020.

[22] L. Qin, Y. Gong, T. Tang, Y. Wang and J. Jin, "Training deep nets with progressive batch normalization on multi-GPUs," *International Journal of Parallel Programming*, vol. 47, no. 3, pp. 373–387, 2019.

[23] H. Y. Jung and Y. S. Heo, "Fingerprint liveness map construction using convolutional neural network," *Electronics Letters*, vol. 54, no. 9, pp. 564–566, 2018.

[24] Y. F. He, A. Carass, Y. H. Liu, B. M. Jedynak, S. D. Solomon *et al.,* "Structured layer surface segmentation for retina OCT using fully convolutional regression networks," *Medical Image Analysis*, vol. 68, no. 136295, pp. 101856, 2020.

[25] L. Liu, C. Shen and A. V. D. Hengel, "Cross-convolutional-layer pooling for image recognition," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 39, no. 11, pp. 2305–2313, 2015.

[26] J. Yuan, H. C. Xiong, Y. Xiao, W. L. Guan, M. Wang *et al.,* "Gated CNN: Integrating multi-scale feature layers for object detection," *Pattern Recognition*, vol. 105, no. 6, pp. 107131, 2019.

[27] R. Xiang, X. Zhu, F. Wu and X. Y. Jiang, "Guided filter based on multikernel fusion," *Journal of Electronic Imaging*, vol. 26, no. 3, pp. 033027.1–033027.8, 2017.

[28] K. He, J. Sun and X. Tang, "Single image haze removal using dark channel prior," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, Miami, IEEE, pp. 1956–1963, 2009.

[29] S. Q. Dong, L. G. Han, Y. Hu and Y. C. Yin, "Full waveform inversion based on a local traveltime correction and zero-mean cross-correlation-based misfit function," *Acta Geophysica*, vol. 68, no. 1, pp. 29–50, 2020.

[30] L. Piermattei, L. Carturan, F. DBlasi and P. Tarolli, "Suitability of ground-based SFM-MVS for monitoring glacial and periglacial processes," *Earth Surface Dynamics*, vol. 4, no. 2, pp. 425–443, 2016.

[31] L. Bo, D. Yuchao and H. Mingyi, "Monocular depth estimation with hierarchical fusion of dilated CNNs and soft-weighted-sum inference," *Pattern Recognition*, vol. 83, pp. 328–339, 2017.

[32] T. S. Zhao, J. H. Wang, Z. Wang and C. W. Chen, "SSIM-based coarse-grain scalable video coding," *IEEE Transactions on Broadcasting*, vol. 61, no. 2, pp. 210–221, 2015.

[33] M. Malarvel, G. Sethumadhavan, P. C. R. Bhagi, S. Kar, T. Saravanan *et al.,* "Anisotropic diffusion based denoising on X-radiography images to detect weld defects," *Digital Signal Processing*, vol. 68, pp. 112–126, 2017.

[34] X. Zhang, D. Li, J. Li and Y. Li, "Magnetotelluric signal-noise separation using IE-LZC and MP," *Entropy*, vol. 21, no. 12, pp. 1190–1204, 2019.

[35] Z. Guo, K. Yu, A. Jolfaei, A. K. Bashir, A. O. Almagrabi *et al.,* "A fuzzy detection system for rumors through explainable adaptive learning," *IEEE Transactions on Fuzzy Systems*, pp. 1, 2021. https://doi.org/10.1109/TFUZZ.2021.3052109.