

## Article

# A Transfer Learning Methodology for Recognizing Unsafe Behavior during Lifting Operations in a Chemical Plant

Hua Li <sup>1</sup>, Xicheng Xue <sup>1</sup>, Yanbin Wang <sup>1,2</sup>, Lizhou Wu <sup>1</sup> and Xinhong Li <sup>1,\*</sup>

<sup>1</sup> School of Resources Engineering, Xi'an University of Architecture and Technology, No.13 Yanta Road, Xi'an 710055, China

<sup>2</sup> Shaanxi Construction Engineering Group Corporation, NO.11 Construction Engineering Group Company Limited, Wexing West Road, Xianyang 712099, China

\* Correspondence: safety\_lxh@163.com

**Abstract:** Large lifting equipment is used regularly in the maintenance operations of chemical plant installations, where safety controls must be carried out to ensure the safety of lifting operations. This paper presents a convolutional neural network (CNN) methodology, based on the PyTorch framework, to identify unsafe behavior among lifting operation drivers, specifically, by collecting 22,352 images of equipment lifting operations over a certain time period in a chemical plant. The lifting drivers' behavior was divided into eight categories, and a ResNet50 network model was selected to identify the drivers' behavior in the pictures. The results show that the proposed ResNet50 network model based on transfer learning achieves a 99.6% accuracy rate, a 99% recall rate and a 99% F1 value for the expected behaviors of eight lifting operation drivers. This knowledge regarding unsafe behavior in the chemical industry provides a new perspective for preventing safety accidents caused by the dangerous behaviors of lifting operation drivers.

**Keywords:** unsafe driver behavior; transfer learning; behavior recognition; chemical operations



**Citation:** Li, H.; Xue, X.; Wang, Y.; Wu, L.; Li, X. A Transfer Learning Methodology for Recognizing Unsafe Behavior during Lifting Operations in a Chemical Plant. *Processes* **2023**, *11*, 971. <https://doi.org/10.3390/pr11030971>

Academic Editor: Xiong Luo

Received: 21 February 2023

Revised: 17 March 2023

Accepted: 18 March 2023

Published: 22 March 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Owing to its rapid development, the chemical industry has become a large-scale operation, with professional management and advanced, large-scale chemical equipment. However, safety accidents often occur during the lifting of large chemical equipment, where the large diameter and weight of chemical equipment can result in many loopholes. The safety problems associated with large chemical equipment lifting operations can be classified into two main categories: one is caused by objective unforeseen circumstances and the other is caused by human irregularities. Note that the unsafe behavior of personnel is known to be the main reason for lifting accidents, according to the “4M” theory of lifting accidents. The traditional safety management of chemical operations relies on human resources for patrol inspection. There is less research on accidents caused by the unsafe behavior of drivers carrying out lifting operations. This paper examines the unsafe behaviors of drivers performing lifting operations, which is very important for understanding the causes of such accidents.

In recent years, the use of computer technology for the identification of unsafe behavior has become a mainstream in research. Li et al. [1] proposed a deep learning-based detection method for the detection and identification of personal protective equipment (PPE), achieving an accuracy of 0.9467 in an experimental scenario by model training. Zhao et al. [2] proposed a deep learning approach to build HAZOP knowledge graphs using deep neural networks with an accuracy of 90.75, a recall of 91.53 and an F1 score of 91.14. Ding et al. [3] proposed a new hybrid model, combining a convolutional neural network (CNN) and long short-term memory (LSTM) for automatic recognition of workers' unsafe behaviors. Wu et al. [4] aimed to develop a blockchain-enabled framework to ensure the effectiveness of on-site safety inspection of tower cranes (OSITC). Guo [5] developed

a skeleton-based real-time recognition method. The method identifies unsafe behaviors by reducing dynamic motion to static states. Zou [6] proposed a new method by which to identify different driving behaviors based on EEG data with an average accuracy of 80.6%. Wang et al. [7] proposed a hierarchical clustering method and an improved “Apriori” algorithm to identify unsafe driver behavior. [8–13] The above research validates the feasibility of deep learning methods for the problem of identifying unsafe behavior of people, but there are various detection problems in the actual identification situations.

In the actual recognition situation, using convolutional neural networks to recognize unsafe behavior during lifting operations suffers from a high rate of detection error, low accuracy and unstable data acquisition. In this paper, we propose a migration learning approach, where a pre-trained network model is used. Specifically, the ImageNet dataset is used to adjust the parameters, changing the output layer for classification, and using the state-farm driver behavior dataset and a self-collected driver behavior dataset from lifting operations to build a new dataset for experimentation, convolution and classification of the pre-processed images of unsafe driver behavior for the final identification of unsafe driver behavior during lifting operations. The method outperforms other CNN models, in terms of accuracy and precision, and can effectively identify the unsafe behaviors of drivers during chemical crane operation, providing a new method for the identification of unsafe driver behaviors.

The rest of this paper is organized as follows: Section 2 briefly describes the background to chemical lifting operations and transfer learning theory. Section 3 describes the recognition process for crane drivers behaviors based on convolutional neural networks, where the construction and pre-processing of datasets for the training of models are evaluated. Section 4 introduces the principles of transfer learning and examines the effects of transfer learning; it also compares the effects of each network model, where it is shown that the migration learning-based ResNet50 network model has high accuracy in classifying and identifying the behavior of drivers in lifting operations. Section 5 gives the conclusions.

## 2. Background to Chemical Lifting Operations and Principles of Transfer Learning

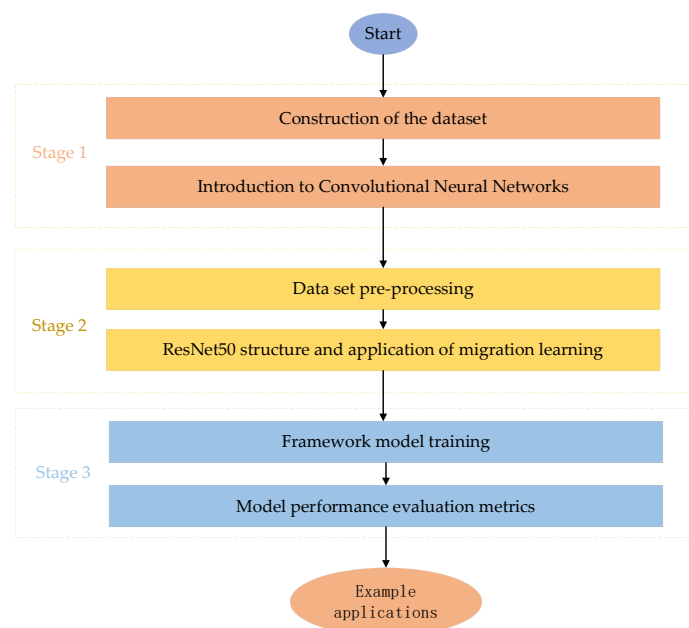
As the scale of petrochemical construction continues to expand, [14–16] the key equipment used in various projects is developing in the direction of extra heavy and extra large. Common lifting methods for large equipment, i.e., lifting with heavy machinery, masts and hydraulic devices, can be divided into the following two categories. First, h equipment, represented by towers and vessels. This type of equipment is large, high and heavy and the operations are difficult due to the large working space and high lifting capacity. Second, highly flexible structures, represented by flares, exhausts, etc. This type of equipment has a high length to slenderness ratio, low stiffness and poor structural stability. Vertical working space is required for lifting, and equipment is prone to deformation and instability during the lifting process. From the above key engineering equipment, it can be seen that the trend in lifting equipment is towards the heavy and the large. With the continuous improvement in lifting capacity the growing prevalence of increasingly large lifting projects and has become the main development trend for lifting technology. As a result, [17] there are increased demands for lifting technology and for the safe management of lifting operations.

Transfer learning is a machine learning method that takes existing knowledge and solves it for different but similar domain problems. In this paper [18–20], based on the transfer learning theory, we first pre-train the model on a large-scale dataset in a related domain, and then extract the weights and image features of the pre-trained model for initialization in the training process of the new small sample dataset. The model is then fine-tuned using a stochastic gradient descent algorithm. The larger and more complex the migration learning dataset, the more the model needs to be retrained. Migration learning works best when we use a dataset smaller than the original pre-trained model. Using a transfer learning model in Pytorch means choosing which layers to freeze and which ones to unfreeze. Freezing the model means that PyTorch retains the parameters (weights) in the specified layer. Unfreezing the model means that the specified layers are available

for training to make their weights trainable. After we have finished training the chosen pre-trained model layer, we can save the newly trained weights for later use.

### 3. Identification of Driver's Operational Behavior

In order to identify the unsafe behavior of drivers in chemical lifting operations, we propose a migration learning-based ResNet50 network model using the convolutional neural network approach (CNN) with the PyTorch framework. The ResNet50 network model was built to identify the operational behavior of crane operators. Note that convolutional neural networks suffer low accuracy and high error rates for unsafe behavior recognition. Nevertheless, we proposed a pre-trained network model to adjust the parameters, where a new dataset for detection was constructed. The proposed method significantly improves recognition rate and accuracy rate. The framework of the behavioral recognition method for drivers in lifting operations is given in Figure 1.



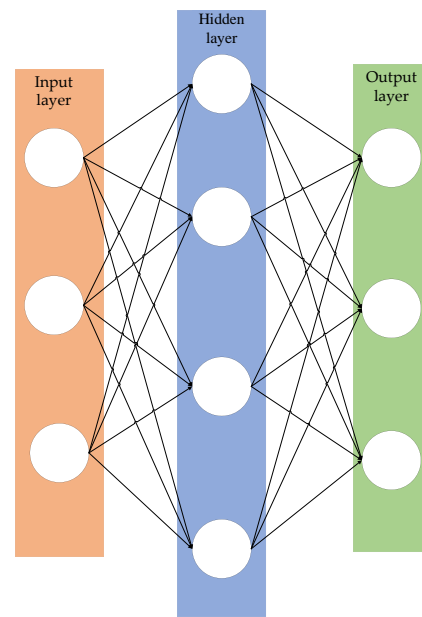
**Figure 1.** Transfer learning-based process for identifying unsafe driver behavior.

#### 3.1. Construction of the Dataset

With the popularization and broad application of video surveillance, collection speeds, processing capacity and transmission speeds of information have been greatly improved, which provides an excellent data source for collecting images of lifting operation drivers' behavior. This paper collected data by intercepting surveillance video streams and crawler technology. The data were collected from the driver while the lifting operation at the chemical plant was in progress, and the preliminary data collection was completed by acquiring images through video surveillance.

#### 3.2. Convolutional Neural Network (CNN) Principle

A neural network is a collection of connected units with input and output mechanisms, each of which has associated weights. Backpropagation is "backpropagation of errors" and has strong applicability to training neural networks. Backpropagation benefits deep neural networks dealing with error-prone projects, such as image recognition. A schematic diagram of the principle of convolutional neural networks is shown in Figure 2 below.



**Figure 2.** CNN schematic.

1. The forward propagation formula is used to obtain the activation value of each output layer.

The input and output of the hidden layer are shown in Equations (1) and (2):

$$\text{net}_j = \sum_{i=0}^m V_{ji} X_i \quad (1)$$

$$H_j = f(\text{net}_j) \quad (2)$$

where:  $\text{net}_j$  is the input of the hidden layer;  $V_{ji}$  is the weight of the  $i$ th neuron of the input layer connected to the  $j$ th neuron of the hidden layer;  $X_i$  is the output of the  $i$ th neuron of the input layer; and  $H_j$  is the  $j$ th neuron of the hidden layer. The result of each neuron,  $f$ , is the activation function.

The inputs and outputs of the output layer are shown in Equations (3) and (4):

$$\text{net}_k = \sum_{j=0}^n W_{kj} H_j \quad (3)$$

$$O_k = f(\text{net}_k) \quad (4)$$

where:  $\text{net}_k$  is the input of the  $k$ th neuron of the output layer and  $W_{kj}$  is the weight of the  $j$ th neuron of the hidden layer connected to the  $k$ th neuron of the output layer; the result is the  $k$ th neuron of the  $O_k$  output layer.

2. Reverse propagation

The cost function is shown in Equation (5):

$$E = \frac{1}{2} \sum_{k=1}^l (Y_k - O_k)^2 \quad (5)$$

where  $E$  is the cost function and  $Y_k$  is the desired output of the  $k$ th neuron in the output layer.

The output and hidden layers  $\delta$  (residuals) are shown in Equations (6) and (7):

$$\delta_k^O = -(Y_k - O_k) \cdot f'(\text{net}_k) \quad (6)$$

$$\delta_j^H = \sum_{k=1}^1 \delta_k^O \cdot W_{kj} f'(\text{net}_j) \quad (7)$$

where:  $\delta_k^O$  is the residual of the  $k$ th neuron in the output layer and  $\delta_j^H$  is the residual of the  $j$ th neuron of the hidden layer.

The weight increments are calculated from the input layer to the hidden layer and from the hidden layer to the output layer, as shown in Equations (8) and (9):

$$\Delta W_{kj} = -\eta \frac{\partial E}{\partial W_{kj}} = \eta \delta_k^O H_j \quad (8)$$

$$\Delta V_{ji} = -\eta \frac{\partial E}{\partial V_{ji}} = \eta \delta_j^H X_i \quad (9)$$

where:  $\Delta W_{kj}$  is the weight increment of the  $j$ th neuron of the hidden layer connected to the  $k$ th neuron of the output layer;  $\Delta V_{ji}$  is the weight increment of the  $i$ th neuron of the input layer connected to the  $j$ th neuron of the hidden layer; and  $\eta$  is the gradient descent step size.

The final weights are obtained as shown in Equations (10) and (11):

$$W_{kj} = W_{kj} + \Delta W_{kj} \quad (10)$$

$$V_{ji} = V_{ji} + \Delta V_{ji} \quad (11)$$

where:  $W_{kj}$  is the weight of the  $j$ th neuron of the hidden layer connected to the  $k$ th neuron of the output layer and  $V_{ji}$  is the weight of the  $i$ th neuron of the input layer connected to the  $j$ th neuron of the hidden layer. The bias term is  $H_0 = X_0 = 1$ .

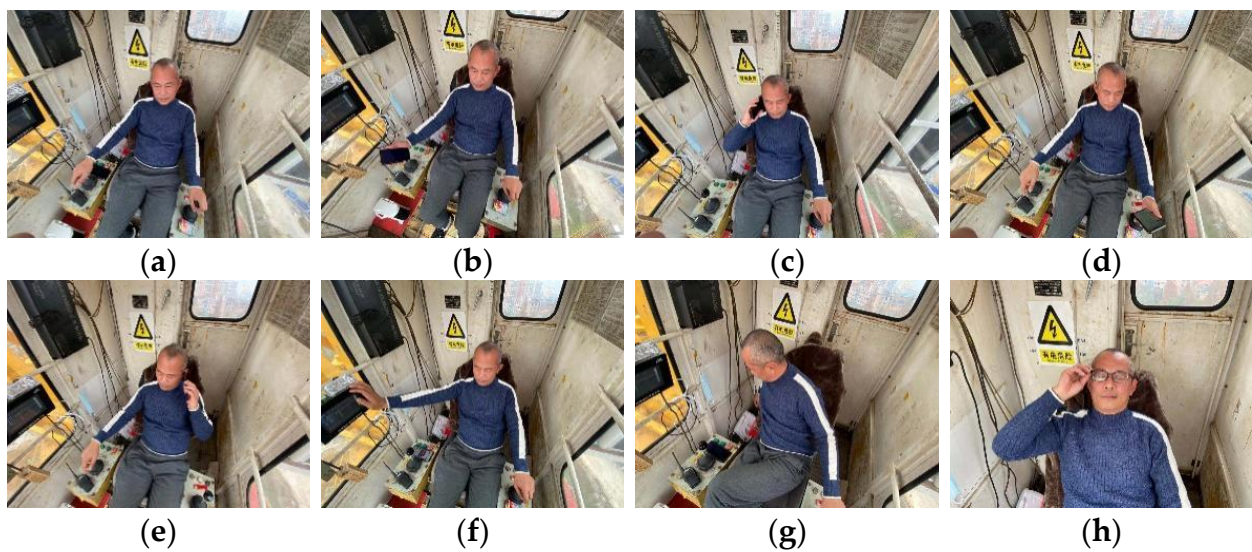
In the process of forwarding propagation, the value of each node in the network is calculated, where the backpropagation process uses these values as well as the chain rule for derivatives in combination with dynamic programming. Repetitive operations are avoided to improve gradient calculation speed. As a result, it is more conducive to the gradient to perform stochastic gradient descent for updating the neural network parameters.

### 3.3. Dataset Preprocessing

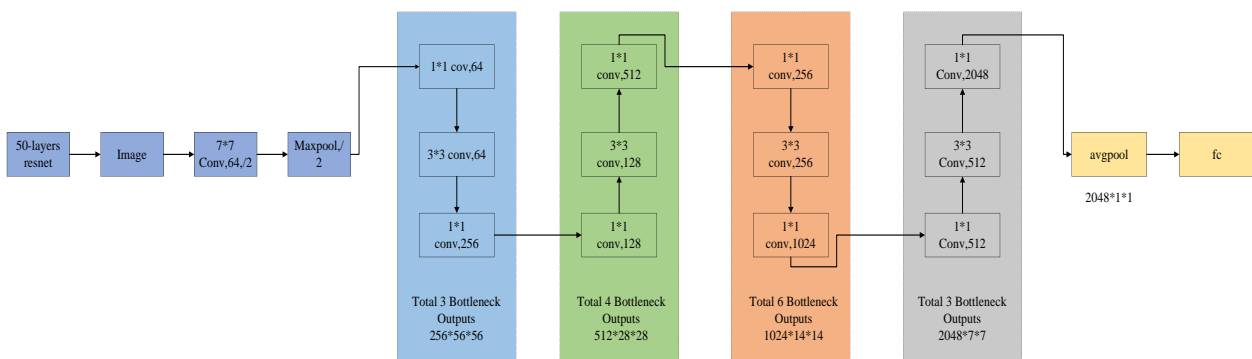
When the behavioral classification is based on the accident causation theory, and other classification bases, it tends to result in fewer classification categories, which eventually leads to a high rate of detection errors. In this paper, we performed the behavioral classification according to human behavior. The collected 102,150 images of the various driver behaviors were classified and divided into eight categories according to the recognition target, namely, everyday driving, i.e., looking at the phone (right hand), looking at the phone (left hand), calling (right hand), calling (left hand), adjusting meters, holding things in the back seat, and holding glasses. An example image of tower crane driver behavior is shown in Figure 3.

### 3.4. ResNet50 Structure

Theoretically, the ability to express is more vital with deeper networks. However, network degradation, gradient disappearance, and explosion problems occurs more often with an increasing number of learning layers. This problem can be solved by the standard initialization and regularization layers, which ensures that dozens of layers of the network can converge. However, it also suffers from the issue of network degradation, gradient disappearance, and explosion. Therefore, this paper used ResNet (residual network), which introduces jump connections, enabling an unimpeded flow of information from the previous residual block. Thus, the proposed method can avoid network degradation, gradient disappearance, and the degradation problems caused by being too deep within the network. The network structure of ResNet50 is shown in Figure 4.



**Figure 3.** Example images of driver behavior. (a) Everyday driving. (b) Looking at the phone (right hand). (c) Calling (right hand). (d) Looking at the phone (left hand). (e) Calling (left hand). (f) Adjusting meters. (g) Holding things in the back seat. (h) Holding glasses.



**Figure 4.** ResNet50 network structure.

When using deeper networks, the expression ability is more vital. However, when the network depth is increased to a certain extent, deeper networks mean more training mistakes. The reason for the elevated error is that the more profound the network, the more pronounced the disappearance of the gradient. In backward propagation, the inability to update the rise to the previous network layer efficiently: this means the preceding network layer parameters cannot be updated, which leads to poor training and testing results. In this paper, we decided to use ResNet (residual network), which adds constant mapping, transferring the current output directly to the next layer of the network so this level of arithmetic is skipped. It also passes the gradient of the lower layer of the network directly to the upper layer in the backward propagation process. Residual networks make feed-forward, back-propagation algorithms work. To a large extent, residual networks make it easier to optimize deeper models. The network structure of ResNet50 is shown in Figure 4.

In Figure 4, conv stands for convolutional operation; Maxpool is Max Pooling; Avgpool is average pooling; and the rectangular box corresponds to the residual module. The size of the input layer is  $244 \times 244 \times 3$ , where a  $7 \times 7$  convolutional kernel is responsible for the extraction of features. The step size of the convolution kernel is 2. This reduces the length and width of the image to half of its original size. Subsequently, a further Maxpool layer is passed through to further reduce the resolution of the image. Firstly, we obtained  $256 \times 56 \times 56$  feature maps with the first bottleneck module. Secondly, the next

$512 \times 28 \times 28$  feature maps were obtained with the second bottleneck module. Thirdly, the third bottleneck module was used to obtain  $1024 \times 14 \times 14$  feature maps. Finally, the  $512 \times 28 \times 28$  feature maps were obtained with the fourth bottleneck module, where we obtained  $1 \times 1 \times 2048$  feature maps in one layer of Avgpool. As a result, we used fully connected layers to obtain eight crane driver behavior prediction probabilities after Softmax classification.

### 3.5. Model Training

Processed images are inputted into a deep-learning framework with multiple iterations and parameter updates. The training process consists of the following steps:

- (1) Data enhancement. The small training sample creates more pieces by flipping, scaling, cropping, panning, adding noise, etc. Thus, we should improve the generalization ability of the neural network. The ResNet50 input image is  $244 \times 244 \times 3$ , and the dataset images are resized to match the training image requirements.
- (2) Improved model generalization capabilities. If the sample is too small, models tend to fall into overfitting during training. In order to maintain a dynamic balance between model complexity and data volume, we should reduce the weight of the high-dimensional parameter values using regularization methods, automatically adapt parameters to different images of unsafe behavior, tune each convolutional layer parameter to a reasonable range, and improve its generalization performance.
- (3) Validation and training set assignment for all data. The role of the validation set in the training process is to determine the model and modify the hyperparameters (e.g., learning rate), reverse the recursive modification of weights, verify the generalization capability of the model, and make the model optimal on the validation set.
- (4) Modification of training-related parameters. Altering parameters, such as batch size and anchor scale, to improve the training result.

### 3.6. Model Performance Evaluation Metrics

The effectiveness of driver behavior recognition using the ResNet50 network can be assessed in terms of accuracy. Accuracy includes four metrics: precision, recall, and F1 value, as well as a confusion matrix.

- (1) Accuracy (precision) is expressed as the percentage of samples with a positive recall, and the number of instances that have been correctly classified as positive. The calculation is shown in Equation (12):

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (12)$$

where TP is the number of samples whose classification is positive and FP is the number of pieces whose sort is negative.

- (2) Recall rate (recall) indicates the number of positive samples that are recalled from the class. The calculation is shown in Equation (13):

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (13)$$

where FN is the number of positive samples incorrectly predicted to be classified as harmful.

F1 value (score), also known as equilibrium F-score, is defined as the summed average of the precision and recall rates. The calculation is shown in Equation (14):

$$\text{F1} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (14)$$

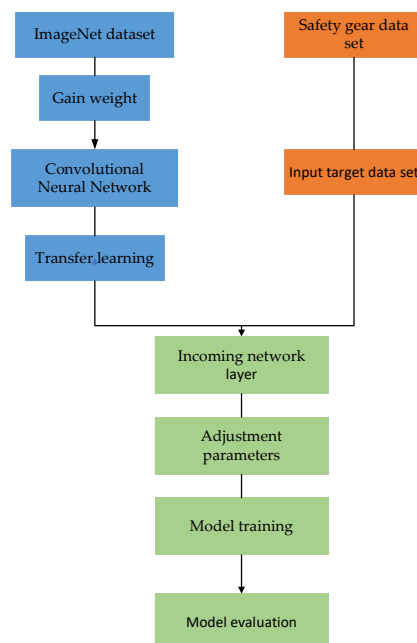
- (3) Each column of the confusion matrix represents a prediction category, where the total of each column indicates the number of data predicted to be in that category. Each row represents the accurate attribution category of the data, where the total number of data in each row indicates the number of data instances of that category.

#### 4. Instance Verification

In the experiments, we built a convolutional neural network framework such as ResNet50 on Windows 10 OS, Inter (R) Core (TM) i7-8750H 6-core 12-thread processor, NVIDIA GeForce GTX 1050 Ti graphics card, 8GB system memory using PyTorch 1.5.1 and compiled it using Python language. In order to avoid the impact of sample imbalance on the model, 22,352 images, of operators in the operating room at a certain time during the lifting operation in the chemical plant, were used as the training and testing dataset for the model. The model creates more samples before training by flipping, scaling, cropping, panning, adding noise, etc, to increase generalization and prevent overfitting.

##### 4.1. Applications of Transfer Learning

The behavior dataset of lifting operation drivers is limited. Consequently, only a limited amount of data can be used for learning algorithms. This inevitably leads to poor accuracy of the resulting model. Moreover, it is easy to cause overfitting in the case of restricted data sets. Although the accuracy of the training samples is exceptionally high, the generalization effect can be inferior. Therefore, the transfer learning idea is adopted to transfer the learned and trained model parameters to the new model, with the aim of helping to train the new model. Since a large amount of data is correlated, sharing the trained model parameters to the new model by transferring learning speeds up and optimizes the learning efficiency of the model. As a result, it improves the recognition accuracy of unsafe behaviors. The process of transfer learning is shown in Figure 5.



**Figure 5.** Transfer learning process.

ResNet50, He et al. [21], is fine-tuned using transfer learning, where the ImageNet dataset and the ResNet50 network are initialized with the trained weight parameters. On this basis, the lifting operation driver behavior dataset is used to train and adjust the parameters.



#### 4.2. Comparison of Transfer Learning Effects

For migration learning, we used the weight parameters of the pre-trained ImageNet dataset as ResNet50 network weight parameters to initialize the model. The model training iteration curve is shown in Figure 6.

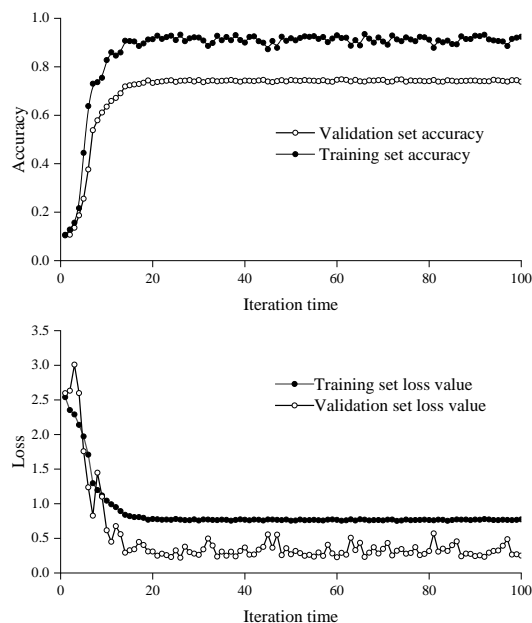


Figure 6. Lost and accuracy values without migration learning.

As seen from Figure 6, there was a slow improvement in accuracy of the training set during training for the ResNet50 network model without migration learning, where the loss value during the iterations hovered around 1.0. At the same time, the accuracy and loss values also changed slowly on the validation set. It can be seen from Figure 7 that, the network can fit the sample accurately using transfer learning, which rapidly improves accuracy and reduces the loss value and the proposed method can converge quickly. It can be seen from the comparison that migration learning can significantly improve the network.

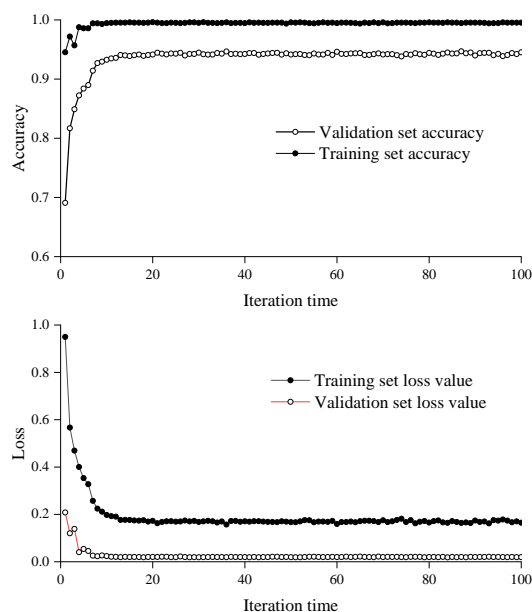


Figure 7. Lost vs. accuracy values after using migration learning.

#### 4.3. Comparison of the Effects of Each Network Model

Using the ResNet50 structure for training, the learning rate was 0.001, and the training period was 100. The evaluation indicators obtained are shown in Table 1.

**Table 1.** ResNet50 model test results.

Network Model		ResNet50		
Assessment Indicators	Precision	Recall	F1	
Everyday driving	1.00	0.99	0.99	
Looking at the phone (right hand)	1.00	1.00	1.00	
Calling (right hand)	1.00	1.00	1.00	
Looking at the phone (left hand)	1.00	1.00	1.00	
Calling (left hand)	1.00	1.00	1.00	
Adjusting meters	1.00	1.00	1.00	
Holding things in the back seat	1.00	1.00	1.00	
Holding glasses	0.99	0.98	0.99	
Average value	0.99	0.99	0.99	

It can be seen from Table 1 that the accuracy is 99% in identifying eight categories of common lifting operation driver behaviors, where the recall rate is 98%, and the F1 value is 98% based on the ResNet50 model. It can be seen that recognition of typical behaviors of lifting operation drivers has a reasonable recognition rate, which enables effective identification of the unsafe behaviors of lifting operation drivers. The proposed method extracts histogram of oriented gradient (HOG) features, which are used for image recognition in the Random Forest (RF) classifier. We constructed a lightweight CNN network with two convolutional layers, two pooling layers, and one fully connected layer. The results of the migration learning based ResNet50 compared with the HOG-RF and light CNN networks are shown in Table 2.

**Table 2.** ResNet50 vs. HOG-RF and light CNN networks.

Network Type	Light CNN	HOG-RF	ResNet50	Maximum Difference
Everyday driving	0.96	0.87	1.00	0.13
Looking at the phone (right hand)	0.96	0.92	1.00	0.08
Calling (right hand)	0.98	0.94	1.00	0.06
Looking at the phone (left hand)	0.98	0.93	1.00	0.07
Calling (left hand)	0.97	0.95	1.00	0.05
Adjusting meters	0.98	0.98	1.00	0.02
Holding things in the back seat	0.97	0.96	1.00	0.04
Holding glasses	0.93	0.93	0.99	0.06
Average value	0.96	0.93	0.99	0.06

As can be seen from Table 2, the recognition rate of the migration learning-based ResNet50 network is higher than both the light CNN structure and the HOG-RF. The recognition rate of the migration learning-based ResNet50 network is 3.25% higher than the lightweight CNN structure, and 6.375% higher than the HOG-RF algorithm. Among them, the recognition rate of hand-held glasses improved the most, by 6%. There was a minimal increase in talking on the phone (right hand) and looking at the phone (left hand), of 2%. The experimental results exemplify the accurate information identified by the migration learning-based ResNet50 network.

The confusion matrix of the lifting operation driver behavior recognition results obtained from the migration learning-based ResNet50 experiments are shown in Table 3.

**Table 3.** ResNet50 driver behavior recognition confusion matrix.

Identification Categories	Holding Glasses	Adjusting Meters	Looking at the Phone (Right Hand)	Calling (Right Hand)	Looking at the Phone (Left Hand)	Calling (Left Hand)	Holding Things in the Back Seat	Everyday Driving
Holding glasses	300	0	0	0	0	0	0	0
Adjusting meters	1	197	0	0	1	0	1	0
Looking at the phone (right hand)	0	0	200	0	0	0	0	0
Calling (right hand)	0	1	0	299	0	0	0	0
Looking at the phone (left hand)	0	0	0	0	299	1	0	0
Calling (left hand)	0	0	0	0	1	299	0	0
Holding things in the back seat	0	0	0	0	0	0	200	0
Everyday driving	0	0	1	0	0	1	0	298

As can be seen from the confusion matrix diagram, holding glasses, looking at the phone (right hand) and holding something in the back seat, are the highest. A small amount of meter adjustment is mistaken for holding glasses and looking at a mobile phone (left hand). The most common type of misclassification is left-handed mobile phone use. As can be seen from the confusion matrix, overall driver behavior recognition accuracy has reached a high level and has an ideal recognition effect.

## 5. Conclusions

- (1) Recently, the lifting of large chemical equipment has become widely used and safety management and the control of lifting operations has been a major concern. Many accidents in lifting operations are caused by improper handling by the driver. In this paper, we aimed to recognize driver behavior during lifting operations. It is well known that convolutional neural networks suffer from a high error rate for the detection of unsafe behavior in lifting operation drivers. We proposed a ResNet50 network model based on migration learning, where the network structure is analyzed and its parameters are adjusted.
- (2) This network has a high accuracy in classifying and recognizing the behavior of eight common types of crane driver behavior. The results show that the recognition method can identify crane driver behaviors more accurately, with an average recognition rate of 99.6%, which is 3.25% higher than the average accuracy of light CNN structures and 6.375% higher than the traditional HOG-RF algorithm.
- (3) For the study of unsafe driver behavior in lifting operations, the dataset from this study is small and does not reflect all the possible behaviors of lifting drivers. In addition, the timing of data collection was relatively fixed, failing to consider other scenarios. Further research needs to consider adding dynamic identification to improve the identification of driver unsafe behavior in order to further contribute to the safety management of chemical lifting operations.

**Author Contributions:** Writing—original draft preparation, H.L.; formal analysis and funding acquisition, X.X.; methodology, Y.W. and L.W.; data curation, X.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was funded by Xi'an Architectural Science and Technology University Engineering and Technology Co., Ltd. (XAJD-YF23N010). The project title: Research on intelligent early warning technology for safety protection detection of construction operations.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Li, J.; Zhao, X.; Zhou, G.; Zhang, M. Standardized use inspection of workers' personal protective equipment based on deep learning. *Saf. Sci.* **2022**, *150*, 105689. [[CrossRef](#)]
2. Zhao, Y.; Zhang, B.; Gao, D. Construction of petrochemical knowledge graph based on deep learning. *J. Loss Prev. Process Ind.* **2022**, *76*, 104736. [[CrossRef](#)]
3. Ding, L.; Fang, W.; Luo, H.; Love, P.E.; Zhong, B.; Ouyang, X. A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory. *Autom. Constr.* **2018**, *86*, 118–124. [[CrossRef](#)]
4. Wu, H.; Zhong, B.; Li, H.; Chi, H.L.; Wang, Y. On-site safety inspection of tower cranes: A blockchain-enabled conceptual framework. *Saf. Sci.* **2022**, *153*, 105815. [[CrossRef](#)]
5. Guo, H.; Yu, Y.; Ding, Q.; Skitmore, M. Image-and-Skeleton-Based Parameterized Approach to Real-Time Identification of Construction Workers' Unsafe Behaviors. *J. Constr. Eng. Manag.* **2018**, *144*, 04018042. [[CrossRef](#)]
6. Zou, B.; Xiao, Z.; Liu, M. Driving Behavior Recognition Based on EEG Data From a Driver Taking Over Experiment on a Simulated Autonomous Vehicle. *J. Phys. Conf. Ser.* **2020**, *1550*, 042046. [[CrossRef](#)]
7. Wang, T.; Chen, Y.; Yan, X.; Chen, J.; Li, W. The Relationship between Bus Drivers' Improper Driving Behaviors and Abnormal Vehicle States Based on Advanced Driver Assistance Systems in Naturalistic Driving. *Math. Probl. Eng.* **2020**, *2020*, 9743504. [[CrossRef](#)]
8. Xu, H.; Liu, Y.; Shu, C.M.; Bai, M.; Motalifu, M.; He, Z.; Wu, S.; Zhou, P.; Li, B. Cause analysis of hot work accidents based on text mining and deep learning. *J. Loss Prev. Process Ind.* **2022**, *76*, 104747. [[CrossRef](#)]
9. Tian, F.; Li, H.; Tian, S.; Tian, C.; Shao, J. Is There a Difference in Brain Functional Connectivity between Chinese Coal Mine Workers Who Have Engaged in Unsafe Behavior and Those Who Have Not? *Int. J. Environ. Res. Public Health* **2022**, *19*, 509. [[CrossRef](#)]
10. Al-Hussein, W.A.; Por, L.Y.; Kiah, M.L.M.; Zaidan, B.B. Driver Behavior Profiling and Recognition Using Deep-Learning Methods: In Accordance with Traffic Regulations and Experts Guidelines. *Int. J. Environ. Res. Public Health* **2022**, *19*, 1470. [[CrossRef](#)]
11. Ni, D.; Guo, F.; Zhang, H.; Li, M.; Zhou, Y. Improving Older Drivers' Behaviors Using Theory of Planned Behavior. *Sustainability* **2022**, *14*, 4769. [[CrossRef](#)]
12. Kareem, M.; Tarek, H. Review of image-based analysis and applications in construction. *Autom. Constr.* **2021**, *122*, 103516. [[CrossRef](#)]
13. Liang, Q.; Zhou, Z.; Ye, G.; Shen, L. Unveiling the mechanism of construction workers' unsafe behaviors from an occupational stress perspective: A qualitative and quantitative examination of a stress–cognition–safety model. *Saf. Sci.* **2022**, *145*, 105486. [[CrossRef](#)]
14. Gao, Y.; Fan, Y.; Wang, J.; Li, X.; Pei, J. The mediating role of safety management practices in process safety culture in the Chinese oil industry. *J. Loss Prev. Process Ind.* **2018**, *57*, 223–230. [[CrossRef](#)]
15. Vianello, C.; Milazzo, M.F.; Maschio, G. Cost–benefit analysis approach for the management of industrial safety in chemical and petrochemical industry. *J. Loss Prev. Process Ind.* **2019**, *58*, 116–123. [[CrossRef](#)]
16. Sarbayev, M.; Yang, M.; Wang, H. Risk assessment of process systems by mapping fault tree into artificial neural network. *J. Loss Prev. Process Ind.* **2019**, *60*, 203–212. [[CrossRef](#)]
17. Khan, A.; Sattari, F.; Lefsrud, L.; Tufail, M. Enhancing regional process safety management. *J. Loss Prev. Process Ind.* **2021**, *71*, 104444. [[CrossRef](#)]
18. Xu, Y.; Huang, Y.; Ma, G. A beetle antennae search improved BP neural network model for predicting multi-factor-based gas explosion pressures. *J. Loss Prev. Process Ind.* **2020**, *65*, 104117. [[CrossRef](#)]
19. Jiao, Z.; Ji, C.; Yuan, S.; Zhang, Z.; Wang, Q. Development of machine learning based prediction models for hazardous properties of chemical mixtures. *J. Loss Prev. Process Ind.* **2020**, *67*, 104226. [[CrossRef](#)]
20. Hemmatian, B.; Casal, J.; Planas, E.; Hemmatian, B.; Rashtchian, D. Prediction of BLEVE mechanical energy by implementation of artificial neural network. *J. Loss Prev. Process Ind.* **2020**, *63*, 104021. [[CrossRef](#)]
21. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *arXiv* **2015**, arXiv:1512.03385.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.