Video Logging Casing Damage Image Recognition Based on Improved Convolutional Neural Network

Hongtao Hu, Yiyao Cheng
School of Computer Science Xi’an Shiyou University, Xi’an, China
huhongtao@xsyu.edu.cn

Abstract. Oil casing damage detection is the key point to ensure the smooth production of oil fields. In recent years, the automatic image recognition technology based on deep learning has become a researchful hot topic. But the common deep learning models have some defects in identifying the target features of casing damage images in the complex environment. This paper proposes an oil casing damage image recognition model based on DS-CNN(deep and shallow convolutional neural network). Based on VGG19, this model integrates the shallow convolution neural network. It combines global features extracted by the shallow network and the local features extracted by the deep network to form the input of the fully connected layer. The joint training of the shallow network and the deep network enables the image to be expressed in multiple scales to improve the recognition accuracy of the entire model. The experimental data is obtained from the downhole casing image dataset of an oil field in Sichuan. Experimental result shows that the macro-average F1 scores of the DS-CNN are 4.41 and 5.74 percentage points higher than those of the VGG19 model and the GoogleNet model, indicating that this model improves the recognition accuracy of oil casing damage images.

1. Introduction

Oil casing is essential equipment for the oil field production. Regularly inspecting oil casings and detecting damaged casings are important steps to ensure the normal production of oil field[1]. In recent years, the new well logging technology of Visible Downhole Television Logging has been widely used in downhole casing inspection. However, the current methods for detecting the casing damage are mainly by manually checking the video log images, whose recognition is inefficient[3]. Currently, Deep Learning has developed rapidly in the field of image recognition. Due to the large number of suspended particles in the well, uneven illumination and liquid flows, the images of the downhole casing obtained by video logging technology are grey, low clarity, and the damage contour features are not obvious[3]. When the conventional deep learning models are used for image recognition of casing damage, they are easy to result in a lower overall recognition rate[4]. Moreover, the available data of the casing image for the training is still small currently, and it is not sufficient to independently training a perfect convolutional neural network(CNN).

In order to solve the above problems, this paper proposes a DS-CNN model combined with the advantages of traditional shallow convolutional neural network and deep convolutional neural network. Multi-scale feature extraction of casing damage images improved image scale invariance. It is easier to train than the conventional CNN, and improve the recognition accuracy. In the process of identifying the casing damage image, the method firstly uses the shallow convolutional neural network to obtain the information of the global image feature, secondarily uses the deep convolutional neural network to obtain the information of the local image feature, then uses feature vectors jointly spliced
by the Inception construction as the input of the last fully connected layer, and uses the Sigmoid function as the full-join classification.

2. DS-CNN Video Log Image Recognition Model

2.1. DS-CNN Overview

With the deepening of convolution, a lot of automatic image recognition model’s identification work will pay more attention on the details of the image. But because of the image detailed sampling is excessive, it not only blurs the target of recognition, but also destroys the globalise of the image. The global features of the image can make the recognition target maintain a good overall contour, and the accuracy of the result will be more accurately[5][6]. While the shallow convolutional network model can keep the overall contour well. According to the characteristics of video logging images, the structure of DS-CNN model is shown in Fig.1. The upper part is a two-layer shallow convolutional neural network compose of Conv1, Conv2, Pooling1, Pooling2 and a fully connected layer. The lower part is a deep network model based on VGG19 and a fully connected layer. Then use a full-connection layer to combine feature vectors and obtain classification through a full-connection layer with Sigmoid function.

![Fig.1 Structure of DS-CNN](image)

2.2. The Shallow Neural Network Part

The shallow part of the structure of the DS-CNN model is mainly used to extract the information of global image feature[7]. IMi is used to represent the ith image in the data set, Shallow-Model is used to represent Shallow Model, and SVi is used to represent Shallow feature vector corresponding to the ith image. Therefore, the shallow model can be expressed as SVi = Shallow – Model (IMi).

As is shown in Fig.2, the shallow neural network structure is a global feature extraction structure composed of a two-layer convolution structure. In the first convolution layer, the convolution kernel is 5×5 and the step size is 1. In the second convolution layer, the convolution kernel is 3×3 and the step size is 1. The ReLU function is used as the activation function of the whole structure, and x is defined to represent the input vector from the upper neural network. The formula is expressed as ReLU(x) = max(0,x).

![Fig.2 Structure diagram of shallow neural network](image)
The output of the last pooling layer is taken as the input of the full connection layer to complete the feature extraction in the shallow convolutional neural network.

2.3. The Deep Neural Network Part

Deep Model is used to represent the Deep Model, and DV\textsubscript{i} is used to symbolize the Deep feature vector corresponding to the \(i\)th image. The deep model can be expressed as a \(DV_i = Deep\ Model(IM_i)\). The alternating structure of multiple convolutional layers and nonlinear activation layers in VGGNet network can better extract the detailed features of the image\[8]\[9]. Therefore, removing the last fully connected layer of the VGG19 model as the deep neural network of the DS-CNN. The parameter settings are shown in Table 1.

| Network name | Network Configurations | Layer |
|--------------|------------------------|-------|
| VGG19        | Conv3-64×2 Max-pool    | 19    |
|              | Conv3-128×2 Max-pool   |       |
|              | Conv3-256×4 Max-pool   |       |
|              | Conv3-512×4 Max-pool   |       |
|              | Conv3-512×4 Max-pool   |       |
|              | FC-4096                |       |
|              | FC-4096                |       |

2.4. The Joint Part

The joint part of the DS-CNN model uses the global features obtained by the shallow convolutional neural network to supplement and correct the abstract features obtained by the deep convolutional neural network. The splicing layer is used to coalesce the output of the deep and the shallow convolutional neural networks, which represented as FVi. The splicing function is Contact, and the splicing layer is FVi = Contact(SVi, DV\textsubscript{i}).

Inception network’s features of the different dimensions of the convolution layer can be used to extract multi-scale features of the image\[10]. The Inception network structure specific parameter settings are shown in Table 2.

| Partial Block | Block 1 | Block 2 | Block 3 | Block 4 |
|---------------|---------|---------|---------|---------|
| Size          | 3×3Conv | 3×3conv | 3×3conv×2 | 3×3Pooling |
|               | 1×1conv | 1×1conv | 1×1conv |         |

The input feature vector dimension is extracted by the depth and shallow convolutional neural networks, and the Dropout layer was added in the output of the joint layer to prevent overfitting. Different image features use corresponding convolution layers to average multiple networks to offset the problems caused by network overfitting. The value of probability is 0.5. After its activation value is set to 0, the vector will rescale.

The input of the full connection layer is the vector from the output of the splicing layer. The output length is the vector of the number of loss types, and the value represents the vector of the probability about the damage type belongs to the judged category. Sigmoid (FVi*W) represent the output of the model, W is the trainable matrix. This layer is expressed as Eq. (1).

\[
\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}}
\]

3. Experiment

3.1. Data Set

Experiments using data from visible light KATEYE™ hawkeye downhole television V2.0 logging equipment of Sichuan oilfield downhole casing video, using Direct Show Video captured technology capture the experimental dataset, including three types of Oil casing damage, casing corrosion, casing deformation, and casing perforation, as Fig.3. The dataset contains 1000 samples, among which 800 images are trainset and 200 images as testset, and the distribution is shown in Table 3.
3.2. Data Preprocessing
Considering the complexity of downhole environment, it will produce influence of pattern noise such as salt-and-pepper noise. Therefore, in the preprocessing of data samples, the median filter is selected to complete noise filtering, to obtain clearer casing damage images and improve the accuracy of image recognition.

3.3. Experiment and Result Analysis
During model training, the train set image feature vector is taken as the input of the convolutional layer. The recognition of casing damage image is regarded as a multi-label classification problem, and the binary cross-entropy is adopted as the loss function to process the training data set received by the input layer, which is expressed as Eq.(2). The output of the part of shallow network and the output of the part of deep network were fused as the input of the joint part of the DS-CNN model, and finally the optimized feature vector was obtained for classification and identification of the image damage.

\[
\text{loss} = - \sum_{i=1}^{n} y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i)
\]

In the Eq.(2), \(\hat{y}\) represents the probability distribution vector of the model prediction, \(y\) is the real probability distribution(Label). Adam optimizer minimizes loss to complete the optimization training of the model.

In order to verify the effectiveness of DS-CNN, a comparative experiment was conducted with VGG19 and GoogleNet. In the three networks, the number of the specific layer and node parameters of network training was set as the same: the initial learning rate was set as 0.01, and the weight attenuation was set as 0.0005. Macro-averaged F1-score was used as the assessment criteria. The higher the macro-average F1 score is, the better the performance of the model is. The final comparison results of the experiment are shown in Table 4.

|Methods| Corrosion| Deformation| Perforation| F1-macro-averaged|
|-------|---------|------------|------------|-----------------|
|VGG19  | 67.83%  | 66.91%     | 67.67%     | 67.6%           |
|GoogleNet| 65.46%  | 66.37%     | 66.98%     | 66.27%          |
|DS-CNN | 71.75%  | 73.47%     | 70.81%     | 72.01%          |

4. Conclusion
In this paper, an improved method of deep learning is proposed for identifying the oil casing damage images captured by the underground video logging equipment-Eagle Eye Downhole TV. Firstly, the images are processed to improve the quality of the training image. The high-quality image training set is used to complete the joint model of the DS-CNN. The global features obtained by the shallow convolutional neural network are fused with the local features obtained by the deep convolutional...
neural network to obtain the overall features. The model is compared with the VGG19 model and the GoogleNet model, and this model achieved better results.

The oil damage image recognition model of the DS-CNN proposed in this paper has several advantages: the joint model can extract multi-scale damage image features and improve the recognition accuracy; the recognition model of DS-CNN combines the advantages of deep convolutional neural network and shallow convolutional neural network, making this model more applicable to downhole video logging applications. Since the types of damages used in model training and testing in this paper are limited to corrosion, deformation and perforation, a further study can be carried out on more types of damages for optimized training, given the advantage of the deep learning on training with various categories. This will cover more kinds of casing damage that can occur in the oil field, which will further improve the accuracy of the oil casing damage identification in general.

References
[1] Yang Xu, Liu Shuhai, Li Feng, Zhang Shimin, Wang Wenming. Research progress in casing detection technology [J]. Petroleum Machinery, 2013, 41 (08): 17-22. (in Chinese).
[2] Cuan Ying, Xue Jijun. Quantitative interpretation of underground TV images[J]. Journal of Xi'an Shiyou University (Natural Science Edition). (in Chinese).
[3] Chai Manzhou, Xiang Xujin, Zhang Qingsheng, Yang Qingrong. Application of Downhole TV Logging System in Casing Inspection[J]. Well Logging Technology, 2002(03): 242-246+264. (in Chinese).
[4] Wang Haiyan, Zhang Qunhui, Zhang Yikuan. Application of Computer Image Processing and Pattern Recognition Technology in On-line Detection of Casing Defects[J]. New Technology & New Process, 2001(2). (in Chinese).
[5] Mumtaz S, Mubariz N, Saleem S, et al. Weighted hybrid features for person re-identification[C]// Seventh International Conference on Image Processing Theory. IEEE, 2018.
[6] ZHANG Yongliang,ZHANG Zhiqin,WU Hongbiao,DONG Lingping,ZHOU Bing.A Method of Perimeter Intrusion Detection Based on Improved Convolutional Neural Network[J].Computer Science,2017,44(03):182-186. (in Chinese).
[7] Yue K, Xu F, Yu J. Shallow and wide fractional max-pooling network for image classification[J]. Neural Computing and Applications, 2017.
[8] He Xueying, Han Zhongyi, Wei Benzheng. Recognition and classification of pigmented skin diseases based on deep convolutional neural network[J]. Journal of Computer Applications, 2018, 38(11): 3236-3240. (in Chinese).
[9] Qassim H, Verma A, Feinzimer D. Compressed residual-VGG16 CNN model for big data places image recognition[C]// 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC). IEEE, 2018.
[10] Yu Mingyang, Yang Peng, Wang Yijun. Erotic image detection based on convolutional neural network[J]. Computer Applications and Software, 2018, 35(01): 232-236+275. (in Chinese).