CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/135231/

This is the author's version of a work that was submitted to / accepted for publication.

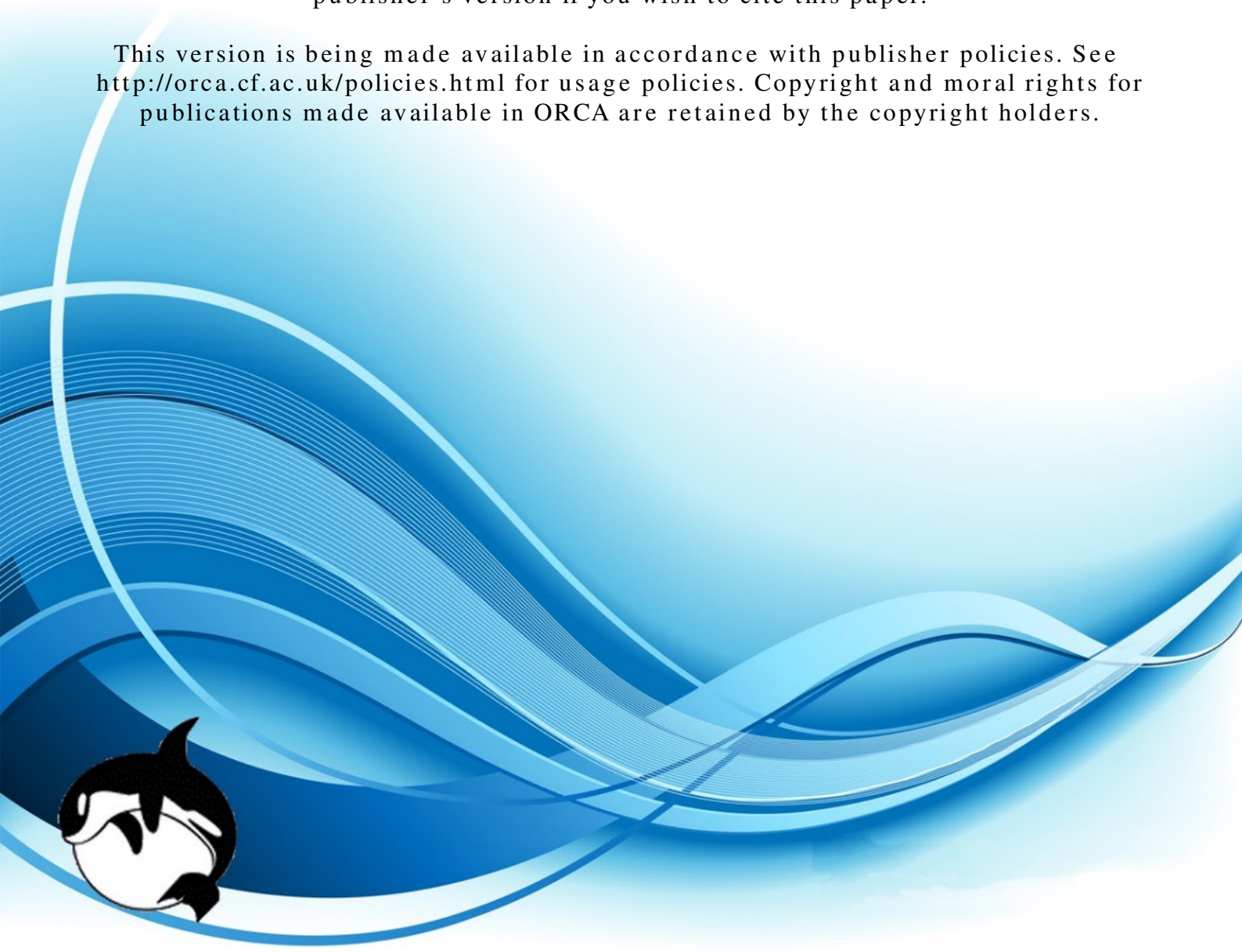
Citation for final published version:

Yao, ZongWei, Huang, Qiuping, Ji, Ze , Li, XueFei and Bi, Qiushi 2021. Deep learning-based prediction of piled-up status and payload distribution of bulk material. Automation in Construction 121 , 103424. 10.1016/j.autcon.2020.103424

Publishers page: https://doi.org/10.1016/j.autcon.2020.103424

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

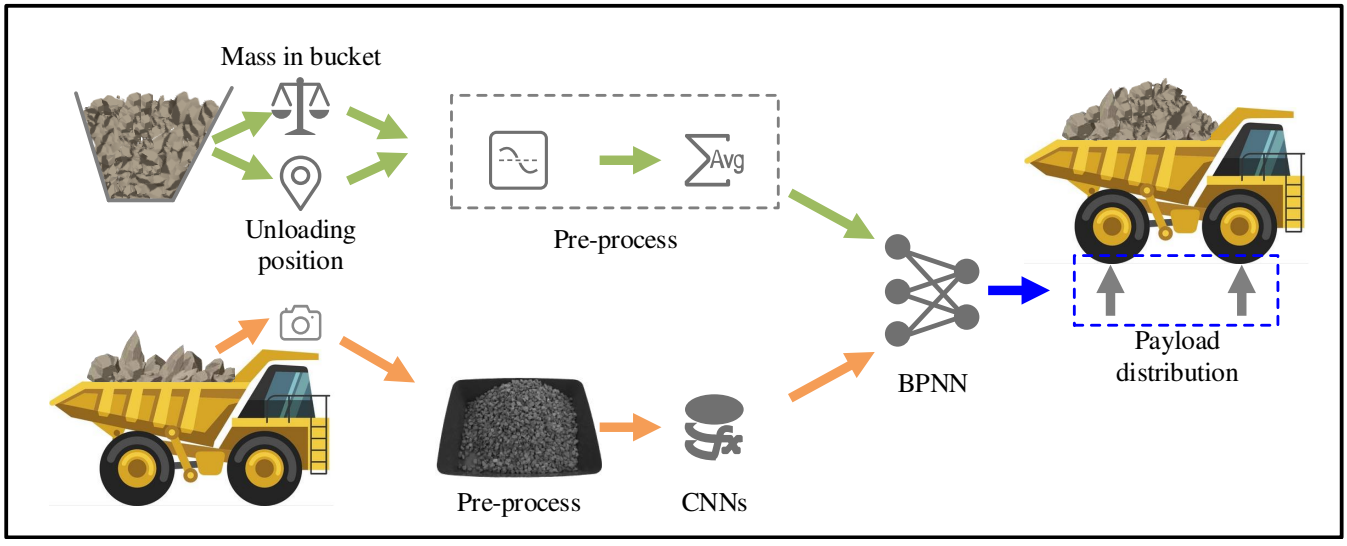
This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Graphical Abstract

Deep learning-based prediction of piled-up status and payload distribution of bulk material

ZongWei Yao, Qiuping Huang, Ze Ji, XueFei Li, Qiushi Bi



Highlights

Deep learning-based prediction of piled-up status and payload distribution of bulk material

ZongWei Yao, Qiuping Huang, Ze Ji, XueFei Li, Qiushi Bi

- The proposed method can automatically predict the Piled-up Status and Payload Distribution (PSPD) of bulk materials in terms of material mass and dumping position from pure images.

- The work introduced a two-stage prediction-regression CNN for predicting the PSPD of bulk materials.

- The proposed method has promising generalization capability in the prediction of the PSPD of bulk material with no special requirements on material characteristics, shape of loader bucket and truck body, dumping position, and so on.

Deep learning-based prediction of piled-up status and payload distribution of bulk material

ZongWei Yao[a], Qiuping Huang[a], Ze Ji[b,*], XueFei Li[a], Qiushi Bi[a,*]

[a] *Key Laboratory of CNC Equipment Reliability, Ministry of Education, School of Mechanical and Aerospace Engineering, Jilin University, Changchun, 130025, China*

[b] *Queen's Building, School of Engineering, Cardiff University, Cardiff, CF24 3AA, UK*

Abstract

The piled-up status of bulk material in a haul truck body determines the load balance, hence affects the mining operations' efficiency. Prediction of Piled-up Status and Payload Distribution (PSPD) of bulk material contributes to providing optimal dumping positions to improve the vehicle's stress state and service life. This work introduces a novel deep learning-based PSPD prediction method from images. A two-stage prediction-regression CNN model is designed to automatically extract image features to obtain the PSPD of the current state. The PSPD prediction is accomplished via a backward-propagation neural network (BPNN). Scaled model experiments are performed using robots for validating the method. Experiments show the trained model accurate and reliable in prediction and computationally efficient. The probability density of the prediction error is subject to the Cauchy distribution with x_0 of -0.00043 and γ of 0.01986. The maximum prediction error is $0.19kg$ (about 3.17% of total weight).

Keywords: Deep learning, Bulk material, Piled-up status and payload distribution

Contents

1 Introduction **1**

2 Related works **3**
2.1 Online payload estimation 3
2.2 Bulk material behavior analysis 4
2.3 Machine learning for handling bulk material 5

3 Method **5**
3.1 Proposed framework overview 5
3.2 Data pre-process 6
3.3 CNNs 7
3.3.1 Classification 7
3.3.2 Regression 9
3.4 BPNN 9

4 Experiments **9**
4.1 Experiment setup 10
4.2 Task specifications 10
4.3 Experiment data 11

5 Results and discussion **12**
5.1 Prediction for the experiment data 12
5.1.1 Classification 12
5.1.2 Regressors 12
5.1.3 Results of BPNN 14
5.2 Comparison with other models 14
5.3 Real-world environment 16

6 Conclusions **18**

1. Introduction

Mining haul trucks are essential equipment in the open-pit mining industry. The transportation efficiency can be heavily influenced by the performance of these vehicles for the whole mining processes. Statistical reports indicate that unbalanced payload distributions for such trucks are common throughout their entire periods of service and could lead to many problems. One typical uneven loading case is shown in Figure 1, where the truck frame is subject to excessive stress. The overloading on one side of the truck not only seriously affects the productivity and efficiency of mining trucks, but also significantly reduces the life span of tires, suspensions and other structural components. Such phenomena will even lead to catastrophic structural failures as shown in Figures 2. Besides, the whole-body vibrations caused by the unbalanced payload distributions during the transportation makes the truck drivers under immense safety and occupational health risks[2].

Mining trucks generally transport blasted bulk materials comprising discrete particles of various sizes, while these materials cannot be self-balanced like fluids, nor can they

*Corresponding author
Email addresses: yzw@jlu.edu.cn (ZongWei Yao), huangqp18@mails.jlu.edu.cn (Qiuping Huang), jiz1@cardiff.ac.uk (Ze Ji), lxf2014@jlu.edu.cn (XueFei Li), bqs@jlu.edu.cn (Qiushi Bi)

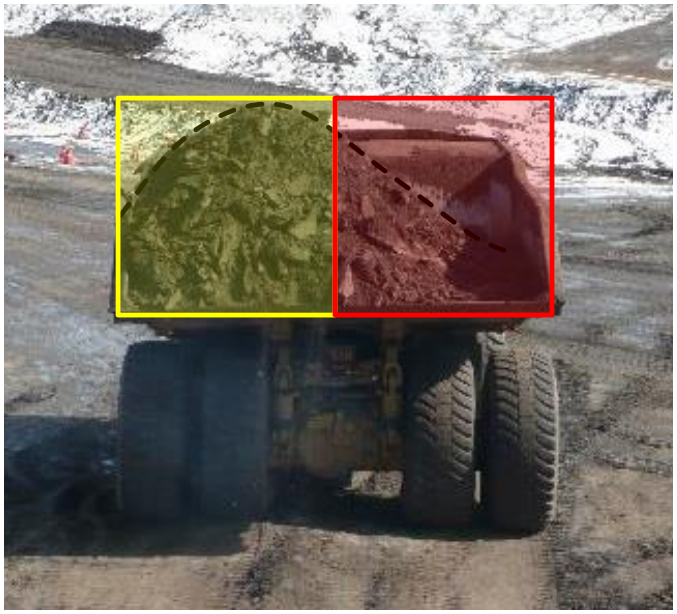


Figure 1: The truck with uneven material distribution [1]

Figure 2: Axle failure on a truck [3]

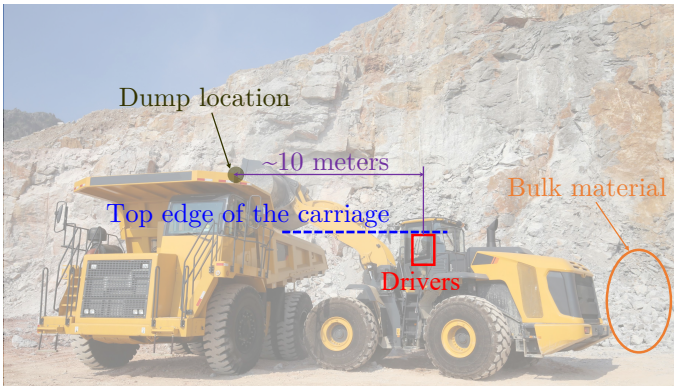


Figure 3: It is difficult for drivers to be fully informed about unloading operations

be loaded in an orderly manner like prefabricated solids with regular shapes. Therefore, the balance of payload distribution can only be carried out by a reasonable selection of dumping positions. For now, operators rely on manual observations to balance the truck payload when dumping from loaders or shovels. Due to the limited field of view of the driver and the obstructions of equipment, it is not possible to clearly observe how the materials are distributed inside the truck body, as shown in Figures 3.

Although the payload distribution in the truck body can be simplified through the pressure signals from the suspension system, the dumped materials could move and, hence, deform overtime due to the granular characteristics. The deformation behaviour is difficult to predict, making the final payload distribution after each dump very difficult to be precisely described. Therefore, even with the help of some dumping assisting software, the determination of dumping positions is still mostly based on personal experience, which causes the unbalanced payload distribution unavoidable.

According to the operation principles of typical shovels-truck systems, it usually takes 5-7 dumps to fill the truck body, meaning the PSPD would change significantly after each dump. However, in actual operations, the randomness of the dumping position cannot be ignored, while the payload distribution can only be estimated after each dump through the steady pressure signals from the suspension system. In other words, even if the suspension system can provide the risk alert, it is too late because the blasted materials cannot be re-dumped that makes the truck tires and the structure subject to excessive loads during the transportation process. In addition, it should be noted that the payload sensing system needs to be frequently calibrated online on a standard platform. Otherwise, data drift would largely impact the measurement results. The payload estimation sensors are subject to complex loads, resulting in short life-span and frequent accidental damages. Once the replacement of the sensors is required, the truck must be stopped and overhauled, which would greatly affect the efficiency of the production chain. Another issue that should be taken into consideration is that the payload estimation system can only function properly when the truck is operating in a quasi-static state. However, it is difficult to eliminate the dynamic loads during the dumping process making the accuracy and reliability of the sensor system hardly meet the requirements of the actual applications.

For these reasons, numerous researchers have attempted to use theoretical and numerical analyses to improve the payload distribution in mining trucks. These studies mainly focus on the following three aspects, which are 1) analyzing the flow behavior of bulk materials in the truck body using the DEM (Discrete Element Method) simulation; 2) the geometry optimization of the truck body aiming to make the material in the truck body achieve a certain degree of self-balancing; and 3) the topology op-

timization of key support components of the vehicle to reduce the stress level and increase the strength limit.

Although these studies provide valuable references to this study, they have not fundamentally solved the problem of payload unbalanced distribution due to some conditional limitations. Therefore, establishing a method to improve the unbalanced distribution of payload in mining trucks remains a challenge. One main reason for the unbalanced payload distribution is due to the lack of reasonable planning of dumping positions. More balanced payload distribution can be achieved through optimizing the dumping position before each cycle reaching a balanced payload distribution result. Serving the vision of optimal planning of dumping position of each load cycle, it is important to effectively and precisely predict the distribution result of bulk material in the truck body before each dump based on the loader dump position, the digging payload and the current payload distribution inside the truck body. The method for payload distribution prediction can significantly reduce the randomness of dumping operations accomplished by drivers and the risks caused by insufficient information and decision-making ability. Meanwhile, the payload distribution prediction method can provide technical support for the autonomous operation of the loader-truck system and contribute to the intelligent process of the mining industry.

It is obvious that experienced drivers of loaders or shovels can mostly control the dumping operation resulted in the fairly balanced payload distribution in the mining truck body, which is evidently better than that of inexperienced ones. In other words, the ability to predict the PSPD of bulk material in the truck body can be "learned" by training through observations from human naked-eyes, without additional precise measurement. The rationale of this work is that it is possible to make accurate enough prediction of the payload distributions in the truck body with relatively "inaccurate" or indirect measurement from partial 2D vision information of the pile surfaces, in order to make an optimal decision for the position of next dump. For the purpose of this project, we assume that the payload distribution can be estimated accurately with images of the surface profiles.

This study proposes a deep learning-based method using a low-cost vision sensor for predicting the PSPD of bulk materials in the truck body under complex environments. In the work, a series of images of material piles in the truck body dumped from different positions are collected and used for training the convolutional neural networks (CNNs), in order to estimate the current PSPD through the surface states. A backward-propagation neural network (BPNN) is introduced to predict the PSPD after dumping, taking the loader dumping position, the digging payload and the current payload distribution inside the truck body as input parameters. Once this function is accomplished, the operators in the cabs of loaders or shovels can adjust the dumping position beforehand to reduce the probability of unbalanced payload distribution and improve the safety and reliability of the truck. The vision-based monitoring system forms one critical component for the "smart construction" system. The proposed method only requires a camera to be mounted above the truck cab for monitoring the dumped materials, and the acquired images and related extracted information can be transferred to the excavator cab through wireless networks. The trained model will provide the driver with information about the distribution of materials in the truck body.

The rest of this paper is organized as follows. Section 2 presents the related existing research works. Section 3 explains the predictive model proposed in this paper, including CNNs and BPNN. Section 4 details the experimental setup, task description, and experimental data. Section 5 discusses the experiment results and evaluates the performance of our model by comparing with other models. This is followed by the conclusion in section 6, summarising the advantages and disadvantages of the proposed model.

2. Related works

As mentioned, the method proposed here is to predict the payload distribution in haul truck body according to the current dumping conditions and the flow behaviour model of bulk material. Therefore, the state of the loaded material inside the truck body should be firstly estimated, as well as the interaction between the bulk materials in the shovel/loader bucket and that in the truck body during dumping. Then, machine learning, specifically deep neural networks, that are proved highly effective in extracting pivotal features, can be adopted for predicting the bulk material state. The main research topics involved in this process include online payload estimation, bulk material behaviour analysis, and machine learning techniques. This section will review related works in the three areas respectively.

2.1. Online payload estimation

Stentz A. et al. [4] introduced an excavator assistant system for loading haul trucks using two scanning laser rangefinders mounted on each side of the excavator's boom searching for the low region in the truck body to ensure the payload distribution to be balanced. However, the system does not consider the characteristics of physical interaction between the bulk materials and the rigid body. In addition, the lasers are also affected by dust and harsh environmental factors, which makes the system less reliable. Duff E. [5] proposed a method using two lasers mounted on a structure scanning the material pile in the truck body to estimate the volume of payload, when the truck is passing under the scanners. Although this method makes it possible to accurately evaluate the payload in the truck body, it can be functioned only after all dumping operations, which cannot meet the requirement for predicting payload distribution. Borthwick J.R. [6] developed a haul truck loading status estimation system using stereo cameras for payload volume evaluation. Since 10% of the truck

body is unavailable in the camera view field, the payload distribution in the blind area is interpolated from the observed surface. The performance using stereo cameras, in terms of accuracy and reliability, is prone to the harsh mining environmental conditions, which are dusty as well as under severe weather conditions.

Chamanara A. et al. [1, 2] developed an algorithm to model the cumulative shovel load passes within the truck body and to provide an appropriate location for successive dumping operations to achieve a balanced payload distribution before the truck leaves the digging face. This work assumes that the material can naturally form regular cone shapes, which are only suitable for sand and gravel with homogeneous particle size distributions. However, the blasted material in the mining site is highly discrete making the model unsuitable for field applications. A method using cluster analysis based on the Iterative Closest Point algorithm (ICP) was introduced by Bewley A. et al. [7, 8, 9] for measuring the volume of digging material in a dragline bucket through a range bearing laser for estimating the material's bulk density in real-time. Results show that the method can achieve an accuracy of above 95%. The application of the method is based on the hypothesis that the material pile in the bucket does not change in shape, which is obviously different from the problem addressed in our research. Huang X. [10] proposed a sensing system that can efficiently and accurately reconstruct 3D shapes under moving conditions based on a single laser distance sensor. However, the method is not suitable for the estimation of piles formed by bulk material.

Innes C. et al. [11, 12] provided a method for representing, tracking and fusing information on excavated material following its movement through a mining process chain. In the method, only the quality of the excavated material is tracked, while the distribution of the material is not involved. Dunbabin M. [13] used a laser scanner to measure the height of the piled-up material in the bucket during the excavation process estimating the fullness of the bucket combining the teeth forces and the height information. Sarata S. et al. [14, 15, 16] performed an on-line measurement of the loading volume of a wheel loader. The method firstly uses stereo vision to establish a 3D model of the large material pile, and then the loading volume is considered as the interference part of the bucket and the 3D model. Rasuli A. et al. [17] dynamically estimated the payload of a cable shovel based on the least square evaluation method and a simplified kinetic model. Anwar H. et al. [18] established an algorithm for estimating the volume of material in the bucket using 3D sensors based on the assumption that the material is evenly distributed in the bucket. The results of experiments prove the effectiveness of the algorithm. An online algorithm was proposed by Bi Q. et al. [19] for estimating the payload of cable shovels based on static analysis verified by scale model experiments. As can be seen, all those methods mainly focus on the value of payload, but not the distribution.

2.2. Bulk material behavior analysis

It is convincing to study bulk material behavior through experiments. The complexity of such bulk materials due to the high variety with the particle sizes or shapes makes it very difficult to simulate the mining operations of real-world experiments in various conditions. It is not possible and unrealistic to simulate all conditions due to the randomness of particle shapes and such a large number of experiments will take too much time with very low efficiency.

The DEM simulation provides convenient solutions for particle behavior modeling and analyses [20]. Many scholars have applied DEM to researches focusing on mining and earth-moving equipment.

Cleary P.W. [21] reviewed the application of DEM in the research topics of bulk material excavation, transportation, sorting, mixing, grinding and landslide, showing that the DEM simulation is an efficient method to understand the particle behaviors of bulk materials. Based on particle dynamics, a mathematical model of hopper discharge was established [22] and the influence of particle shape on the entire system was analyzed [23]. Coetzee C. et al. [24, 25] used DEM to establish a simulation model of the material fulfilling process of excavator, which was validated through experiments. The comparison results show that the error of DEM simulation is less than 20%. DEM is also used by Teufelsbauer H. et al. [26] to analyze the interaction between particles and obstacles validated experimentally in the laboratory environment. Lu G. et al. [27] found that the particle shape has a great influence on the static and dynamic properties of particle systems. Dynamic hopper discharge of irregular non-spherical particles was analyzed by Wang S. et al. [28], and the accuracy of the model was verified by the results of both experiments and simulations. However, most of the particles in the current studies are simplified based on analytical numerical studies that can be described with mathematical equations. Even the so-called irregular particles are just irregular balls or polyhedron, which are not strictly irregular. As can be seen in Figure 3, the shape of the blasted rock is almost impossible to describe by equations. In addition, the contact models applied in the DEM simulation are restricted, which cannot cover all the complex contact conditions of the rock materials in the mining site.

In DEM simulations, computational efficiency is a pivotal performance indicator. In order to speed-up the calculation, Lommen S. et al. [29] proposed a reduced stiffness method. Gopalakrishnan P. et al. [30] proposed a parallel computing method. Similarly, Gan J. et al. [31] used GPU for improvement for computation efficiency and Yue X. et al [32] utilized parallel computing and GPU simultaneously for acceleration. To simulate the macroscopic particle behavior of materials, the number of particles involved in the simulation may exceed 10 millions. Meanwhile, in order to ensure the stability of the simulation, the time step has to be set to a small value, which leads to

huge computational demands [33]. Therefore, even with parallel computing or other efficient algorithms, it is unrealistic to achieve real-time PSPD with DEM simulation on a reasonable hardware system for real mining operations.

For the purpose of real-time prediction, Hagenbuch L. [34] assumed that loose bulk material dumped on a flat surface will form a tapered or elongated cone, and Joseph T. et al. [35] considered that the piled-up status of the material in the compartment was formed by the intersection of cones limited by the shape of the cabin. Errors of the method are considerably small when the particles are uniform that there is no need to utilize DEM or other algorithms that consume huge amounts of computational resources. But the accuracy of the method needs to be verified for the bulk materials with different sizes and shapes.

2.3. Machine learning for handling bulk material

As mentioned before, the capability of predicting the PSPD of bulk material can be learned and improved through training. In recent years, deep learning has been used widely and successfully in various fields. Schenck C. et al. [36] used deep learning to control the robotic arm to form bulk materials into a pre-set shape. The application process includes the determination of the current material piled-up profile and the prediction of its shape after the operation of dumping. Although the parameters involved in Schencks study are not the same as what we need, it is still with great value for our research. However, to obtain accurate results, the sensors must be able to output depth information with high accuracy for the point cloud measurement used as the main source in Schencks study. Besides, the robot applied for material handling needs 6 elements for motion planning, which are much more complicated than that for earth-moving equipment such as loaders or shovels.

Takei T. et al. [37] applied the optimization method using a genetic algorithm to the path planning for wheel loaders on scraping and loading operations. Rozo L. et al. [38] developed a parametric hidden Markov model based on force feedback to teach robots of pouring skills. Yamaguchi A. et al. [39] proposed a differential dynamic programming method based on the directed graph structure, which can pour fluid-like materials with different properties into containers forming different shapes.

Zhao Z. et al. [40] used the Gray model data as input and seed distribution as output to train the BPNN, which can predict the relationship between planter vibration parameters and the seed distribution in the rectangular vibrating tray. Kumar R. et al. [41] used ANN to predict the relationship between key parameters of the material handling system (inner angle, bulk density, average diameter, friction coefficient, et al.) and discharge efficiency. Rabault J. et al. [42] applied an ANN training method through a deep reinforcement learning agent to perform active flow control.

Benvenuti L. [43, 44] and Ye F. [45] adopted ANN and BPNN to predict the relationship between particle contact parameters and the repose angle for DEM simulations. Chen S. et al. [46] established a CNN model with parameters of particle residence time, diameter and density as inputs, and classification of particle height in the initial packing as output. Wan Z. Y. et al. [47] combined a fluid mechanics model with machine learning to propose a method describing the motion state of spherical particles in fluid. Goldstein E.B. et al. [48] predicted the settling velocity of frictionless particles using a machine learning approach based on the genetic programming. Clarke S. [49] proposed a learning method using audio-frequency vibrations from contact events to estimate the flow rate and mass of granular materials during scraping and loading tasks.

3. Method

In view of the extreme complexity of the working environment and operational requirements of the loaders and haul trucks, several coupling factors affecting the bulk material behavior impede the establishment of an accurate theoretical and numerical model for describing the material handling operation. During the digging and dumping process, the loader-truck system, as a typical complex electromechanical system, would generate a large amount of noises that greatly interfere the signals, which can be used for payload distribution estimation. Compared with other methods, the vision technology applied in the mining industry is relatively easy to obtain with high stability. Therefore, it is considered more reliable to use images than other sources for prediction of payload distribution. With the powerful capabilities of characterizing complex systems, processing big data, and automatically extracting features, deep learning has feasibility and superiority in the prediction task. Therefore, this paper trains the deep learning model by images supplemented with a small amount of data of material quality and coordinates to predict the PSPD in the truck body after the current dumping operation.

3.1. Proposed framework overview

To achieve the dynamic prediction of the PSPD of the bulk material in the truck body before each dump, this paper presents a synthesis model by using a deep learning-based forecasting framework that integrates CNN and BPNN. Figure 4 shows the flowchart of this framework, which involves 3 parts.

Part I Data pre-processing. For image data, the region of interest (ROI) is firstly extracted and the data outside the ROI is deleted to purify the image to improve the regression accuracy of the CNN, and then perform a grayscale process to the images to reduce the consumption of computing resources. For the quantity of material in bucket and the dumping coordinates, the average value can be obtained after eliminating the zero drift and the

filtering process, as shown in *Part I: Data Pre-process* in Figure 4.

Part II CNN classifier and regressors. CNNs, which have deep architectures with multiple convolutions and pooling layers, and the abilities of automatically extracting pivotal features of data, are applied in this study. We will take the advantages of these features of CNN to estimate the PSPD in the current state based on the processed image data, as shown in *Part II: CNNs* in Figure 4.

Part III BPNN regressor. In order to train the BPNN to accomplish the prediction of PSPD, the results obtained in Part II and the processed data in Part I are taken as input parameters with the predicted PSPD after dumping as output, as shown in *Part III: BP Neural Network* in Figure 4.

The detailed approach of each part is discussed as follows.

3.2. Data pre-process

Considering the interference of dust and light in the mining sites, the background segmentation and grayscale conversion are needed on the images in this work to improve the accuracy and the calculation efficiency of the prediction. Due to the identifiable chromatic aberration between the truck body, dumped material, and the background, a trapezoidal-shaped mask can be used to effectively identify the ROI, while the regions outside of the ROI are removed. The original RGB images are converted to grayscale by the weighting method. The weighted gray value $I(i,j)$ of the pixel (i,j) can be calculated according to Equation (1) based on the red component $R(i,j)$, the green component $G(i,j)$, and the blue component $B(i,j)$:

$$I(i,j) = 0.30R(i,j) + 0.59G(i,j) + 0.11B(i,j) \quad (1)$$

It should be noted that the weights (0.30, 0.59 and 0.11) in Equation (1) were determined empirically after testing multiple combinations of parameters. In the subsequent training processes, the efficiency was found improved by about 3% compared to using the default parameters provided by MATLAB®. However, there is no significant difference in terms of prediction performance between the trained models with different parameters.

Regarding the mass of the material, the measurement feedback from the load cells would vary over time, presenting the data drifting problem. It is necessary to eliminate the zero drift before any applications. Then the data reflecting the PSPD of the material can be obtained after the process of filtering and averaging.

For the design of the experiments, 9 different candidate dumping positions of the scaled bucket are predefined with respect to a fixed position of the cabin representing the truck body. However, due to the influence of the inertia of the robot arms and the payload distribution status in the bucket, the coordinates of the dumping positions would change to a certain degree. Considering the movement of dumping is performed by a KUKA LBR iiwa robot, which cannot accurately simulate the motion of the z-bar linkage of a loader, the dumping position is defined as the average value of the center of gravity of the cabin with the dumped

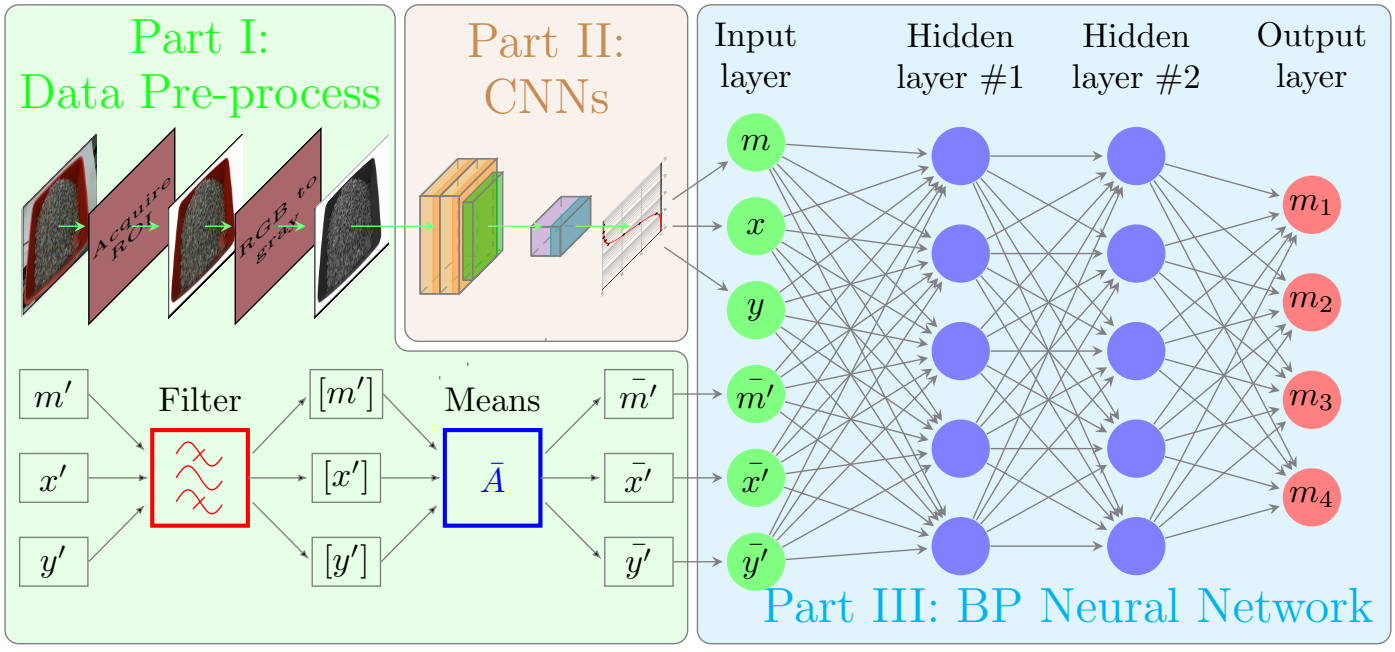


Figure 4: The model presented in this paper.

material after every first dump of all the cycles.

$$(X_i, Y_i) = \frac{1}{N_i}\left(\sum_{j=1}^{N_i} x_i^j, \sum_{j=1}^{N_i} y_i^j\right) \tag{2}$$

where, the subscript $i\,(i = 1, 2, \cdots, 9)$ indicates the values of the ith dumping point, (X_i, Y_i) represents the average value of coordinates of the ith dumping point, N_i indicates the total number of cycles in which the first bucket of payload is dumped at the ith dumping point, $\left(x_i^j, y_i^j\right)$ represents the coordinates at the jth cycle of the ith dumping point.

3.3. CNNs

Despite the great success of CNNs for deep learning tasks with images as input, high-accuracy results can usually only be obtained with the availability of massive data for training the network. The data set in this work will be limited for training a deep CNN from scratch. Therefore, it is necessary to take advantage of the pre-trained deep CNNs. A feasible way to use deep CNNs is to conduct transfer learning. To be more specific, a pre-trained model that has been trained with massive data is applied to specific problems trained with dedicated data, which share some commonality at low-level features with the massive data, and a small learning rate. The commonly used deep CNNs include AlexNet, Densenet201, Googlenet, Inception v2, Inception v3, Resnet101, Resnet50, Resnet18, VGG16, VGG19, and so on. Generally, these networks are designed for identifying and classifying targets, while the goal of our study is to perform regression. In addition, the low-level features of the data sets (CIFAR-10/100 or ImageNet) used to train these networks are very different from that of the images applied in this paper. Hence, it is difficult to accomplish the purpose of this study by simply applying the method of transfer learning.

Based on the above reasons, we proposed a two-stage CNN prediction method, including a classifier and some regressors, as shown in Figure 5.

Step 1: Classification task. To increase the training efficiency and estimation accuracy of PSPD and the material mass in the cabin, we divide the images into several different categories based on their weights. For that purpose, 6 sets of images are prepared out of the total images, according to the mass intervals of material as follows: $(0, 1]$, $(1, 2]$, $(2, 3]$, $(3, 4]$, $(4, 5]$, $(5, 7]$, where $(0, 1]$ represents the mass that fall between 0 and 1kg. The highest weight in the cabin is 6.68kg. There are very limited data with only 46 groups that are above 6kg in the validation set. So the data above 6kg is combined into the intervals of $(5, 6]$ to form $(5, 7]$. The images can be classified according to the mass of material that can greatly reduce the variation of the target for the regression task and improve the prediction accuracy.

Step 2: Regression task. In this task, we introduced a compact neural network for predicting the mass and PSPD. Six networks of the same architecture are trained separately for different categories according to the output from Step 1 above. The regression target here contains two outputs, which are the total payload and the coordinates. A single universal network can be deployed for this work. Unfortunately, the results are not as satisfactory as using two separate networks. Therefore, we propose a framework comprising two separate networks for predicting the mass of the material in the cabin and the coordinate of its centre of gravity.

An alternative option is to design a network for estimating the mass directly in continuous space. However, correctly estimating the mass in continuous space from images would be difficult using our limited data. For simplicity, we used a classification network as described above. Considering that the number of prediction regression networks will change correspondingly with respect to the number of classes, the number of classes cannot be too large, as that would introduce extra computational burden. Also, as mentioned, the number of images in each of the intervals needs to be sufficient for training. In this work, empirically, six was found to be a well-balanced decision with our collected data.

3.3.1. Classification

The performance of those popular networks listed in Section 3.3 is evaluated for the classification task in terms of several performance metrics, including accuracy, resource consumption, etc. In order to make the networks applicable to the problem addressed in this article, the last fully connected layer of each model is deleted and replaced by a 6-output fully connected layer. Other layers remain unchanged with their structures, parameters, and weights. When performing the training process, the multi-class cross-entropy is taken as the loss function,

$$Loss = -\sum_i \ln(\hat{x_i}) \tag{3}$$

where $\hat{x_i}$ is the probability that the prediction being classified as the ith class.

Accuracy is employed as one performance metric, defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

where TP, FP, TN, and FN denote the true positive, false positive, true negative, and false negative respectively.

The training processes are shown in Figure 6, and some result data are listed in Table 1.

Since the entire model is connected in serial with multiple networks, the accuracy of the entire model will be lower than that of any sub-network. The accuracy of the entire model will, therefore, be dependent on the accuracy

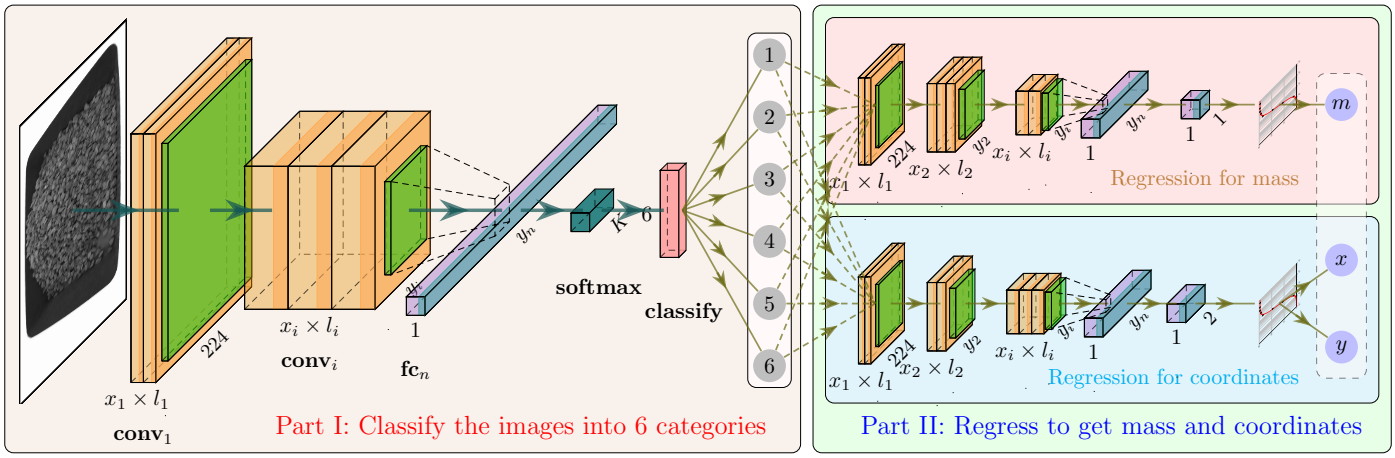


Figure 5: Schematic of the proposed CNNs: The figure does not reflect all details of the actual network and is only to present the overall architecture for illustration purposes. The details will be described later.

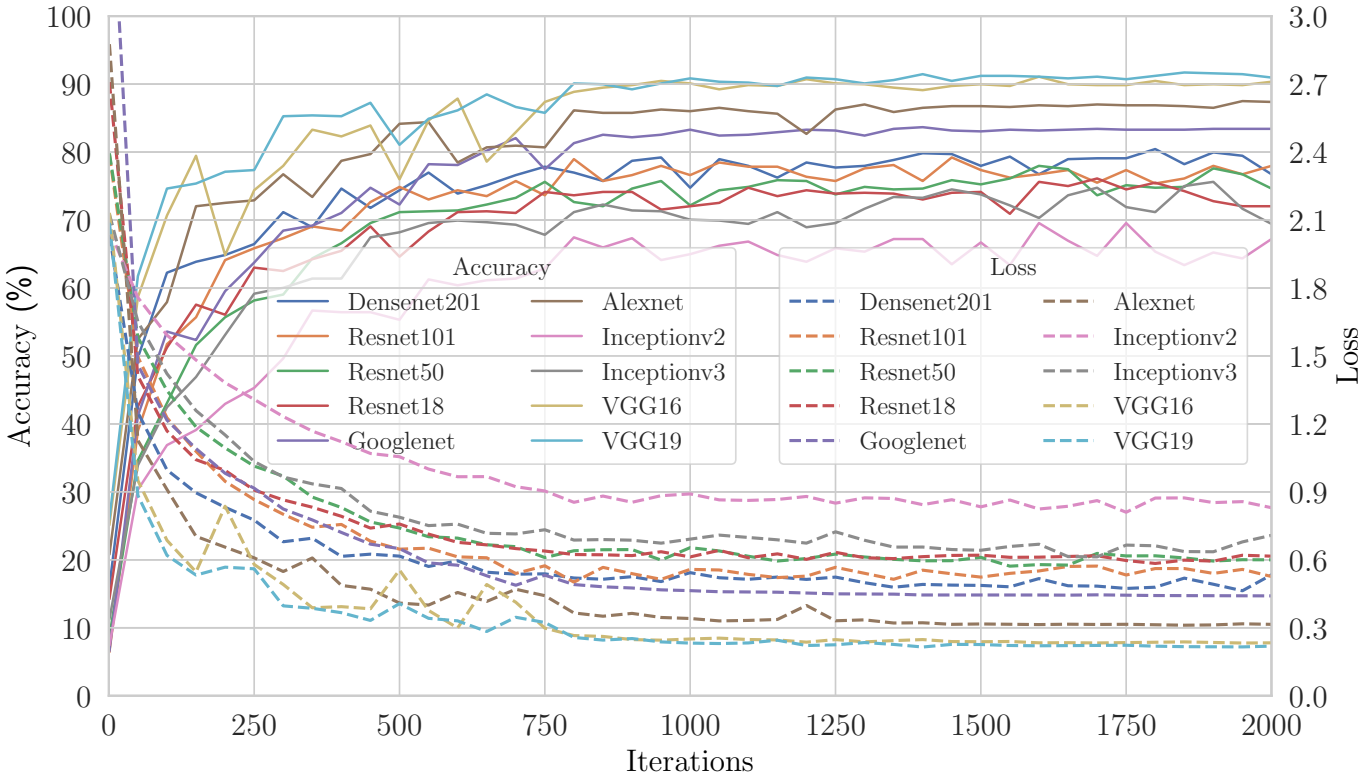


Figure 6: Accuracy and loss of popular networks in the classification task.

of each sub-network, which needs be as high as possible in order to reach a satisfying level. Since the subsequent regression networks are designed separately for each class, the prediction performance will be poor for mis-classified samples. However, practically, the images in two adjacent intervals, e.g. (1,2] and (2,3], would appear more similar than with more isolated intervals, e.g. (4,5]. The impact of misclassifications between adjacent intervals would be limited on the PSPD prediction networks.

As can be seen from Table 1, VGG19 performs best compared to other networks, so it is adopted as the basis for transformation. The modified classifier of the VGG19, as shown in Figure 7, consists of 47 layers, including 1 input layer, 16 convolutional layers, 5 pooling layers, 18 active layers, 3 full connect layers, 1 softmax layer, 1 normalization layer, and 2 dropout layers with the probability of 0.3. The sizes of the convolution kernels can be found out from the figure.

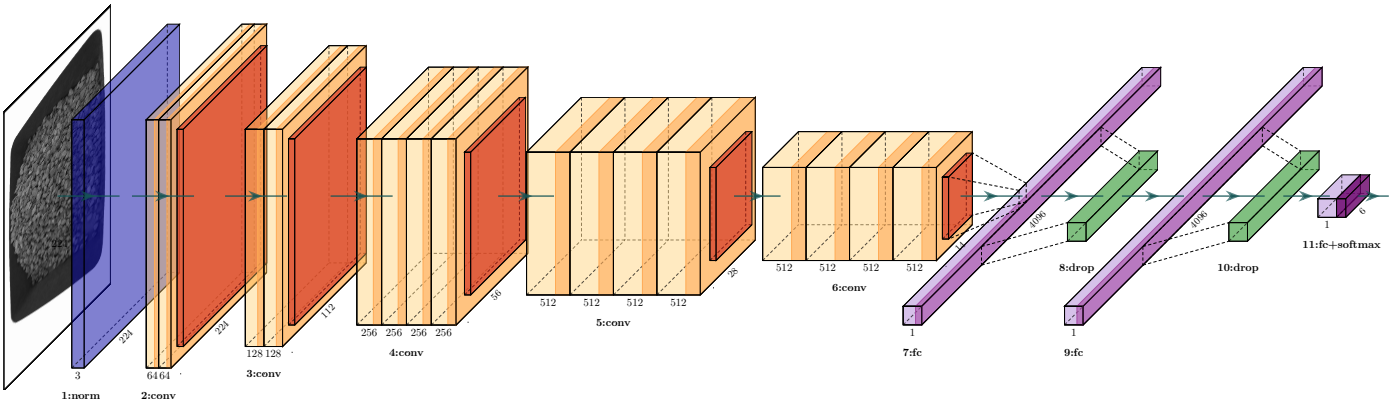


Figure 7: Network architecture for classification task. Due to the limited space, the layer details are abbreviated as follows: normalization layer as norm, convolutional layer as conv, fully connected layer as fc, dropout layer as drop, and softmax layer as softmax. There are a relu layer after each convolution layer and 2 fully connected layers, #7 and #9. The orange rectangles after the convolutional layer groups depict the pooling layers.

Table 1: Important results of popular networks in the classification task

Base model	Accuracy (%)	Loss	Time cost (s)
VGG19	91.34	0.2167	1159
VGG16	89.73	0.2366	984
Alexnet	86.63	0.3128	116
Googlenet	83.42	0.4396	491
Densenet201	84.53	0.4995	5981
Resnet101	83.17	0.5470	2001
Resnet18	79.46	0.6253	274
Resnet50	81.68	0.6324	912
Inception v3	81.56	0.6345	1832
Inception v2	72.15	0.8462	6739

3.3.2. Regression

The proposed regressor in this paper is based on VGG19, as shown in Figure 8, consisting of 41 layers, including 1 input layer, 12 convolutional layers, 5 pooling layers, 15 relu layers, 5 full connect layers, 1 regression layer, 1 cross channel normalization layer, and 1 dropout layer with the probability of 0.7. The sizes of the convolution kernels can be found out from Figure 8.

The continuous form of cross-entropy is used as the loss function for the training of the regressors. The parameters of $\{(d_0, z_0), (d_1, z_1), \ldots, (d_{N-1}, z_{N-1})\}$ are used to represent the values of image-label pairs and f for the regression function. For any image-label pair (d_i, z_i), the probability $\Pr(\hat{z}_i)$ can be derived through the regression function f. Since the true value of z_i changes continuously within a certain range, $\Pr(z_i)$ follows a continuous probability distribution, that is, the cross-entropy between $\Pr(z_i)$ and $\Pr(\hat{z}_i)$ should be expressed in a continuous form as:

$$H_i = \int_{-\infty}^{+\infty} \delta(z - z_i) \ln \Pr_z(\hat{z}_i) dz = \ln p(z_i|s) \tag{5}$$

where $p(z_i|s)$ represents the distribution of $\hat{z}_i$, which follows a Gaussian distribution if function f was linear regression with the parameter s. $\delta(\cdot)$ is the probability density function of z_i which is a Dirac delta function.

Therefore, the overall cross-entropy loss function would be expressed as a sum over all H_i:

$$H = \sum_{i=0}^{N-1} H_i = \sum_{i=0}^{N-1} \ln p(z_i) \tag{6}$$

Since $p(z_i|s) \propto \exp \frac{(z_i - \hat{z}_i)^2}{2s^2}$, the sum of cross-entropy would be proportional to $\sum_{i=0}^{N-1} (z_i - \hat{z}_i)^2$, which is equivalent to least square.

This work uses RMSE as the criterion for evaluating the quality of training results:

$$RMSE = \left[\frac{1}{N} \sum_{i=0}^{N-1} (z_i - \hat{z}_i)^2 \right]^{\frac{1}{2}} \tag{7}$$

3.4. BPNN

The data for the network are relatively simple, comprising 6 inputs and 4 outputs, a fully connected BPNN with two hidden layers is proved to be effective and accurate. Figure 4 in *Part III: BP Neural Networks* depicts the network architecture.

4. Experiments

Machine learning has been successfully applied in many fields. One common characteristic for these applications is the availability of massive labelled data for training the neural networks. Also, the labelled data are mostly collected from controlled working environments, such as laboratories, production assembly lines and structured urban transportation networks. However, the environmental conditions for mining trucks are much more severe than the laboratory environment. Factors such as vibrations, dust, and lights may cause failures in field experiments. Also,

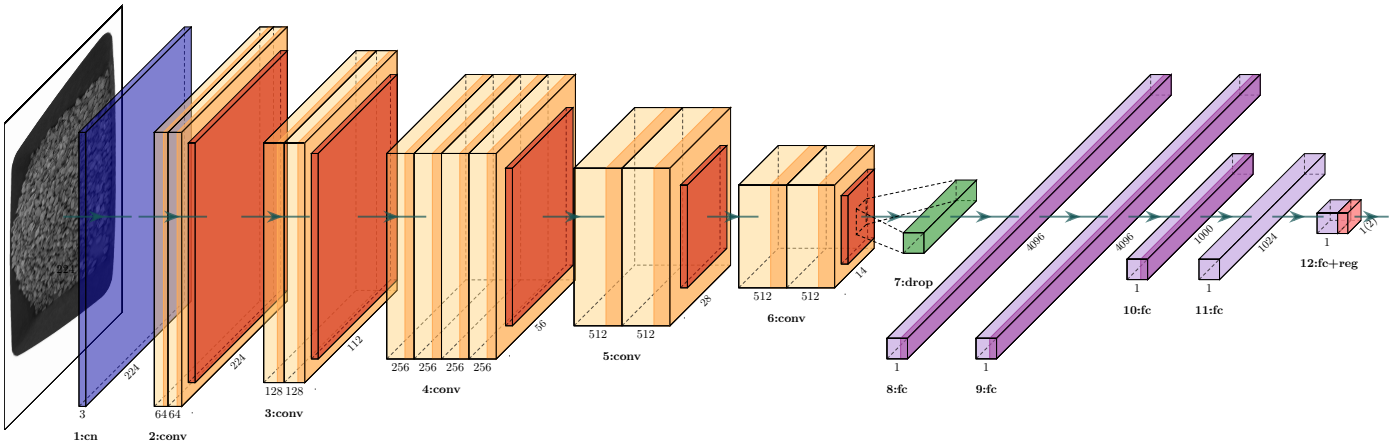


Figure 8: Network architecture for regression task. Abbreviations are used in the diagram to save space, where c-n represents a cross channel normalization layer and reg denotes a regression layer. There are a relu layer after each convolution layer and several fully connected layers (#8, #9, #10). The *1(2)* under the last arrow indicates that the structure can be used for both mass (1) and coordinates (2) prediction. Other instructions are consistent with Figure 7.

labelled data for the real-world environment are very limited. Therefore, it is difficult to understand the causes of the failures or inaccuracies with the prediction results that may come from the prediction model or/and the extreme environmental conditions. Repeated experiments potentially can be applied to minimise the influence of the environmental factors. However, it is obviously not an economical or practical method to pause the mining chain simply for the validation of the payload prediction algorithm. It would be desirable to only perform large-scale field experiments after the framework's validity has been fully verified in laboratory.

Therefore, this work designs a series of laboratory experiments for validation of the proposed prediction framework.

4.1. Experiment setup

The physical setup for the experiment is shown in Figure 9. Two LBR iiwa 14 R820 type robots are applied for the experiments. A custom 3D-printed plastic bucket mounted to the end-effector flange of one robot to simulate the dumping operations of a loader. A red box simulating the truck body is put between the dumping robot and the camera robot. Four load cells are assembled at four bottom corners of the red box for the payload description inside the box after each dump. The specific parameters of the load cells are shown in Table 2. The data logger USB-1608FS-Plus is applied to record the data acquired by the load cells. The mass of the material contained in the bucket can be determined by the difference between the sensor data obtained by two consecutive dumping operations. As can be seen from Figure 9, a black box containing irregularly shaped limestone of about 10mm in size is placed on the side of the dumping robot, simulating the material pile in the mining site. A USB camera is mounted to the end-effector of the other robot and calibrated with respect to the world coordinates. The camera faces down towards the red box to capture RGB images of the surface in different positions. The data logger and the camera communicate with the system through USB respectively, while the robots are connected via TCP/IP. MATLAB® is applied to automate the process of data collection by controlling the two Kuka robots that greatly accelerated the process of data acquisition. The MATLAB® toolbox KUKA Sunrise is used to control the robot and log position data.

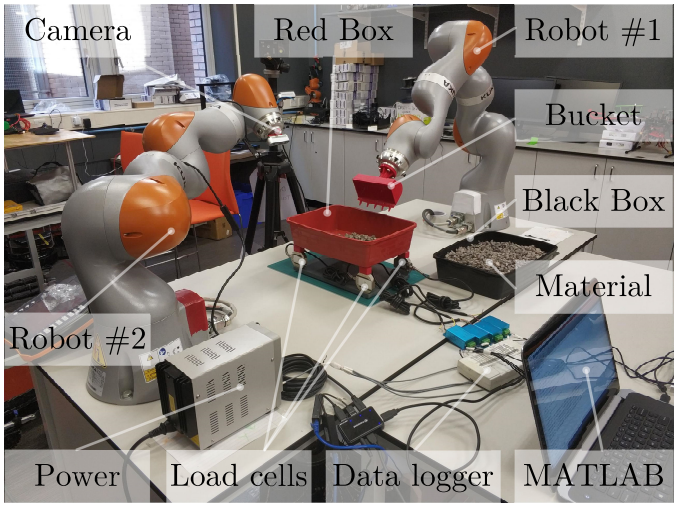


Figure 9: The physical setup for the experiment

4.2. Task specifications

The experiments contain the following main steps:

Step 1: Set experimental parameters.

- Put the empty red box onto the load sensors and then clear the initial data of the load sensors.

Table 2: The specifications of the load cells

Parameter	Value	Units
Rate Load	5000	g
Null Balance	± 1	%F.S
Sensitivity	2.0 ± 0.05	mV/V
Non-Linearity	± 0.03	%F.S
Hysteresis	± 0.03	%F.S
Repeatability	± 0.03	%F.S
Recommend Excitation	$5 \sim 15$	V
Input Resistance	350 ± 20	Ω
Output Resistance	350 ± 5	Ω
Insulation Resistance	2000	MΩ
Safe Load Limit	7500	g
Total Error	0.03	mV/V

- Select a predetermined dumping position relative to the black box.
- Set the relative positions of the camera and the red box.

Step 2: The robot excavates the material from the black box. Practically, operators are usually required to dig up as much material as possible (fullness is about 90% - 110%). If the digging conditions can be fully controlled, the material mass of each bucket would be consistent and can be described as within a natural variability. However, the uncertainty of the material pile will encounter difficult working conditions. For example, stones with irregular shapes will lower the fullness to less than 50%. On the other hand, under certain conditions, material above the boom may "collapse" into the bucket, increasing the fullness level to above 130%. Therefore, the variability of material mass will be large. To simulate the phenomenon of large-scale variation, the material mass of each excavation in this work was deliberately prepared within a large variation.

Step 3: Move the robot to the unloading position and keep still for a short time (0.5s) to stabilize the bucket and material to eliminate the effects of inertial motion, while allowing load cells to record stable data.

Step 4: Dump the material into the red box, and then return to the initial position where the material is excavated. The camera takes pictures at the predetermined positions and the load cells record the data again while the robot moving away.

Step 5: Repeat Step 2 to Step 4 until the material loaded in the red box reaches a certain amount, and then dump the material back into the black box after the measurement.

Step 6: After obtaining a sufficient amount of data (see section 4.3), change the dumping position relative to the black box, and then repeat Step 1 to Step 5.

The red box coordinate origin O is defined at the position of the load cell at the lower-left corner, where the horizontal axis is x and the vertical axis is y. The load cells at each corner are numbered counter-clockwise from the origin of the coordinate frame. The measured value of each load cell is represented by m_i, where $i = 1, 2, 3, 4$, and the mass of the red box is m_h, Therefore, the centroid of the material can be projected on the xOy plane at the following coordinate:

$$\begin{cases} x_{\mathrm{COG}} = \dfrac{m_2 + m_3 - \frac{1}{2}m_h}{\sum_{i=1}^{4} m_i - m_h} \\ y_{\mathrm{COG}} = \dfrac{m_3 + m_4 - \frac{1}{2}m_h}{\sum_{i=1}^{4} m_i - m_h} \end{cases} \tag{8}$$

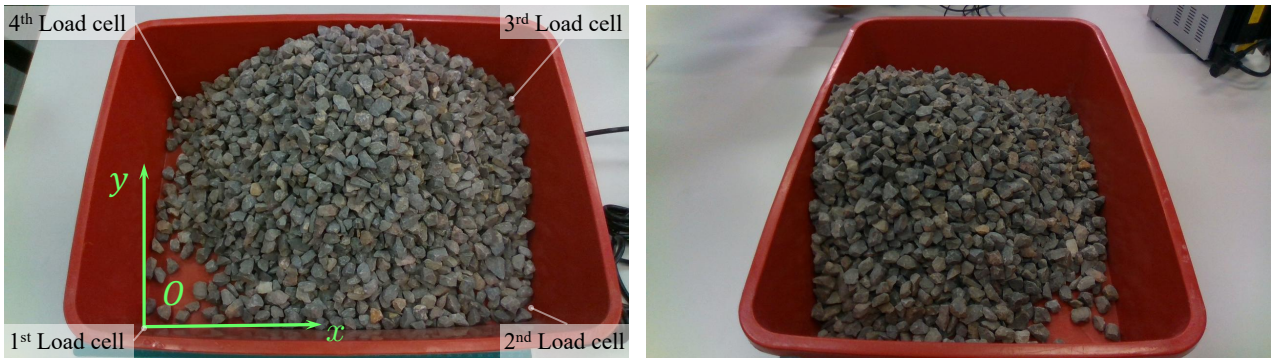
4.3. Experiment data

We predefined 9 dumping positions and conducted about 1000 experiments at each position, totaling 8840 sets of valid data (each set contains the images and corresponding sensor data). The experiments were conducted iteratively. Initially, about 500 sets of data were acquired at each position. We then evaluate the performance difference by adding more data iteratively. We collect 100 sets of data per round and this process stops until no noticeable performance improvement. Figure 10 shows some acquired images, and the corresponding sensor data are shown in Figure 11.

As can be seen from Figure 11, the changes of the sensor data are relatively smooth. It should be noted that although the zero line and the data are plotted in the same graph, they are collected during different time periods, where the zero line data are acquired before the start of each dump (when the red box has not been placed on the load cells) and the load data are collected after dumping. In addition, the relative values of the load data for each load cell appear consistent with the material distribution in the red box shown in Figure 10, where $m_3 > m_4 > m_2 > m_1$.

The acquired data include the current distribution image of the material in the red box, the current mass of material in the bucket, the dumping position, and the PSPD of the material in the red box after the dumping operation. The first three items of the data are set as input for the predictive model, and the last one is used as the label, providing the labelled training data.

Statistical analysis of the material mass in the bucket is shown in Figure 12, which is subject to the Normal distribution $\mathcal{N}(0.4379, 0.06866)$, with $RMSE = 2.92 \times 10^{-3}$kg, and $SSE = 3.15 \times 10^{-4}$.



(a) Images took from the oblique upper view angle of the longer side

(b) Picture took from the oblique upper view angle of the shorter side

Figure 10: Pictures at different locations in the same unloading pass: (a) Picture took from the oblique upper view angle of the long side, and (b) of the short side.

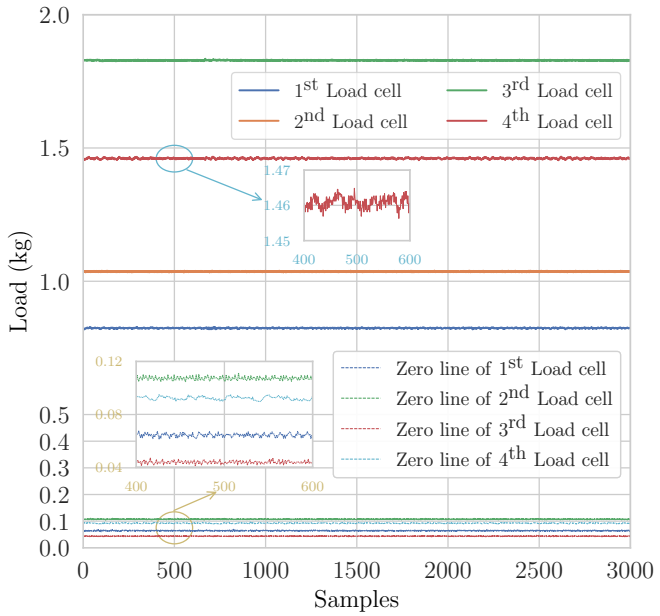


Figure 11: Data obtained from load cells : zero lines and loads varying with time.

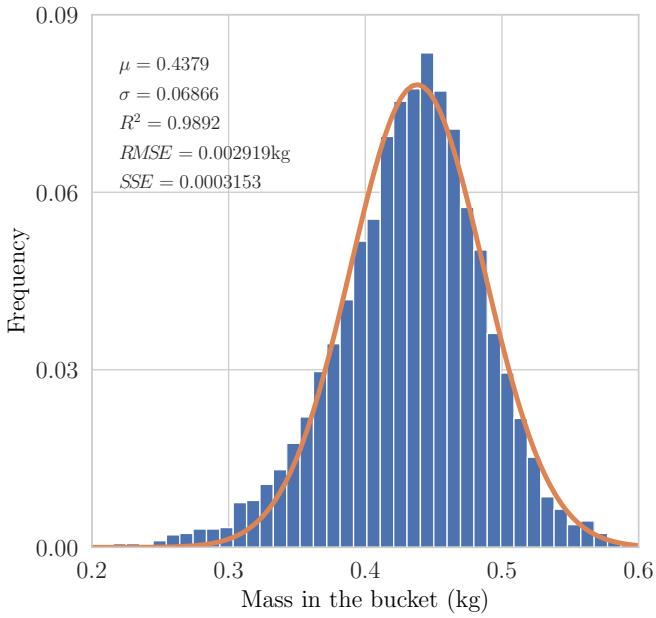


Figure 12: Quality of material in bucket subjects to Normal distribution.

5. Results and discussion

5.1. Prediction for the experiment data

5.1.1. Classification

Figure 13 shows the training results in terms of accuracy and loss of the proposed network (see Section 3.3.1), where the hyper-parameters are empirically decided based on common values that have been extensively used in other literature, where the learning rate is 10^{-6}, the batch size is 20 and the number of epochs is 6. It can be seen from the training and validation results that the network does not appear over-fitting or under-fitting, implying that the model has a good generalization ability. The final accuracy is 97.26% and the loss is 0.082, which are considerably better than the performance of the original networks listed in Table 1.

5.1.2. Regressors

According to Section 3.3.1, all data (images) are divided into 6 groups according to the weight range that are used as the inputs for training six networks individually. Meanwhile, the data obtained from the load cells at the time, when the images were captured, are used as the output to train the networks. In the work, we set the learning rate as 10^{-6}, the batch size as 20 and epochs as 6, the training results are shown in Figure 13. Correspondingly, the RMSE and loss for each range are listed in Table 3. *Relative RMSE* and *Relative Loss* in Table 3 are the ratios of RMSE values with respect to the corresponding median

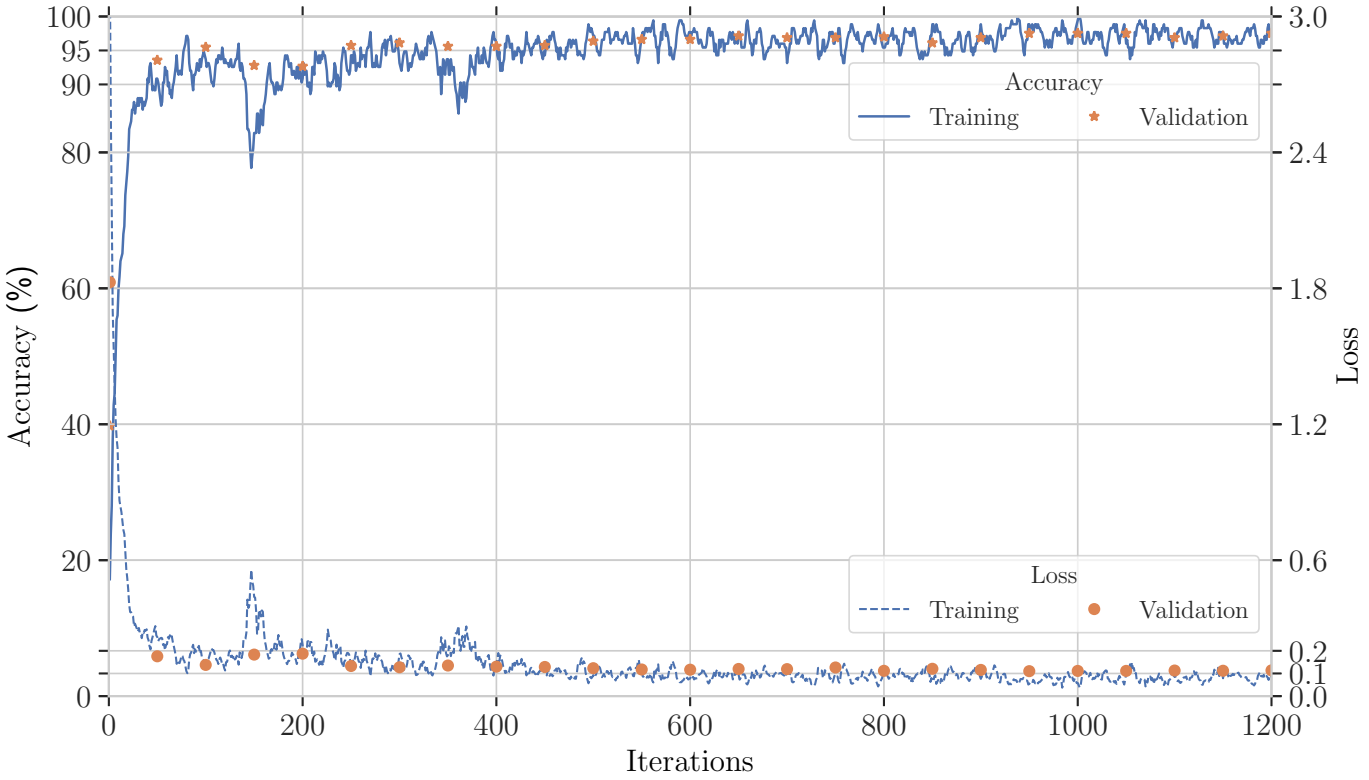


Figure 13: Accuracy and loss of the proposed classifier.

Table 3: RMSE and loss of mass regression for each range

Range(kg)	RMSE(kg)	Relative RMSE	Loss	Relative Loss
$(0, 1]$	0.0486	0.0972	0.00118	0.0024
$(1, 2]$	0.0737	0.0491	0.00271	0.0181
$(2, 3]$	0.0935	0.0374	0.00438	0.0018
$(3, 4]$	0.1515	0.0433	0.01148	0.0033
$(4, 5]$	0.1568	0.0348	0.01230	0.0027
$(5, 7]$	0.1934	0.0322	0.01871	0.0031

values. Take the data in the first row as an example, *Relative RMSE* (0.0972) equals to divide RMSE (0.0486kg) by the median (0.5kg) of the range $(0, 1]$.

It can be seen from Figure 14 that with the increase of mass of material, the absolute value of RMSE predicted by the model increases, while the relative RMSE gradually decreases, indicating that the regression performance of the model has improved slightly with the increase of the mass of material. The possible reason is that due to the limitation of the red box simulating the truck body. The viscosity or granularity characteristics of such materials are different at different layers, while the bottom layers would spread slightly flatter, showing less variations in the weight distribution. When the payload exceeds a certain amount, the bottom part of the material pile cannot be changed while the viscosity of the newly dumped material is reduced. However, when there is only one dump of material inside the red box, the movement of the particles would be less affected by the sides of the container. The

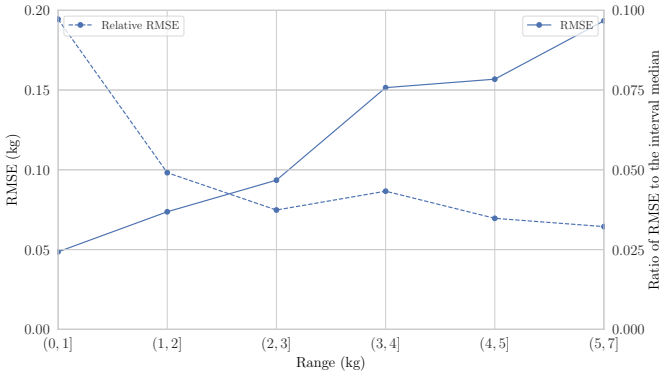


Figure 14: RMSE and relative RMSE of regressed mass varying with the material weights.

flowing behavior of the material would show greater uncertainty in that case. Therefore, in such circumstances, even if there is only a small change in the surface shape of the material, the mass of the material might be greatly different.

Training under the same hyper-parameter settings, the RMSE and loss for coordinates regression of the gravity centers of the payload are listed in Table 4. The curve in Figure 15 shows the RMSE of the regressed coordinates and it can be seen that there is no obvious correlation between the performance of the model and the mass of material.

Table 4: RMSE and loss of coordinates regression for each range

Range(kg)	RMSE	Loss
$(0, 1]$	0.0217	0.000236
$(1, 2]$	0.0226	0.000256
$(2, 3]$	0.0181	0.000164
$(3, 4]$	0.0222	0.000246
$(4, 5]$	0.0218	0.000239
$(5, 7]$	0.0337	0.000566

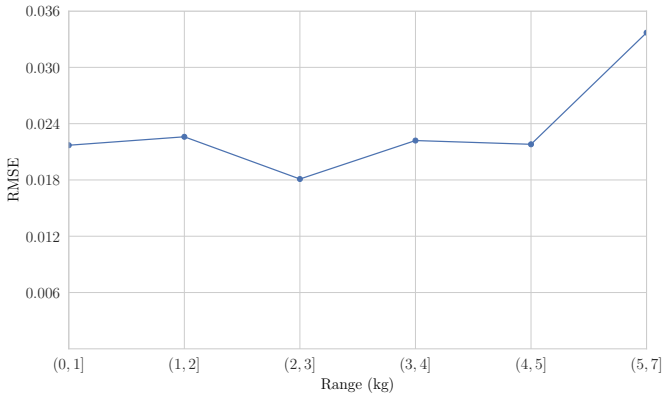


Figure 15: RMSE of regressed coordinates varying with the material quality.

5.1.3. Results of BPNN

The RMSE of the training process are shown in Figure 16. It can be seen that RMSE reaches a relatively low level in a short time. Although there is a certain difference between the training and validation, the absolute difference at the final stage is negligibly small with the average value of the material mass in the red box, proving that the model provides satisfactory results for the application in terms of generalization. It should be noted that the input data used for training the BPNN is the data obtained from the load cells, which serve as the ground truth of the corresponding images, not the output predicted by the model established in Section 3.3.

The predicted results are shown in Figure 17 and Figure 18. As can be seen in Figure 17, the predicted and the target values have a linear relationship with a slope close to

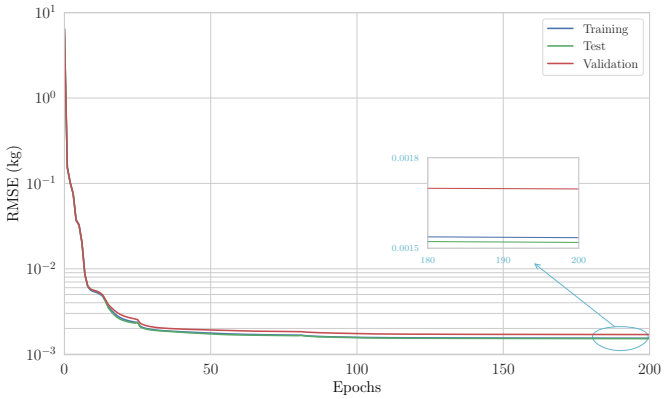


Figure 16: RMSE of the BPNN varying with epochs.

1 (the maximum is 0.99414 and the minimum is 0.99127), and the intercept close to zero (the minimum is 0.00302 and the maximum is 0.00674). Figure 18 shows that the probability density of the prediction error on the entire data set is approximately subject to the Cauchy distribution with the fitted position parameter x_0 close to zero, with the scale parameter γ around 0.015, indicating that the prediction errors are negligible.

5.2. Comparison with other models

For comparison purposes, we transform some popular networks, with minimum changes, in order to complete the task of the CNNs (as specified in *Part II: CNNs* in Figure 4), by 1) removing the softmax layer and the classification layers behind the last full connect layer of each network and then 2) adding a 3-output full connect layer with the same parameters as the last full connect layer, while keeping the structure, parameters, weights, etc. of the other layers. The training process is shown in Figure 19, and the results are listed in Table 5. After that, the trained results of CNNs, the mass of the material in the bucket and dumping coordinates, are used as the input data of the trained BPNN to predict the final PSPD.

Table 5: CNN results of the most popular networks.

Base model	RMSE	Loss	Time elapsed
VGG19	0.0800	0.00321	306min56sec
Inception v2	0.0817	0.00334	997min23sec
VGG16	0.1036	0.00536	242min30sec
Googlenet	0.1333	0.00888	74min08sec
Inception v3	0.1501	0.01126	785min45sec
Resnet18	0.1708	0.01459	52min31sec
Resnet101	0.1807	0.01633	344min32sec
Resnet50	0.2321	0.02693	169min08sec
Alexnet	0.2415	0.02915	74min10sec
Densenet201	0.2706	0.03662	834min04sec

It can be seen from Figure 20 that a few networks, including our proposed model, VGG16 and VGG19, perform

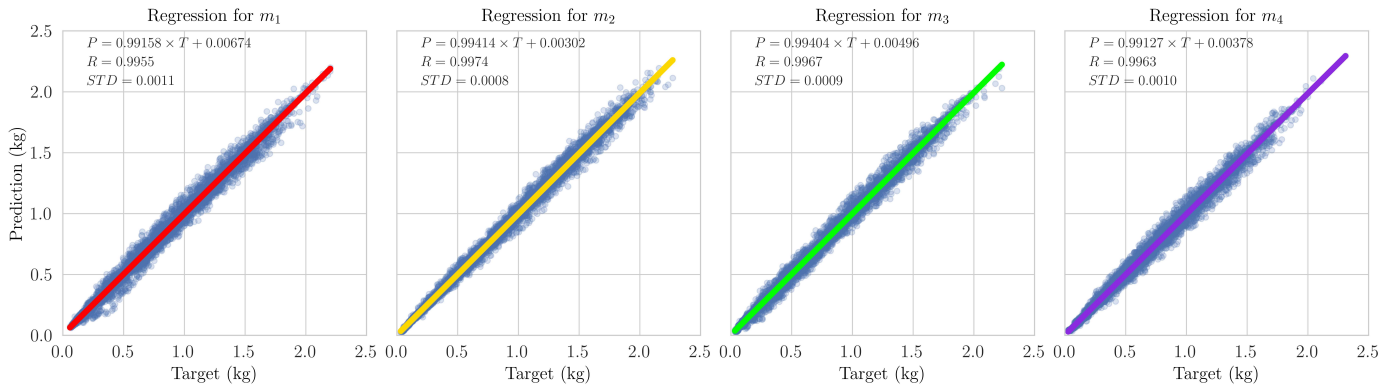


Figure 17: Relationship between the predicted and the target values. From left to right are the corresponding results of m_i $(i = 1, 2, 3, 4)$. The equation obtained by linear fitting, the correlation coefficient R, and the standard deviation STD are shown at the upper left corner.

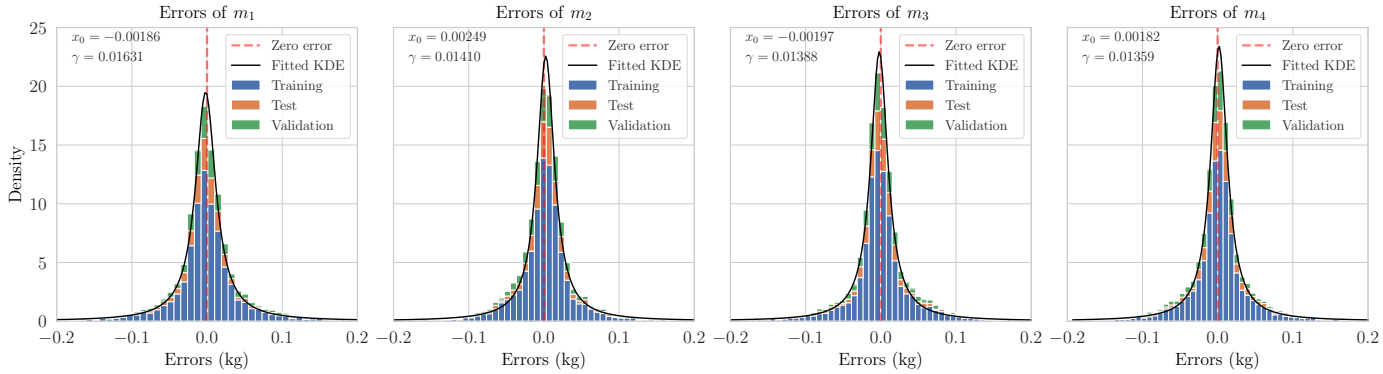


Figure 18: Error distribution for training, validation, and test sets. The *Fitted KDE* in the figure fits the error probability density of ensemble dataset, including the training, validation, and test set, and its distribution approximately subjects to Cauchy distribution $f(x; x_0, \gamma) = \frac{1}{\pi}\left[\frac{\gamma}{\gamma^2+(x-x_0)^2}\right]$. The position parameters x_0 and the scale parameter γ are marked at the upper left corner.

better in terms of predicting the mass of the material, showing strong correlations between the prediction and the actual target mass. With other networks, it can be observed that there exist certain numbers of outlier samples in the predicted values that deviate far from the fitted line, especially with the network of Resnet101, Resnet50, Resnet18, and Densenet201. For the PSPD prediction, those models are reliable enough or at least not directly usable for practical applications by default without additional tuning or modifications. Besides, the slopes of the fitted lines of VGG16, Googlenet, and Alexnet are clearly lower than others, indicating that the three networks have a larger prediction bias for large-mass samples. Specifically, the slope of the fitted line of Alexnet is only 0.88558, which cannot meet the requirements.

It can be seen from the violin figure (Figure 21) that the proposed model, VGG16 and VGG19, perform considerably better than all other networks in terms of accuracy, especially for Densenet201, Resnet101, Resnet50, and Resnet18, whose maximum prediction errors have exceeded 1kg. Although the maximum prediction error of Alexnet is lower than 0.5kg, the distribution of its peak values deviate from zero clearly (as can be seen from Figure 22, the *Skew* of Alexnet has exceeded −0.37kg). The peaks of the prediction errors of VGG16, VGG19, Googlenet, Inception v2 and Inception v3 are close to zero. However, their results are more scattering than the model proposed in this paper.

Table 6 shows that the largest prediction error of the proposed model is the lowest among all models, where the maximum error is 0.192kg and the minimum error is −0.189kg (both are about 3.17% of the total weight). Due to the visualisation effect caused by the line width, the curve in Figure 21 appears to slightly exceed ±0.2kg. It should be noted that, when using the model of Densenet201, Resnet101, Resnet50, and Resnet18 for prediction, negative values appear in the results. However, it is obvious that negative results are practically meaningless. Further study confirmed that those models are less reliable for the PSPD prediction. Because the focus of this research is on the proposed model, there is no further analysis of this phenomenon, but to simply delete the samples corresponding to negative results when comparing different models. It can be seen from the comparison results that the method proposed in this paper has great advantages in terms of accuracy and reliability.

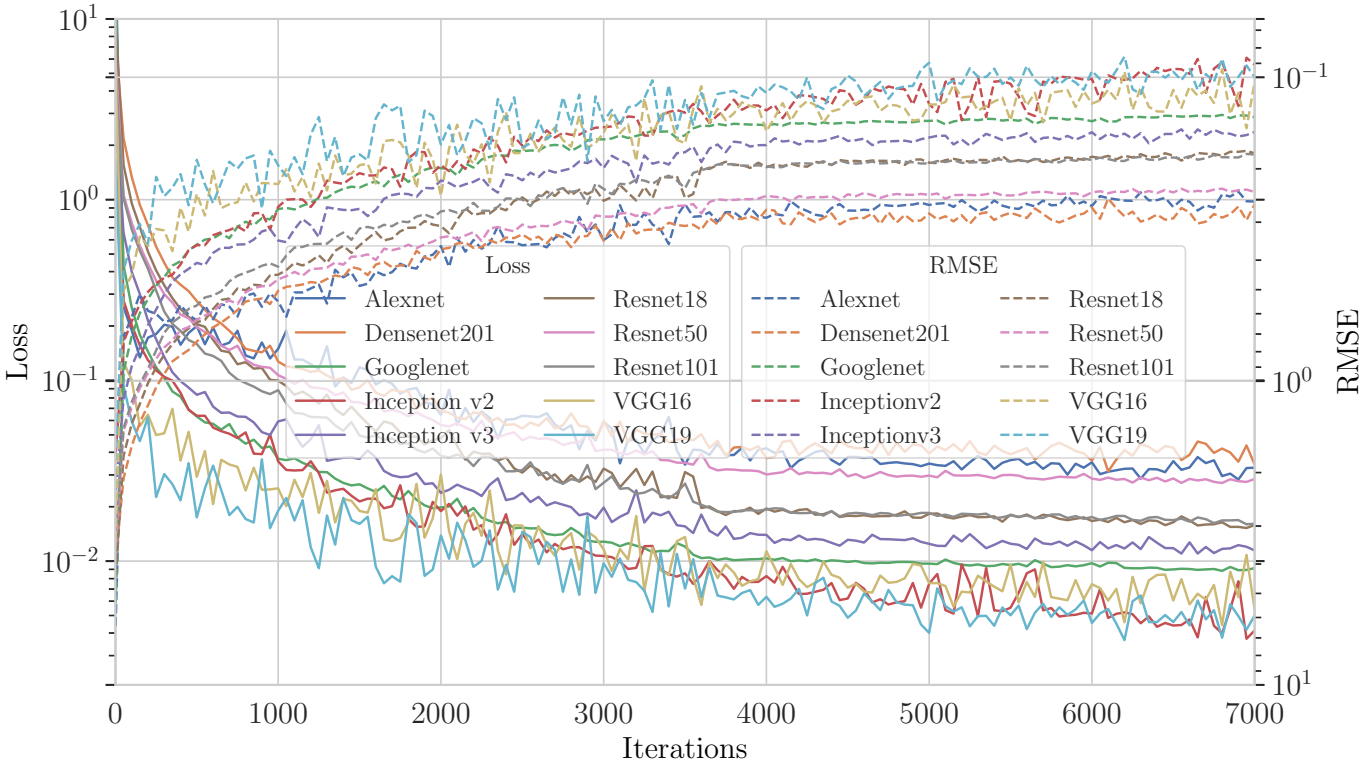


Figure 19: RMSE and Loss of popular networks. Since both RMSE and Loss decrease with iteration, the coordinates of RMSE (right-side ordinate) increase from top to bottom for understanding.

Table 6: Maximum and minimum errors and time required for a single prediction of each model.

Model	Max. error (kg)	Min. error (kg)	Time cost (ms/prediction)
Proposed model	0.19205	−0.18929	104.8
VGG19	0.28447	−0.23852	34.18
VGG16	0.28814	−0.22554	27.40
Inception v3	0.29956	−0.53738	34.98
Inception v2	0.33447	−0.32081	74.11
Googlenet	0.43461	−0.34711	7.37
Alexnet	0.47534	−0.17423	2.11
Resnet18	0.60868	−1.03483	6.60
Resnet101	0.64185	−1.16618	29.26
Densenet201	1.01221	−1.39022	47.11
Resnet50	1.36430	−0.84171	19.30

5.3. Real-world environment

Applying this method to real-world construction applications should overcome some engineering challenges, including 1) limited labelled training data for the real-world environment, 2) the difference of the data between laboratory and real-world environments, 3) the unstructured nature of real-world data, and 4) noisy images and sensor data from the real-world environment. Below are our strategies to deploy the work for real-world environment.

Part I deals with pre-processing of sensor data from both the camera and load cells. On construction sites, image quality will be unavoidably degraded by the adverse effects of light scattering due to the presence of dust particles that will hinder the performance of the proposed system. In order to enhance the observed image quality, removal of the interference caused by dust will be required. Deep Learning-base methods will be deployed for reconstructing high-quality images by training neural networks using synthetic images [50, 51] and using CNNs or generative networks, such as Generative Adversarial Networks (GAN) and Variational Autoencoder (VAE), to produce clean images [51, 52]. For dust-specific image enhancement, it is also worth investigating in manipulating image parameters in the colour space, such as the halo-reduced dark channel prior (DCP) dehazing technique [53]. In addition, trustworthy labels are critical to the success of the model

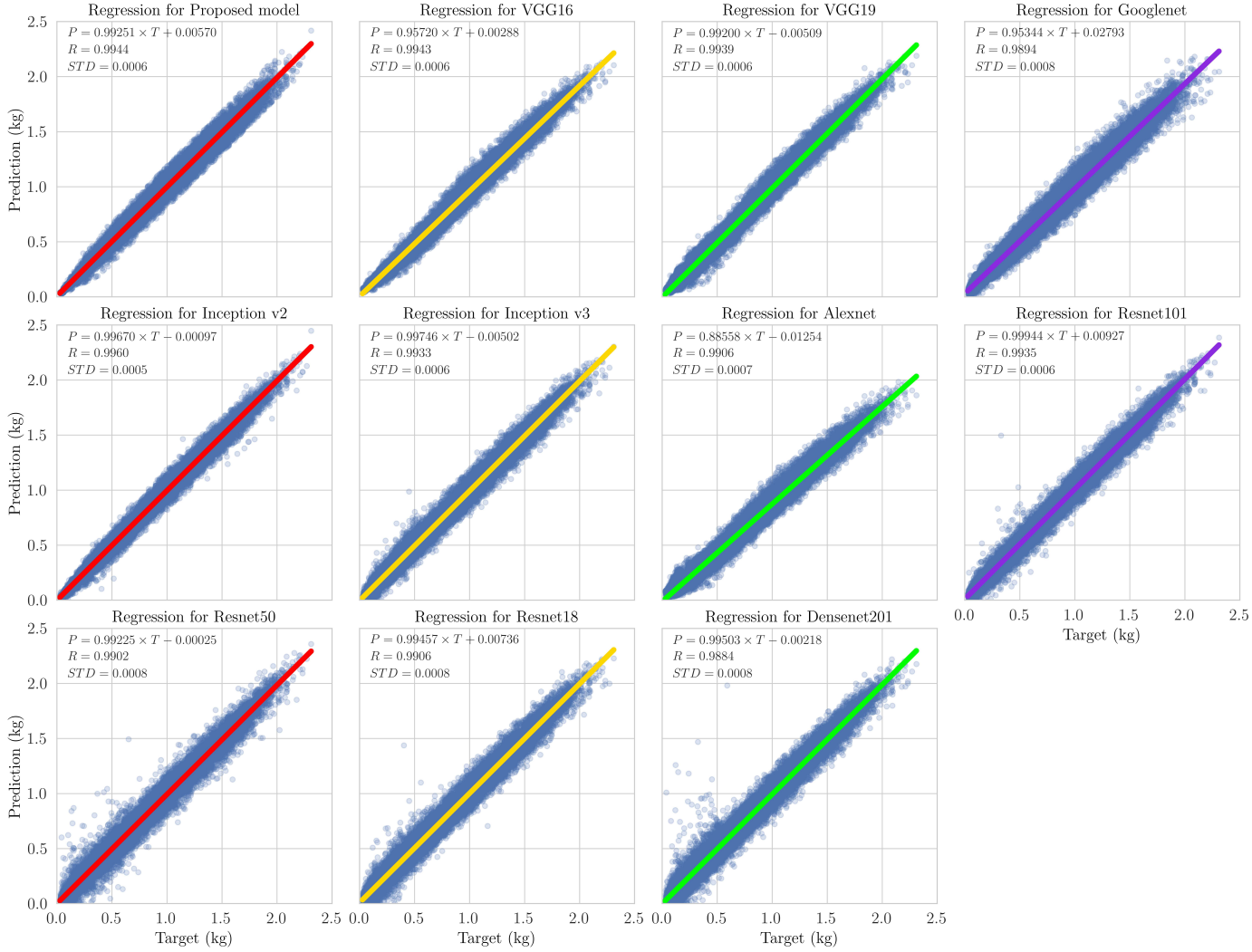


Figure 20: Relationship between target and predicted values. The equation expression, correlation coefficient R, and standard deviation STD obtained by linear fitting are labelled in the figure.

training. However, as mentioned in Section 1, it is difficult to keep the vehicle horizontal after each load from shovels caused by the suspension system and rough terrain, which lead to inaccurate force measurement from the four corners, hence not able to provide true representations of the distributions of the materials. Sensors, such as gyroscope, will be needed to provide tilt angles of the truck body to compensate the force measurement. It should be noted that once the model has been trained, it is not necessary to use such sensors, when applying it to the vehicle, because the camera will stay relatively stationary with the truck body.

Secondly, the CNNs (Part II) are not directly applicable for real world environment due to the different feature spaces from the laboratory experiments. The most obvious solution is through transfer learning, which allows the use of pre-trained models to be mapped to a different domain, the real-world dataset here [54, 55]. Since there will be limited data obtained from the real-world environment, it is unrealistic to retrain the whole network from scratch, i.e. the feature levels at the convolutional layers. We believe that the materials that we tested in the laboratory share some similarities with the real-world images at its low-level representations. In that case, the low-level generic features from the laboratory experiments can be reused for the real-world data and the network layers will be selectively freezed or fine-tuned during retraining. It will be challenging if the labelled real-world data are scarce, and focus will be on the strategy of how to balance freezing or fine-tuning the generic feature layers. The network structure in Part III is relatively simpler and should be more straightforward to be retrained. Further experiments and investigation will be focusing on incorporating prior domain knowledge to further constrain the training process to improve the performance and reliability in the

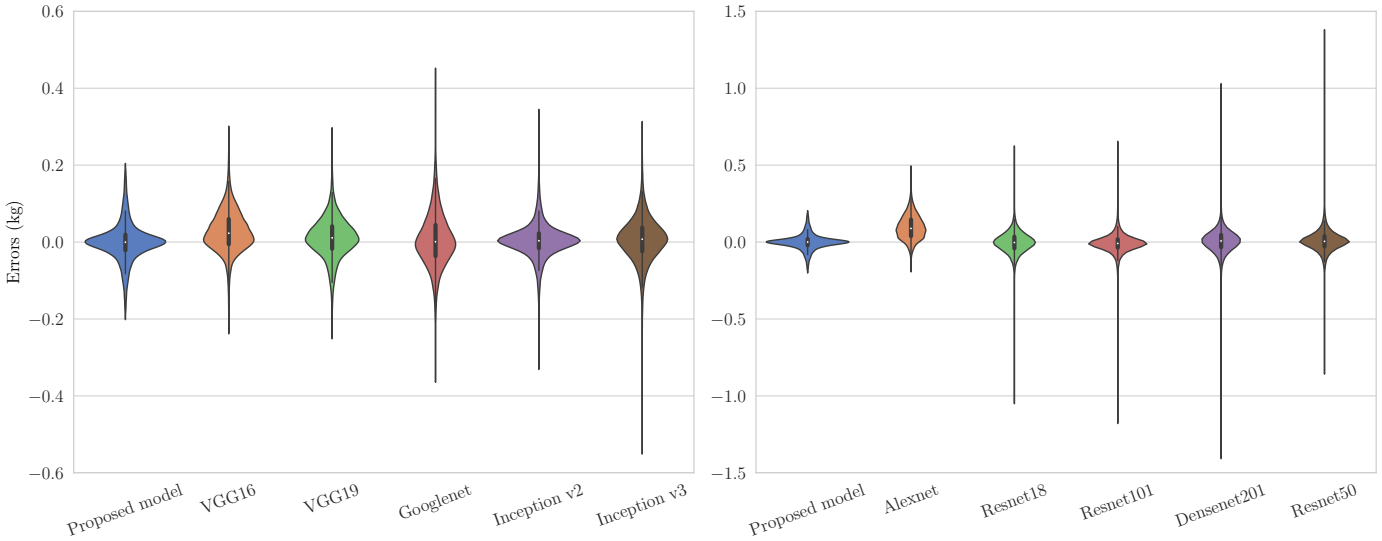


Figure 21: Volin figure of prediction errors for each model. The limitations of the y coordinate for the left and right sub-figures are different. Besides, the first models on both the left and right sub-figures represent the Volin plots of the data of the proposed model, to demonstrate the errors of different scales with different networks for better visualisation.

real-world environment.

6. Conclusions

The pilot study of the application of artificial neural networks on predicting the PSPD is proposed in this paper. The effectiveness of the solution has been verified experimentally with real data collected from a laboratory-based setup using two robots, simulating the mining site environments.

The prediction model consists of three main parts, namely data pre-processing, CNN and BPNN. Firstly, the operations of background removal and grayscale conversion are applied to images. Then, filtering and averaging are applied for data processing of material mass and coordinates to remove irrelevant noises, enhance data availability, speed up network training, and improve model prediction accuracy. Secondly, based on the model of VGG19, a two-stage prediction-regression CNN model is designed to automatically extract image features to obtain the PSPD in the current state. Finally, a shallow BPNN with 6 inputs and 4 outputs is established to predict the PSPD after the current dumping operation. In order to verify the proposed model, we designed experiments and obtained a large amount of data allowing for model training and testing. The proposed model has demonstrated better performance, in terms of accuracy, reliability, and so on, compared to other standard neural networks.

Our work has two main contributions. Firstly, a new data pipeline or model is developed by combining machine vision, data processing, CNN and BPNN, which can perform regression prediction on images. This idea can be also accommodated for solving prediction problems in other similar engineering fields related to the distribution of bulk materials, such as estimation of the material quantity in loader buckets, storages in granary, etc. Secondly, the algorithm proposed in this paper accomplishes the real-time prediction of the PSPD in the laboratory environment. It can be expected that, when the method is transplanted into the actual production environment, drivers can adjust the dumping position in advance to achieve the material balance in the truck body. Besides, the proposed model lays the foundation in supporting future research in autonomous operations of mining equipment.

Although the proposed model provides satisfactory predictive performance, there are still some limitations. The characteristics of data collected from the laboratory and data from real-world working conditions may be inconsistent. For example, this work was carried out in a laboratory environment under controllable lighting conditions. Also, the impact of adverse factors such as dust and vibrations are nearly negligible here. However, in real operational situations, lighting conditions will vary greatly. The impact of illumination, dust and vibrations on the quality of the images cannot be ignored. The results in this paper demonstrate the feasibility of our proposed method for predicting the PSPD of bulk material. However, it is foreseeable that a considerably large amount of data would be required for real-world conditions to cover many different conditions, compared to laboratories. In future, we will also consider other machine vision methods to preprocess images and enhance image quality in conditions with reduced visibility due to illumination, dust and vibrations. The influence of the material's characteristics on the results will also be investigated. In this work, the bulk materials are of the same type, which, however, will not be always true in real-world operations that the materials would be likely to be a mixture of multiple types.

In addition, this article only discusses the possibility of dumping control methods based on prediction optimiza-

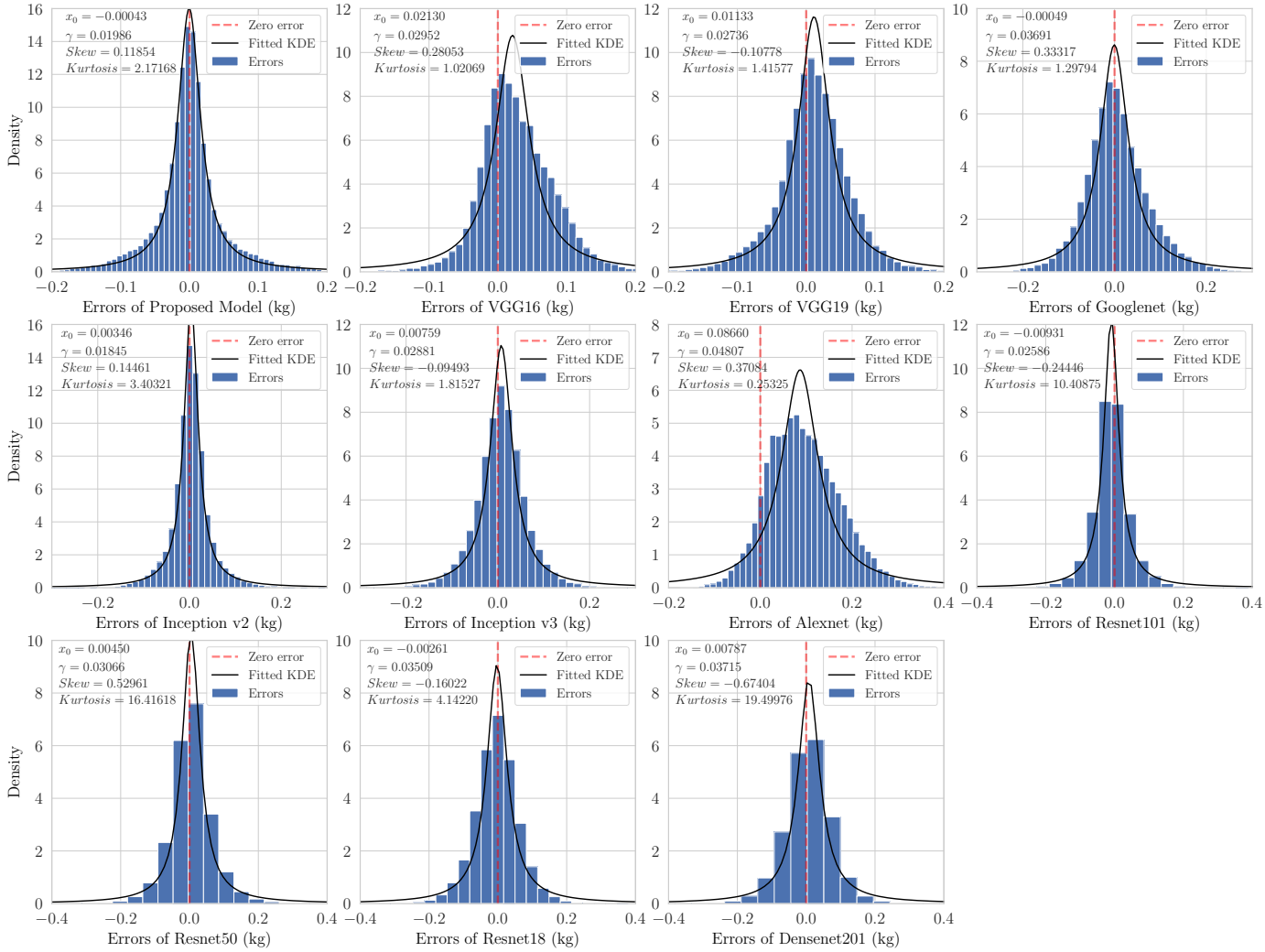


Figure 22: Distribution of prediction error for each model. The *Fitted KDE* in the figure is the curve obtained by fitting the probability density with the Cauchy distribution. The fitting parameters x_0 and γ corresponding to each result are marked in the upper left corner of the graph as well as the *Skew* and *Kurtosis*.

tion in mining shovels, and detailed research is needed in the future. Although drivers can manually select the dumping position based on the results of the prediction model, it is still difficult to achieve the optimal balance result for imprecise operations. Therefore, in the subsequent work, a multi-objective optimization study of the dumping position will be performed based on the prediction model.

Acknowledgements

This work was supported by the National Natural Science Foundation of China [grant numbers 51875232, 51875233]. In addition, thanks to the Key Laboratory of CNC Equipment Reliability, Ministry of Education, China, for providing the high performance computer on which we trained the CNN models. The authors would also like to acknowledge the support of the Royal Society under the research grant agreement RGS\R1\201159.

References

[1] A. Chamanara, Enhancing mine haul truck KPIs via payload balance, Ph.D. thesis, University of Alberta, 2013. doi:10.7939/R3JQ0T24T.

[2] J. C. Henze, T. G. Joseph, R. Hall, M. Curley, Haul truck body payload placement modeling, International Research Journal of Geology and Mining 7 (2017). URL: https://pdfs.semanticscholar.org/961b/44e67e424bb877749984040610fc53ebdc6f.pdf, accessed May 5, 2020.

[3] M. Hardigree, Caterpillar mining truck axle failure, https://jalopnik.com/caterpillar-mining-truck-axle-failure-5385130, 2009. Accessed May 5, 2020.

[4] A. Stentz, J. Bares, S. Singh, P. Rowe, A robotic excavator for autonomous truck loading, Autonomous Robots 7 (1999) 175–186. doi:10.1023/A:1008914201877.

[5] E. Duff, Automated volume estimation of haul-truck loads, in: Proceedings of the Australian Conference on Robotics and Automation, CSIRO, 2000, pp. 179–184. URL: https://www.researchgate.net/publication/245573520, accessed May 5, 2020.

[6] J. R. Borthwick, Mining haul truck pose estimation and load

profiling using stereo vision, Ph.D. thesis, University of British Columbia, 2009. doi:10.14288/1.0070913.

[7] A. Bewley, B. Upcroft, P. Lever, S. Leonard, Automatic in-bucket volume estimation for dragline operations, in: Australian Mining Technology Conference, 2009, pp. 153–167. URL: https://www.researchgate.net/publication/43526477, accessed May 5, 2020.

[8] A. Bewley, R. Shekhar, S. Leonard, B. Upcroft, P. Lever, Real-time volume estimation of a dragline payload, in: 2011 IEEE International Conference on Robotics and Automation, 2011, pp. 1571–1576. doi:10.1109/ICRA.2011.5979898.

[9] B. Upcroft, R. C. Shekhar, A. J. Bewley, P. J. Lever, Measurement of bulk density of the payload in a dragline bucket, 2015. URL: https://patents.google.com/patent/US8930091B2/en, US Patent 8,930,091.

[10] X. Huang, 3D shape reconstruction by dynamic sensing with a range sensor, Ph.D. thesis, University of Tokyo, 2013. URL: https://ci.nii.ac.jp/naid/500000582516/, accessed May 5, 2020.

[11] C. Innes, E. Nettleton, A. Melkumyan, Estimation and tracking of excavated material in mining, in: 14th International Conference on Information Fusion, IEEE, 2011, pp. 1–8. URL: https://ieeexplore.ieee.org/document/5977565.

[12] C. Innes, A stochastic method for representation, modelling and fusion of excavated material in mining, Ph.D. thesis, University of Sydney, 2012. URL: https://ses.library.usyd.edu.au/handle/2123/13333, accessed May 5, 2020.

[13] M. Dunbabin, P. Corke, Autonomous excavation using a rope shovel, Journal of Field Robotics 23 (2006) 379–394. doi:10.1002/rob.20132.

[14] S. Sarata, Y. Weeramhaeng, T. Tsubouchi, Approach path generation to scooping position for wheel loader, in: Proc. IEEE Int. Conf. Robotics and Automation, 2005, pp. 1809–1814. doi:10.1109/ROBOT.2005.1570376.

[15] N. Koyachi, S. Sarata, Unmanned loading operation by autonomous wheel loader, in: Proc. ICCAS-SICE, 2009, pp. 2221–2225. URL: https://ieeexplore.ieee.org/document/5335253.

[16] T. Takei, T. Hoshi, S. Sarata, T. Tsubouchi, Simultaneous determination of an optimal unloading point and paths between scooping points and the unloading point for a wheel loader, in: Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS), 2015, pp. 5923–5929. doi:10.1109/IROS.2015.7354219.

[17] A. Rasuli, S. Tafazoli, W. G. Dunford, Dynamic modeling, parameter identification, and payload estimation of mining cable shovels, in: Proc. IEEE Industry Application Society Annual Meeting, 2014, pp. 1–9. doi:10.1109/IAS.2014.6978451.

[18] H. Anwar, S. M. Abbas, A. Muhammad, K. Berns, Volumetric estimation of contained soil using 3d sensors, in: Intl. Commercial Vehicle Technology Symposium, 2014, pp. 11–13. URL: http://cyphynets.lums.edu.pk/images/SoilEstimCVT2014.pdf, accessed May 5, 2020.

[19] Q. Bi, G. Wang, R. Yang, Y. Liu, Y. Lu, S. Xing, Study on theory and methods of payload online estimation for cable shovels, Journal of the Brazilian Society of Mechanical Sciences and Engineering (2019) 41–53. doi:10.1007/s40430-018-1546-1.

[20] P. A. Cundall, O. D. L. Strack, A discrete numerical model for granular assemblies, Gotechnique 29 (1979) 47–65. doi:10.1680/geot.1979.29.1.47.

[21] P. W. Cleary, Large scale industrial dem modelling, Engineering Computations 21 (2004) 169–204. doi:10.1108/02644400410519730.

[22] J. Hilton, P. Cleary, Granular flow during hopper discharge, Physical Review E 84 (2011) 011307. doi:10.1103/PhysRevE.84.011307.

[23] P. W. Cleary, DEM prediction of industrial and geophysical particle flows, Particuology 8 (2010) 106–118. doi:10.1016/j.partic.2009.05.006.

[24] C. Coetzee, A. Basson, P. Vermeer, Discrete and continuum modelling of excavator bucket filling, Journal of Terramechanics 44 (2007) 177–186. doi:10.1016/j.jterra.2006.07.001.

[25] C. Coetzee, D. Els, The numerical modelling of excavator bucket filling using DEM, Journal of Terramechanics 46 (2009) 217–227. doi:10.1016/j.jterra.2009.05.003.

[26] H. Teufelsbauer, Y. Wang, M. C. Chiou, W. Wu, Flow-obstacle interaction in rapid granular avalanches: DEM simulation and comparison with experiment, Granular Matter 11 (2009) 209–220. doi:10.1007/s10035-009-0142-6.

[27] G. Lu, J. Third, C. Müller, Discrete element models for non-spherical particle systems: from theoretical developments to applications, Chemical Engineering Science 127 (2015) 425–465. doi:10.1016/j.ces.2014.11.050.

[28] S. Wang, Y. Fan, S. Ji, Interaction between super-quadric particles and triangular elements and its application to hopper discharge, Powder Technology 339 (2018) 534–549. doi:10.1016/j.powtec.2018.08.026.

[29] S. Lommen, D. Schott, G. Lodewijks, DEM speedup: Stiffness effects on behavior of bulk material, Particuology 12 (2014) 107–112. doi:10.1016/j.partic.2013.03.006.

[30] P. Gopalakrishnan, D. Tafti, Development of parallel DEM for the open source code MFIX, Powder Technology 235 (2013) 33–41. doi:10.1016/j.powtec.2012.09.006.

[31] J. Gan, Z. Zhou, A. Yu, A GPU-based DEM approach for modelling of particulate systems, Powder Technology 301 (2016) 1172–1182. doi:10.1016/j.powtec.2016.07.072.

[32] X. Yue, H. Zhang, C. Luo, S. Shu, C. Feng, Parallelization of a DEM code based on CPU-GPU heterogeneous architecture, in: International Conference on Parallel Computing in Fluid Dynamics, Springer, 2013, pp. 149–159. doi:10.1007/978-3-642-53962-6_13.

[33] S. Jeong, B. Solenthaler, M. Pollefeys, M. Gross, et al., Data-driven fluid simulations using regression forests, ACM Transactions on Graphics (TOG) 34 (2015). doi:10.1145/2816795.2818129.

[34] L. G. Hagenbuch, P. T. Brinkman, Process for three-dimensional modeling and design of off-highway dump bodies, 2012. URL: https://patents.google.com/patent/US7369978B1/en, uS Patent 8,113,763.

[35] T. Joseph, A. Chamanara, Hauler body payload balance, CIM Journal 3 (2012) 9. URL: https://store.cim.org/en/hauler-body-payload-balance, accessed May 5, 2020.

[36] C. Schenck, J. Tompson, D. Fox, S. Levine, Learning robotic manipulation of granular media, arXiv preprint arXiv:1709.02833 (2017). URL: https://arxiv.org/abs/1709.02833, accessed May 5, 2020.

[37] T. Takei, K. Ichikawa, K. Okawa, S. Sarata, T. Tsubouchi, A. Torige, Path planning of wheel loader type robot for scooping and loading operation by genetic algorithm, in: Proc. 13th Int. Conf. Control Automation and Systems (ICCAS 2013), 2013, pp. 1494–1499. doi:10.1109/ICCAS.2013.6704123.

[38] L. Rozo, P. Jimnez, C. Torras, Force-based robot learning of pouring skills using parametric hidden markov models, in: Proc. 9th Int. Workshop Robot Motion and Control, 2013, pp. 227–232. doi:10.1109/RoMoCo.2013.6614613.

[39] A. Yamaguchi, C. G. Atkeson, Differential dynamic programming for graph-structured dynamical systems: Generalization of pouring behavior with different skills, in: Proc. IEEE-RAS 16th Int. Conf. Humanoid Robots (Humanoids), 2016, pp. 1029–1036. doi:10.1109/HUMANOIDS.2016.7803398.

[40] Z. Zhao, M. Jin, C. Tian, S. X. Yang, Prediction of seed distribution in rectangular vibrating tray using grey model and artificial neuralnetwork, Biosystems Engineering 175 (2018) 194–205. doi:https://doi.org/10.1016/j.biosystemseng.2018.09.017.

[41] R. Kumar, C. M. Patel, A. K. Jana, S. R. Gopireddy, Prediction of hopper discharge rate using combined discrete element method and artificial neural network, Advanced Powder Technology 29 (2018) 2822–2834. doi:10.1016/j.apt.2018.08.002.

[42] J. Rabault, M. Kuchta, A. Jensen, U. Réglade, N. Cerardi, Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control, Journal of Fluid Mechanics 865 (2019) 281–302. doi:10.1017/jfm.2019.62.

[43] L. Benvenuti, C. Kloss, S. Pirker, DEM parameter identification by means of artificial neural network for iron ore sintering,

in: Eleventh International Conference on CFD in the Minerals and Process Industries CSIRO, Melbourne, Australia, 2015, pp. 7–9. URL: http://www.cfd.com.au/cfd_conf15/PDFs/020BEN.pdf, accessed May 5, 2020.

[44] L. Benvenuti, C. Kloss, S. Pirker, Identification of DEM simulation parameters by artificial neural networks and bulk experiments, Powder Technology 291 (2016) 456–465. doi:10.1016/j.powtec.2016.01.003.

[45] F. Ye, C. Wheeler, B. Chen, J. Hu, K. Chen, W. Chen, Calibration and verification of DEM parameters for dynamic particle flow conditions using a backpropagation neural network, Advanced Powder Technology 30 (2019) 292–301. doi:10.1016/j.apt.2018.11.005.

[46] S. Chen, L. A. Baumes, A. Gel, M. Adepu, H. Emady, Y. Jiao, Classification of particle height in a hopper bin from limited discharge data using convolutional neural network models, Powder Technology 339 (2018) 615–624. doi:10.1016/j.powtec.2018.08.048.

[47] Z. Y. Wan, T. P. Sapsis, Machine learning the kinematics of spherical particles in fluid flows, Journal of Fluid Mechanics 857 (2018). doi:10.1017/jfm.2018.797.

[48] E. B. Goldstein, G. Coco, A machine learning approach for the prediction of settling velocity, Water Resources Research 50 (2014) 3595–3601. doi:10.1002/2013WR015116.

[49] S. Clarke, T. Rhodes, C. G. Atkeson, O. Kroemer, Learning audio feedback for estimating amount and flow of granular material, in: A. Billard, A. Dragan, J. Peters, J. Morimoto (Eds.), Proceedings of The 2nd Conference on Robot Learning, volume 87 of *Proceedings of Machine Learning Research*, PMLR, 2018, pp. 529–550. URL: http://proceedings.mlr.press/v87/clarke18a.html, accessed May 5, 2020.

[50] C. Li, S. Anwar, F. Porikli, Underwater scene prior inspired deep underwater image and video enhancement, Pattern Recognition 98 (2020) 107038. doi:https://doi.org/10.1016/j.patcog.2019.107038.

[51] P. M Uplavikar, Z. Wu, Z. Wang, All-in-one underwater image enhancement using domain-adversarial learning, in: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2019. URL: https://arxiv.org/abs/1905.13342.

[52] S. Bianco, L. Celona, F. Piccoli, R. Schettini, High-resolution single image dehazing using encoder-decoder architecture, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, Long Beach, CA, USA, 2019, pp. 1927–1935. doi:10.1109/CVPRW.2019.00244.

[53] Z. Shi, Y. Feng, M. Zhao, E. Zhang, L. He, Let you see in sand dust weather: A method based on halo-reduced dark channel prior dehazing for sand-dust image enhancement, IEEE Access 7 (2019) 116722–116733. doi:10.1109/ACCESS.2019.2936444.

[54] S. J. Pan, Q. Yang, A survey on transfer learning, IEEE Transactions on Knowledge and Data Engineering 22 (2010) 1345–1359. doi:10.1109/TKDE.2009.191.

[55] K. Weiss, T. M. Khoshgoftaar, D. Wang, A survey of transfer learning, Journal of Big Data 3 (2016). doi:10.1186/s40537-016-0043-6.