International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 3 March 2023, pp: 799-804 www.ijaem.net ISSN: 2395-5252

ResNet-CA: A Novel End-to-End Visibility Estimation Method

Yu Xin, Hao Peng
[1]Student, Chengdu University of Information Technology, ChengDu, China
[2]Student,Chengdu University of Information Technology, ChengDu, China

Date of Submission: 05-03-2023 Date of Acceptance: 15-03-2023

ABSTRACT: Low visibility always leads to serious traffic accidents worldwide, although extensive works are studied to deal with the estimation of visibility in meteorology areas, it is still a tough problem. As one of the basic indicators of atmospheric observation, the accurate estimation of visibility has an important impact on people's daily life and travel. To improve the accuracy of visibility estimation, we propose a visibility estimation network based on the fusion of deep learning and attention mechanisms: ResNet-CA. we use the learning ability of deep learning networks to learn the features for visibility estimation and subsequently feed the features to the classifier for classification. In addition, to improve the learning ability of the network, we embed the attention mechanism in the divine network and use transfer learning to guide the network for training. Also, to address the problem that there is no public dataset for visibility estimation, we construct a visibility estimation dataset for evaluating our method. According to the experimental results, ResNet-CA can effectively estimate visibility with better accuracy than some classical networks.

KEYWORDS: Transfer learning; Deep learning; ResNet-CA; visibility esitimation

I. INTRODUCTION

The International Commission on Illumination (International Com-mission on Illumination) defines visibility: the farthest distance that can be identified to an object without any help from the naked eye is called the current visibility distance [1]. And in meteorology it is usually defined as the maximum horizontal distance of a target object that can be seen and discerned from the sky background when a person with vision conditions within the normal range is in good weather conditions, and this distance can be objectively measured or expressed by meteorological optical formulas. Visibility is an important indicator of atmospheric transparency and is mainly influenced by the atmospheric extinction coefficient caused by various suspended particles in the atmosphere. Extreme visibility weather can greatly affect people's life and production, and although the current forecasting technology for weather conditions has made great progress, it is mainly for rain, sunny and cloudy weather conditions, while the forecasting and real-time detection of visibility is still more difficult [2]. In the early days, people often used instrumental and visual methods to estimate the visibility detection. The instrumentation method mainly detects visibility by some optical components, but such instruments often have high usage costs, while the visual inspection method has a large major factor, and the estimation results often have human errors. With the rapid development of image processing technology, researchers started to study image-driven visibility detection and estimation techniques. Early on, the visibility estimation was mainly based on theories such as atmospheric scattering model, etc. With the deployment and application of a large number of monitoring devices and the rapid development of CNN (Convolutional Neural Networks), researchers began to turn With the deployment of a large number of surveillance devices and the rapid development of CNN (Convolutional Neural Networks), researchers have started to shift their attention to big data-driven visibility estimation algorithms, which have the advantages of fewer constraints and higher accuracy than traditional image processing methods. Graves et al. combined local contrast with dark channel a priori theory to estimate the visibility of images [3], but the accuracy of the obtained transmittance was not high. Subsequently, He et al. added a guided filtering algorithm to refine the transmittance in the process of image transmittance estimation, which led to a more accurate visibility estimation [4]. The calibration requirements in this method are high and

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 3 March 2023, pp: 799-804 www.ijaem.net ISSN: 2395-5252

have a large deployment difficulty, which is not very practical. Mu and Wang et al. used CNN for visibility detection of images, but the methods of this type all apply deep learning directly to visibility estimation, and thus the accuracy rate needs to be improved [5, 6].Tang et al. proposed a CNN visibility estimation algorithm based on continuous video, which has the advantage of extracting the ROI region of the image into the network, but the method has a cumbersome process for ROI region extraction [7].Zhang J et al. used haze images and fog-free images of the same scene to jointly guide HazDesNet for training and used the structural similarity (SSIM) scores in both images as regression targets for detecting outdoor image haze concentration [8]. The biggest problem of this method is the difficulty of obtaining fog and fog-free image pairs in the same scene. The contributions of this paper are as follows:

1. propose a deep learning combined with attention mechanism for visibility estimation: ReNet-CA.
2. construct a visibility estimation dataset, which contains four levels of visibility data with a total of 1558 sheets.

II. METHODOLOGY

In this section, we will introduce the specific structure and detailed parts of ResNet-CA, along with migration learning, attention mechanism, and loss function.

A. Attention Mechanisms

Numerous experiments have shown that attention mechanisms can help networks achieve better results in most computer vision tasks.SENet as a classical attention mechanism module effectively builds dependencies between channels, but it tends to ignore location information, which is the key to generate spatially selectiveattention maps [9].Hou et al. introduce a new attention module called CA (Coordinate Attention), a novel attention module. Its structure diagram is shown in Figure 1.

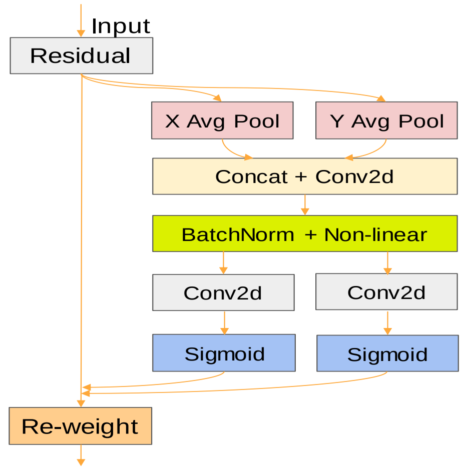

Figure. 1 Coordinate Attention

On the ImageNet dataset, CA has performed much better than SE and CBAM for the following reason: CA is not only able to use the obtained location information for more accurate regions of interest, but also capture the relationship between channels.

B. Transfer learning

Transfer learning [10] is commonly expressed as using an existing foundation to learn new knowledge. Specifically, we usually call the existing knowledge as the source domain and the knowledge to be learned as the target domain, and transfer learning is to transfer the knowledge from the source domain to the target domain as a way to help the target domain learn. Numerous studies have shown that transfer learning has achieved great success in the field of deep learning.

C. ResNet34

In this paper, we choose the modified ResNet34 model and apply transfer learning as a deep feature extractor for visibility images.ResNet was proposed by Kaiming He and four others from Microsoft Research, and its main idea is to add directly connected channels in the network to allow the original input information to directly access the deeper network layers, and its basic structure is shown in Fig. 2. ResNet34 is constructed by superimposing residual blocks, and its structure is shown in Table 1.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 3 March 2023, pp: 799-804 www.ijaem.net ISSN: 2395-5252

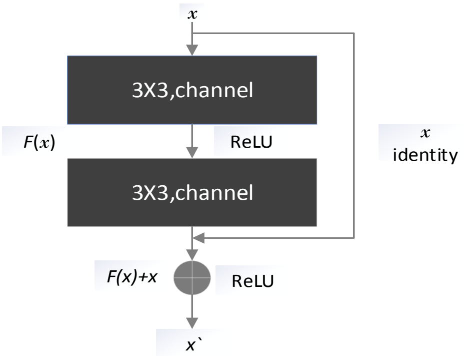

Figure. 2 Residual Basic-Block

Table 1.The structure of ResNet-34

Layer name	Output size	34-layer
Conv1	112 × 112	7×7, 64, stride 2
Conv2_x	56 × 56	3×3 max pool, stride 2 $\begin{bmatrix} 3\times3, & 64 \\ 3\times3, & 64 \end{bmatrix} \times 2$
Conv3_x	28 × 28	$\begin{bmatrix} 3\times3, & 128 \\ 3\times3, & 128 \end{bmatrix} \times 4$
Conv4_x	14 × 14	$\begin{bmatrix} 3\times3, & 256 \\ 3\times3, & 256 \end{bmatrix} \times 6$
Conv5_x	7 ×7	$\begin{bmatrix} 3\times3, & 512 \\ 3\times3, & 512 \end{bmatrix} \times 3$

Traditional CNNs have more or less gradient disappearance, gradient explosion, information loss, and loss in information transfer [16]. When training deep networks, these problems can lead to "degradation" in the training of deep networks. The problem is solved to some extent by the emergence of the shortcut structure in ResNet, where the network extracts deeper semantic information from the image as the number of layers increases, but also loses some of the shallow information.

D. ResNet-CA

To improve the accuracy of visibility estimation, we propose a visibility estimation network based on the fusion of deep learning and attention mechanisms: ResNet-CA. we use the learning ability of deep learning networks to learn the features for visibility estimation and subsequently feed the features to the classifier for classification. In addition, to improve the learning ability of the network, we embed the attention mechanism in the divine network and use transfer learning to guide the network for training. The structure diagram of ResNet-CA is shown in Figure 3.

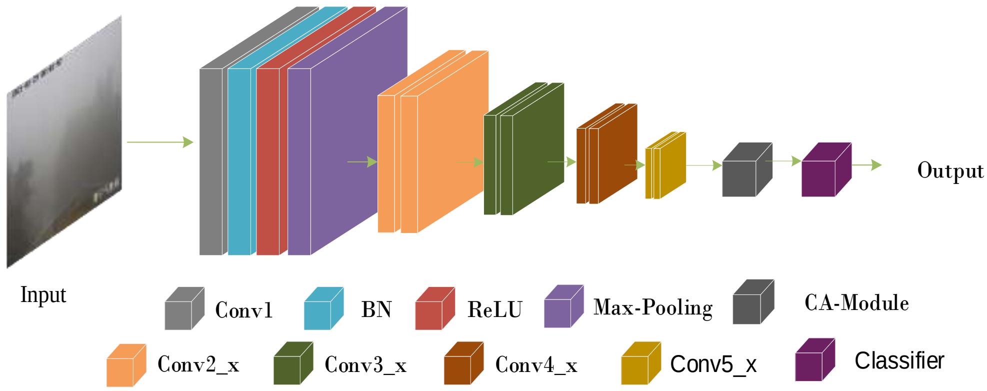

Figure 3 The structure diagram of ResNet-CA

III. EXPERIMENTAL RESULTS AND ANALYSIS

A. Experimental environment

The experiments were conducted under Intel(R) Xeon(R) Gold 5218R CPU @ 2.10GHz, 256G RAM, Ubuntu 20.04.1 LTS 64-bit OS, NVIDIA GeForce RTX 3080, python 3.6, OpenCV 3.4.0, Pytorch 1.9.0, CUDA 11.1 environment. CUDA 11.1 environment. the training and test sets of MLID are divided in the ratio of 8 : 2, in the deep feature extraction part epoch is 100, the learning rate is set to 0.001, the learning decay rate is 0.01, and the optimizer is Adam [25]. In experiments 4.3 to 4.6 the weights were not frozen in any of the model training, and the loss function was used in Cross Entropy. as shown in the formula (1)[11].

$$loss = -\sum_{i=1}^{N} \hat{y}_i * Log(y_i) \quad (1)$$

Where is the real label and is the predicted label.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 3 March 2023, pp: 799-804 www.ijaem.net ISSN: 2395-5252

B. Evaluation Metrics

The confusion matrix is often used as a visualization tool and evaluation index in deep learning, including precision, recall, accuracy, and specificity. A confusion matrix is used to evaluate the proposed model in this paper. Equations (2)~(5) show their calculation methods[12].

$$precision = \frac{TP}{TP+FP} \quad (2)$$

$$recall = \frac{TP}{TP+FN} \quad (3)$$

$$specificity = \frac{TN}{TN+FP} \quad (4)$$

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (5)$$

where TP, FP, TN, and FN are separately presented as true positive,false positive, true negative, and false negative.

C. Datasets

The datasets of FRIDA1 and FRIDA2 are often used for the research of image dehazing, which contains 420com-posites in 84different scenes and depth information is given for each scene[21,22]. We use the atmospheric scattering model, next change the value of β then combine the corre-sponding depth information to generate 1558 foggy images with different visibility values, last use these images to construct a dataset named VID Ⅰ. Figure 4 shows images of different visibility levels in VID Ⅰ

(a) (b) (c) (d)

Figure 4. Images with different visibility levels

With reference to the national criteria for the classification of visibility and the situation of data collection, the data ofvisibility values are divided into four levels:50m~500m, 500m-1km, 1km-10km, and over10km.

D. Comparison experiment

In order to verify the effectiveness and superiority of the method proposed in this paper, we design an experiment to compare the accuracy of ResNet-CA and other classical methods such as ResNet18,DenseNet121[14,15], Resnet34, Resnet 50[16] and AlexNet on VID I. Meanwhile, At the same time, we compare the recall, precision, and specificity of STCN-Net and other classical methods on the two datasets. The results are shown in Tables 2 and Tables 3.

Table 2.Summary of comparison test results

Method	accuracy
AlexNet	93.2%
Resnet18	93.4%
Resnet34	94.9%
Resnet50	93.1%
DenseNet121	92.4%
GooleNet	89.4%
ResNet-CA(our method)	96.8%

Table 3.Summary of comparison test results

Method	recall	precision	specificity
AlexNet	94.2%	95.2%	95.6%
Resnet18	95.6%	96.3%	96.3%

Resnet34	96.7%	97.3%	98.1%
Resnet50	94.6%	96.2%	97.1%
DenseNet121	96.2%	95.3%	97.3%
GooleNet	93.5%	94.1%	97.4%
ResNet-CA(our method)	98.2%	97.6%	98.6%

Although AlexNet is a representative model in CNN, it works in a mediocre performance with accuracy on both datasets. Meanwhile, because AlexNet only adopts the basic structures of CNN[17], there will cause an"information disappearing" pro-blem with the deepening of the network.. Therefore the residual structure proposed by ResNet improves the problem of "information disappearing" effectively. And it brings a higher accuracy of visibility estimation compared with AlexNet. DenseNet inherits the idea of ResNet, but it builds a dense connection between all the previous layers and the following layers.ResNet-CA outperforms the other classical methods on both datasets, especially in the real-scene dataset VID I, and compared with other methods, its accuracy of visibility estimation reaches 96.8%, which is a significant improvement.

E. Relevant data and results display

We use our proposed method for visibility estimation of some scenes and find that our method can all predict them accurately, and we show some of the results in Figure 5.

Figure 5.Relevant data and results display

IV. CONLUSION

In this paper, we propose a deep learning-based visibility level classification model to classify the visibility level of an image, expressing both shallow and deep semantic information of the image, fully considering the problem of shallow information loss caused by the deep convolution process of the network too deep and trying to solve it, and finally achieving a better classification accuracy of visibility level. At the same time, compared with other traditional visibility detection methods, the method is more convenient and faster, and less costly, and can be used for long time visibility detection. In the future, the method will be improved and its accuracy will be further improved.

REFERENCES

[1]. Dumont E , Cavallo V. Extended Photometric Model of Fog Effects on Road Vision[J]. Transportation Research Record, SAGE Publications Inc, 2004, 1862(1) : 77–81.

[2]. Holmstrom M, Liu D, Vo C. Machine learning applied to weather forecasting[J]. Meteorol. Appl, 2016, 10: 1-5.

[3]. N. Graves and S. Newsam, "Using visibility cameras to estimate atmospheric light extinction , " 2011 IEEE Workshop on Applications of Computer Vision (WACV),

2011 , pp. 577-584 , doi: 10.1109/WACV.2011.5711556.

[4]. Y. He, J. Ding, H. Teng, X. Han, Y. Chen and W. Zhou, "Visibility detection and prediction of foggy highway based on lane line detection and Winters additive model , " 2021 40th Chinese Control Conference (CCC), 2021, pp. 7254-7259 , doi : 10.23919/CCC52363.2021.9549291.

[5]. M. Xiyu, X. Qi, Z. Qiang, R. Junch, W. Hongbin and Z. Linyi, "An Improved Diracnet Convolutional Neural Network for Haze Visibility Detection, " 2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP), 2021, pp. 1-5 , doi : 10.1109/MLSP52302.2021.9596249.

[6]. J. Wang and L. Zhang, "Research on Deep Learning Model of Fog Visibility Estimation Based on CNN," 2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2021, pp. 1355-1359, doi: 10.1109/IMCEC51613.2021.9482258.

[7]. [8] Tang P , Song S , Zhao T , et al. Visibility estimation of foggy weather based on continuous video information[J]. Remote Sensing Letters, 2021, 12(10):1061-1072.

[8]. Zhang J , Min X , Zhu Y , et al. HazDesNet: An End-to-End Network for Haze Density Prediction [J]. IEEE Trans on Intelligent Transportation Systems, 2020, online.

[9]. Qin H, Qin H. Image-Based Dedicated Methods of Night Traffic Visibility Estimation[J]. Applied Sciences-Basel, Basel: Mdpi, 2020, 10(2): 440.

[10]. GUO ZQ,HU Y W,LIU P,et al.Outdoor weather image clas-sification based on feature fusion[J.Journal of Computer Ap-plications,2020,40(4):1023-1029.

[11]. JIN L S,CHEN M,JIANG Y Y,et al.Multi-Traffic Scene Perception Based on Supervised Learning[J].IEEE Access,2018, 6:4287-4296.

[12]. ZHANG Z,MA H D,FU H Y,et al.Scene-free multi-classweather classification on single images[J]. Neurocomputing , 2016,207(26):365-373.

[13]. HE K M,SUN J, TANG X O.Single Image Haze RemovalUsing Dark Channel Prior[J].IEEE Transactions on PatternAnalysis and Machine Intelligence,2011,33(12):2341-2353.

[14]. ARBANE M,BENLAMRI R,BRIK Y,et al. Transfer Learningfor Automatic Brain Tumor Classification Using MRI Images[C]// 2020 2nd International Workshop on Human-CentricSmart Environments for Health and Well-being(IHSH).Boumerdes , Algeria,2021:210-214.

[15]. Chakraborty S, Mondal R, Singh P K, et al. Transfer learning with fine tuning for human action recognition from still images[J]. Multimedia Tools and Applications, 2021(3).

[16]. He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[A]. 2016: 770–778.

[17]. LIU W B,ZOU Z Y,XING w W. Feature Fusion Methods inPattern Classification[J.Journal of Beijing University of Postsand Telecommunications,2017,40(4):1-8.