Research on Visibility Estimation Model Based on DenseNet

Guang Li
Department of Launch and Power Engineering,
Rocket Officer Academy
Wei'fang, China
Email: 865413666@qq.com

Zhiqiang Chang
Department of Launch and Power Engineering,
Rocket Officer Academy
Wei'fang, China
Email: 865413666@qq.com

***Abstract*—In recent years, the road visibility detection method based on video has been paid more and more attention. It has overcome the deficiency of laser visibility meter to some extent. Deep learning has a good effect in image processing and analysis. This paper firstly analyzes the current situation of deep learning, and then compares DenseNet and ResNet to propose a visibility estimation model based on deep DenseNet. The model firstly integrates airport video data and visibility data. Secondly, the DenseNet algorithm is used to automatically extract the features of the airport data set. Finally, Softmax classifier is constructed to evaluate the visibility accuracy. They reduce the problem of disappearing gradient, enhance feature propagation, encourage functional reuse, and greatly reduce the number of parameters, well train the deep model, has a good visibility estimation effect. On this basis, this paper based on Canny operator lane dividing line extraction edge extraction and visibility analysis based on edge detection, and do the corresponding test. Finally, a video visibility analysis model based on Kalman filter is built based on the given data, and Gaussian process regression model is used to predict the fog change trend.**

Keywords-Densenet; Visibility Estimation; Softmax Classifier

I. INTRODUCTION

A. Background

Visibility is a common indicator in meteorology, highway traffic and aircraft flight. The unit is usually meters. The main factors affecting visibility are fog and haze. As we all know, visibility is very important to highway safety. When visibility is very low, highway managers usually close the road in order to drive safety. However, in the aviation field, runway visibility is used to reflect the size of fog and haze near the airport, which is defined as the maximum distance at one end of the runway along the direction of the runway to identify the runway or the target object close to the runway (runway side lights at night). Normally, flights are banned when visibility at the airport is only about 400 meters. When visibility at the airport is only about 600 to 800 meters, flights can take off and land normally. However, due to security concerns, the airport will take temporary measures to control the flight flow, widening the flight departure interval, which is easy to cause flight delays. Therefore, visibility prediction is the highway management department and airlines are very concerned about the problem.

In recent years, video-based road condition (runway) visibility detection method [1,2] has attracted people's attention, which to some extent

overcomes the deficiency of laser visibility meter. Video visibility detection method is a combination of atmospheric optical analysis, image processing and artificial intelligence technology, through the analysis and processing of video images, establish the relationship between the video image and the real scene, and then calculate the visibility value indirectly according to the changes of image features. Wang Yaxue et al [3] established estimation and prediction models for visibility under different fog conditions. Through the establishment of multiple classification of multiple regression model. Han Huihui [4] introduced the ideas of closed-loop control and transfer learning into deep convolutional neural networks and designed a model of intelligent fog classification and visibility estimation based on deep transfer learning. Liu Jianlei et al. [5] analyzed the characteristics of anisotropy, continuity and horizontality of the inflection point line. Then, an inflection line detection filter is constructed according to these characteristics to improve the accuracy and speed of the detection. With the application of machine learning in various fields, the strong feature learning ability of deep learning network model has been applied to the classification and prediction of meteorological information, which has played a great role in promoting the improvement of forecast accuracy. In foreign countries, the artificial neural network model was initially used to forecast the visibility of winter weather in Milan, Italy [6], and the effect was better than the traditional method. By extending the standard probabilistic neural network method, Bremnes et al. [7] used multiple models to obtain an average output to improve the visibility forecast. Marzban et al. [8] combined the output of the numerical model with the ground observation data, and mapped the data to 39 different neural networks. The results showed that the performance of neural networks was generally due to logistic regression and MOS model. In China, Zhou Yongjiang et al. [9] used meteorological parameters such as temperature, air pressure, atmospheric precipitable water volume and PM 2.5 data of the same period to establish a haze prediction model integrating time series network and regression network to detect and forecast haze weather, proving that the fusion network model with meteorological parameters is more adaptable and accurate than the single network model.

In order to solve the visibility prediction problem in foggy weather, this paper constructs the visibility estimation model mainly through the method based on deep learning. The following is the introduction of deep learning-related issues and the main work of this paper.

B. Research status of deep learning technology

The concept of deep learning stems from research on artificial neural networks. Deep learning is a feature learning method that transforms raw data into a higher level and more abstract representation through some simple but nonlinear models. A more abstract high-level representation (attribute category or feature) is formed by combining low-level features to discover distributed feature representation of data [9]. In the process of learning the internal rules and representation hierarchy of sample data, deep learning has the ability to analyze, learn and interpret the acquired text, image and sound data, making it achieve remarkable achievements in image recognition, speech recognition, natural language understanding, weather prediction, gene expression, content recommendation and other aspects. Nowadays, deep learning plays a crucial role in various tasks of image analysis by discovering or learning information features that describe the inherent laws or patterns of data.

C. The Comparison between DenseNet and ResNet

In the field of computer vision, convolutional neural network (CNN) has become the most mainstream method. A milestone event in the history of CNN is the emergence of ResNet model, which can train deeper CNN model and achieve higher accuracy. The core of the ResNet model is to train a deeper CNN network by establishing a "short-circuit connection" between the front layer and the back layer, which facilitates the back propagation of the gradient during the training process. The DenseNet(Densely connected convolutional networks) model has the same basic idea as ResNet, but it establishes the dense connection between all the front layers and the back layers. Another feature of DenseNet is feature reuse through the concatenation of features on channels. These features allow DenseNet to achieve better performance than ResNet with fewer parameters and computational costs,

II. Related Work

Based on the video data and visibility data of an airport, this paper establishes a deep learning model for visibility estimation based on video data. The main work is to carry out data fusion of the video data and visibility data of the airport. Secondly, the convolutional network DenseNet algorithm is used to carry out automatic feature extraction of the merged airport data set. Softmax classifier was built for visibility and accuracy evaluation.

Data fusion technology refers to the information processing technology that utilizes the computer to analyze and synthesize the observed information obtained in time order under certain criteria to complete the required decision-making and evaluation tasks. Data layer fusion: fusion is carried out directly on the collected original data layer, and data synthesis and analysis are carried out before the original measurements of various sensors are processed. Data layer fusion generally adopts centralized fusion system for fusion processing

A. Camera calibration

In this paper, the mapping relationship between image coordinates and road coordinates is established for the video images collected by the camera, and the image distance information is converted into road distance information. The mapping relationship between image coordinates and road coordinates is completed by camera self-calibration technology, and its working steps are as follows:

1) Establish the road condition camera imaging model, as shown in Figure 3, in which three coordinate systems are defined, among which the ground coordinate system XW-YW-ZW and the camera coordinate system XC-YC-ZC are used to represent the three-dimensional space; Image plane coordinate system Xf -- Yf represents the imaging plane. The world coordinate system is established, whose origin is the intersection point of the camera optical axis and the ground. The camera coordinate system is established, and the origin is the camera optical center position. Set the distance between the optical center of the camera and the origin of the world coordinate system as l, the pitch Angle of the camera as t, the declination Angle as p, and the rotation Angle as s, and the area between parallel lines on the ground plane corresponds to the highway pavement in the camera's view field.

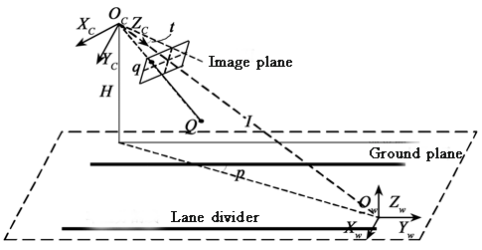


Figure 1. Highway camera model

2) Based on the defined camera spatial azimuth parameter [9], the coordinate transformation relationship between the ground coordinate system and the two-dimensional image coordinate system under the ideal penetration model can be established, as shown in Equations (1) and (2).

$$X_w = \frac{\left[l\sin p(x_f \sin s + y_f \cos s) + l\cos p\sin t(x_f \cos s + y_f \sin s)\right]}{x_f \cos t\sin s + y_f \cos t\cos s + f\sin t} \quad (1)$$

$$Y_w = \frac{\left[-l\cos p(x_f \sin s + y_f \cos s) + l\sin p\sin t(x_f \cos s - y_f \sin s)\right]}{x_f \cos t\sin s + y_f \cos t\cos s + f\sin t} \quad (2)$$

3) Parameter solution. The corresponding relationship between uncalibrated camera parameters and image feature parameters is established using highway lane lines as reference. The parallelogram based on the corner of the divider is used as the calibration block. According to the parallel correspondence between corner points, unknown camera parameters p, t, s, f, l in Equations (8) and (9) can be solved, as shown in Equations (1) and (2).

$$t = -\arcsin\left\{\frac{v_0^{\,2}(V_A - V_B + V_C - V_D)}{A \times B}\right\} \quad (3)$$

Among them:

$$A = \left[(V_0 - V_D)u_A - (V_0 - V_C)u_B + (V_0 - V_B)u_C - (V_0 - V_A)u_D\right] \quad (4)$$

$$B = \left[-(V_0 - V_B)u_A + (V_0 - V_A)u_B - (V_0 - V_D)u_C + (V_0 - V_C)u_D\right] \quad (5)$$

$$\left.\begin{array}{l} f = V_0 / \tan(t), l = \dfrac{H}{\sin t}, \tan s = -\dfrac{V_0 - V_1}{u_0 - u_1} \\ p = \arctan\left\{\dfrac{C * \sin(t)}{V_0(V_A - V_B + V_C - V_D)}\right\} \end{array}\right\} \quad (6)$$

Among them:

$$C = \left[(V_0 - V_D)u_A + (V_0 - V_C)u_B + (V_0 - V_B)u_C - (V_0 - V_A)u_D\right] \quad (7)$$

Where: (u,v) is the image coordinate system, (u0,v0) is the extinction point determined by the lines xaxd and xbxc, (u1,v1) is the extinction point determined by the lines xaxb and xdxc ,H camera column height.

B. Description of DenseNet network

Further, more accurate and effective training can be performed if the convolutional network contains shorter connections between the input layer and the output layer. In this paper, we accept this observation and introduce a dense convolutional network (DenseNet) that connects each layer to each other in a feedforward manner. A traditional convolutional network with L layers has L connections -- one connection between each layer and the one that follows -- and our network has L (L + 1) / 2 direct connections. For each layer, the feature maps of all previous layers are used as inputs, and its own feature maps are used as inputs for all subsequent layers. DenseNets have several compelling advantages: they alleviate the problem of disappearing gradients, enhance feature propagation, encourage functional reuse, and dramatically reduce the number of parameters.

The structure of the DenseNet network is shown in the figure 2. Each layer of the network is connected to all subsequent layers in a dense connection way.

C. Network backpropagation

In the training of DenseNet model, error information of training samples can be transmitted Back to the hidden layer through Back Propagation (BP) algorithm to realize constant updating and iteration of weight matrix between hidden layers until network convergence.

BP network is a feedforward neural network, which was proposed in 1985. In this kind of network, there are two kinds of signals flowing: one is the working signal, which is the input signal

applied and propagated forward until the actual output is generated at the output end of the signal, is a function of the input and weight; The second is the error signal. The difference between the actual network output and the theoretical output is the error, which starts from the output end and propagates back layer by layer. The output of the JTH unit at the output end in the NTH iteration is $y_j(n)$, then the error signal of the unit is $e_j(n) = d_j(n) - y_j(n)$, and the square error of unit j is defined as $\frac{1}{2} y_j^2(n)$. The instantaneous value of the total square error of the output is:

$$\varepsilon(n) = \frac{1}{2} \sum_j \mathrm{e}_j^2(n) \tag{8}$$

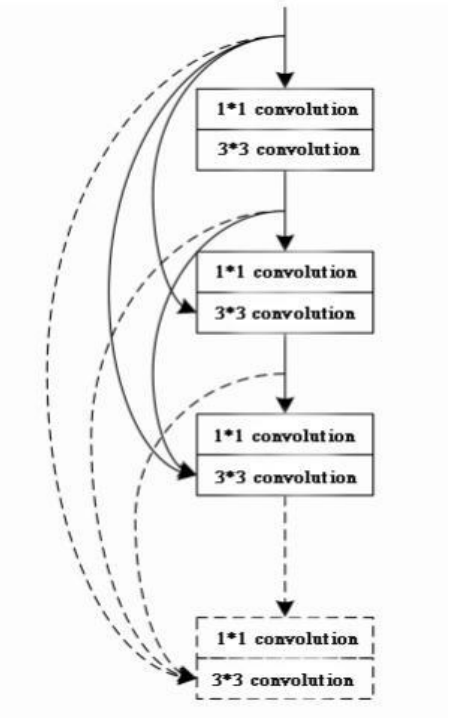


Figure 2. Schematic diagram of structural network

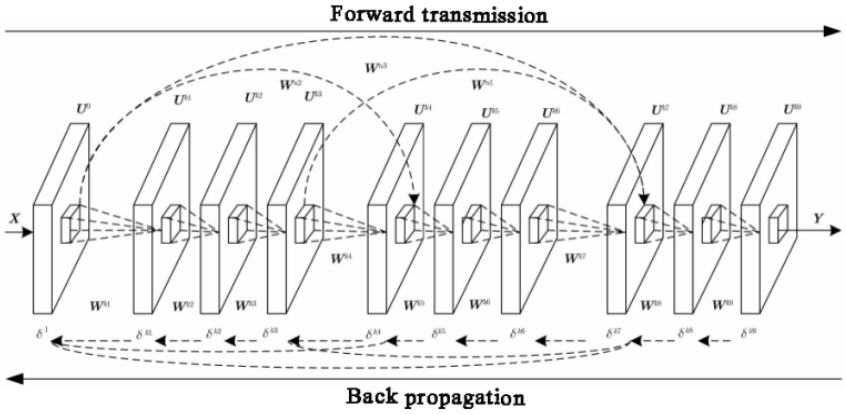


Figure 3. Training of DenseNet

Densenet can not only make efficient use of multi-dimensional feature information through the "feature re-calibration" strategy, but also slow down the attenuation of error terms of each hidden layer through the reverse transmission mechanism of the network itself, ensure the stability of gradient weight information, and enhance the learning and expression ability of the deep network, so as to further improve the network performance.

The data set used in this paper is the airport video and airport AMOS observations provided by the 2020 Modeling Competition. Firstly, the original airport video information and meteorological data were cleaned. Using python data analysis technology to carry out data fusion

of features; In addition, when using DenseNet for feature extraction, direct input of disordered discrete data will mislead the internal training mechanism of the network and confuse the importance of features. Discrete data is encoded by assigning values in turn according to the frequency of variable occurrence. For continuous data, feature values are mapped to [0,1] by minmax normalization to remove dimensional differences between features.

The classification prediction module mainly consists of three layers of network structure: global pooling layer, full connection layer and Softmax classifier. Through the combined operation of the three-layer network structure, the n×n×k dimension size matrix is converted into the probability corresponding to the occurrence of each category. Figure 4 below shows the integrated visibility analysis Organization.

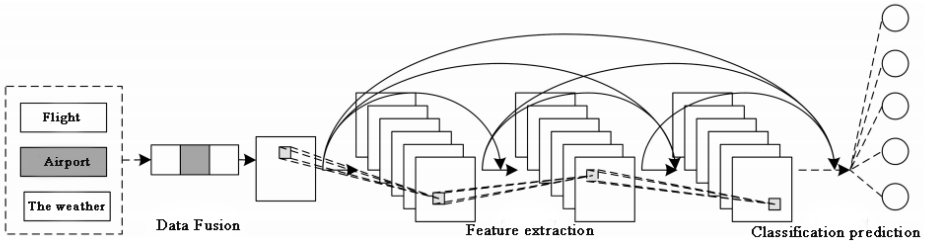


Figure 4. Airport visibility analysis structure

III. ANALYSIS OF EXPERIMENTAL RESULTS

The airport video used in the experiment in this paper has a total of 994379 samples of image data captured by python. The visibility of the human eye is 3000 meters, divided into 27 categories.

The network model is shown in the figure 5.

Classification accuracy: The depth network verifies the degree of model fitting mainly through the final accuracy of the test set, that is, the ratio of the sum of all correctly predicted flight data to the total number of test sets, expressed as:

$$\omega = \sum_{i=0}^{i=q} TP_i / N \tag{9}$$

Where, N is the total number of test sets, q is the number of visibility levels, the maximum is 27, TP_i represents the number of correct predictions in visibility levels.

Model: "dense_net"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	multiple	3200
batch_normalization (BatchNo	multiple	256
max_pooling2d (MaxPooling2D)	multiple	0
dense_block (DenseBlock)	multiple	209280
transition_layer (Transition	multiple	19296
dense_block_1 (DenseBlock)	multiple	226176
transition_layer_1 (Transiti	multiple	26096
dense_block_2 (DenseBlock)	multiple	234624
transition_layer_2 (Transiti	multiple	29880
dense_block_3 (DenseBlock)	multiple	238848
global_average_pooling2d (Gl	multiple	0
dense (Dense)	multiple	6723

Total params: 994,379
Trainable params: 984,171
Non-trainable params: 10,208

Figure 5. Network model

TABLE I. ACCURACY AND LOSS VALUES OF DENSENET

NUMBER OF NETWORK LAYERS	DENSENET(%)	LOSS
17	94.14	0.1459

Generally, the network judges whether the model finally converges by the loss value of the training set. The smaller the loss value is, the closer the predicted value of the model is to the real value. With the increase of the number of iterations, the model gradually converges and the loss value finally decreases to a fixed range. The loss value of DenseNet is reduced to about 0.15, and DenseNet can train the deep model well.

IV. FOLLOW-UP WORK

The video-based visibility detection method of road condition (runway) has attracted more and more attention. It overcomes the deficiency of laser visibility instrument to some extent. Video visibility detection method is a combination of atmospheric optical analysis, image processing and artificial intelligence technology, through the analysis and processing of video images, establish the relationship between the video image and the real scene, and then calculate the visibility value indirectly according to the changes of image features. In the follow-up work, we mainly focus on the edge extraction of lane divider based on Canny operator and the visibility analysis of single image based on edge detection, and do corresponding tests. Finally, we make the video visibility analysis model based on Kalman filter on the given data. In addition, a video provided by the mathematical modeling topic was used to draw the curve of the highway visibility changing with time in this period.

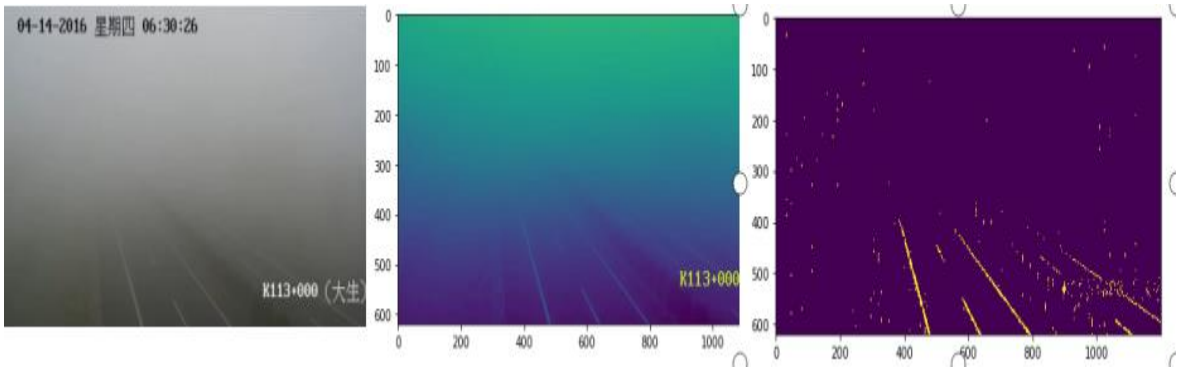


Figure 6. Canny operator edge extraction results

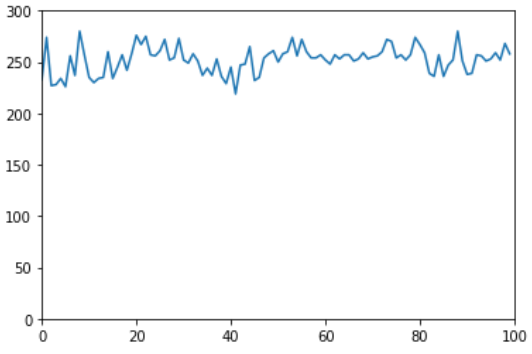


Figure 7. Highway visibility curve with time

Finally, a mathematical model is established to predict the change trend (worsening or weakening) and when the fog will disperse (reaching the specified visibility, such as MOR=150m). In this paper, Gaussian process regression model is proposed to predict the change trend of fog. Because Gauss process (GP) is a general supervised learning method designed to solve regression and probabilistic classification problems.

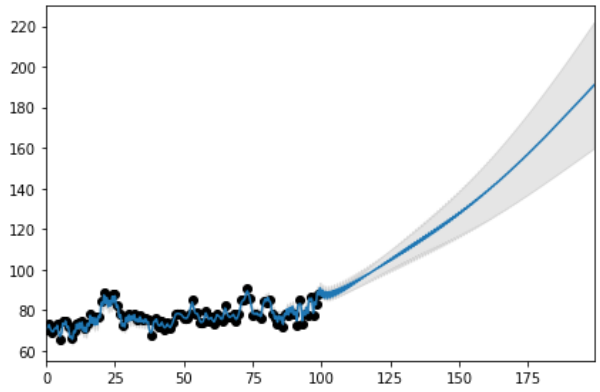


Figure 8. Gaussian process regression

V. CONCLUSION

In recent years, the video-based road condition (runway) visibility detection method has been paid more and more attention. It has overcome the deficiency of laser visibility meter to some extent. In fact, fogs form and dissipate in their own way, often related to near-ground meteorological factors. Based on the video data and visibility data of an airport, this paper establishes a deep learning model for visibility estimation based on video data. They reduce the problem of disappearing gradient, enhance feature propagation, encourage functional reuse, and greatly reduce the number of parameters. The loss value of DenseNet is reduced to about 0.15, and the deep model is well trained. With good visibility estimation effect, the visibility estimation model is used to forecast the dissipation of fog.

With the higher and higher level of numerical prediction, the accuracy of weather forecast is also getting higher and higher. Visibility estimation based on deep learning can be used to predict the dissipation of fog in the future, thus improving the transformation of weather to decision-making information and improving service quality.

REFERENCES

[1] S. K. Nayar, S. G. Narasimhan, Vision in bad weather, ICCV '99.

[2] 2.R. T. Tan, Visibility in bad weather from a single image, CVPR, 2008

[3] Wang Yaxue, Li Cheng, Liu Ting. Visibility Estimation and Prediction under Fog -- Based on the 2020 Graduate Mathematical Contest in Modeling E [J]. 2021.

[4] Han H H. Research on intelligent fog classification and visibility estimation method based on deep transfer learning.

[5] Liu Jianlei, Liu Xiaoliang. Fog visibility detection algorithm based on inflection line [J]. Journal of Computer Applications, 2015, 35(2):528-530.

[6] Pasini, A., Pelino, V. and Potesta, S. (2001) A Neural Network Model for Visibility Nowcasting from Surface Observations: Results and Sensitivity to Physical Input Variables. Journal of Geophysical Research, 106, 14951-14959. https://doi.org/10.1029/2001JD900134

[7] Bremnes, J.B. and Michaelides, S.C. (2007) Probabilistic Visibility Forecasting Using Neural Networks. Pure & Applied Geophysics, 164, 1365-1381.

[8] Marzban, C., Leyton, S. and Colman, B. (2006) Ceiling and Visibility Forecasts via Neural Networks. Weather & Forecasting, 22, 466-479. https://doi.org/10.1175/WAF994.1

[9] Zhou Yongjiang, Yao Yibin, Yan Xiao, et al. Prediction of haze by BP neural network based on GNSS meteorological parameters [J]. JGG, 2019, 39(11): 1148-1152.

[10] Li Bo, Dong Rong, Chen Qimei. Video Contrast Road Visibility Detection without Manual Marking [J]. Journal of Computer Aided Design & Computer Graphics,2009,21(11):1575-1582. (in Chinese)

[11] Xiao Hong, Cao Maojun, LI Panchi, WANG Haiying, Process Neural network Training Based on Piecewise Linear Interpolation [J]. Computer Engineering, 37(20) : 211-212, 215,2011.

[12] Christos Sakaridis · Dengxin Dai · Luc Van Gool.Semantic Foggy Scene Understanding with Synthetic Data, International J. Computer Vision, 2018, 3

[13] N. Hautiere, J-P Tarel, J Lavenant D. Aubert, Automatic fog detection and estimation of visibility distance through use of onboard camera, Machine Vision and Application, 2006, 17（1）: 8-20

[14] Gao Huang,Zhuang Liu,Laurens van der Maaten.Densely Connected Convolutional Networks.