

Received April 2, 2020, accepted April 9, 2020, date of publication April 24, 2020, date of current version May 7, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2990200

Coal/Gangue Recognition Using Convolutional Neural Networks and Thermal Images

MURAD SALEH ALFARZAEI¹, QIANG NIU¹, JIAQI ZHAO¹, (Member, IEEE),
REFAT MOHAMMED ABDULLAH ESHAQ², AND ERYI HU²

¹School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China

²School of Mechatronic Engineering, China University of Mining and Technology, Xuzhou 221116, China

Corresponding author: Qiang Niu (niuq@cumt.edu.cn)

This work was supported in part by the State's Key Project of Research and Development Plan of China under Grant 2016YFC0600908 and Grant 2018YFC0604503, in part by the National Natural Science Foundation of China under Grant 51674255, Grant 61806206, and Grant 51875568, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20180639, and in part by the China Postdoctoral Science Foundation under Grant 2018M642359.

ABSTRACT Recognition and separation of Coal/Gangue are important phases in the coal industries for many aspects. This paper addressed the topic of Coal/Gangue recognition and built a new model called (CGR-CNN) based on Convolutional Neural network (CNN) and using thermal images as standard images for Coal/Gangue recognition. The CGR-CNN model has been developed, augmentation principle has been applied in order to increase the dataset and the best experimental results have been achieved (99.36%) learning accuracy and (95.09%) validation accuracy, in the prediction phase (160) new images of coal and gangue (80 for both) have been tested to measure the efficiency of the work, the prediction result comes with (100%) for coal recognition accuracy and (97.5%) gangue recognition accuracy giving an overall prediction accuracy (98.75%).

INDEX TERMS Coal, coal gangue, convolutional neural network, object classification, dataset augmentation, thermal images, separation system.

I. INTRODUCTION

The coal industries around the world are still widespread and the demand for technologies to help raise the efficiency of these industries is still high, the process of coal mining and the associated overlap of coal pockets with other geological components of mines and the mixing of coals with gangues is inevitable so the process of separating coal from gangue is still urgent, in China coal is still the primary source of fossil energy with 90% [1], on the other hand coal is the primary source of China's carbon emissions [1], [2], the importance of Coal/Gangue separation based on the need of reducing the environment hazards and increasing production efficiency is still an important open research topic [3], [4], "Gangue is defined as non ore rock surrounding or associated with the ore" [5], it is important to separate coal from gangue for many aspects such as environmental, productive and worker safety [6]–[13]. The separation process could be manual or mechanical, manual separation process needs the effort of the expert workers to recognize the gangue from the

coal on the transmitting belt this process consumes money, restricts mineral process and may badly influence the workers health [13], [14], on the other hand mechanical separation could lead to environmental pollution and affect the quality of the produced coals [7], [11], to increase the efficiency of separation systems it is good to combine the advantages of the two separation processes to come out with an efficient separating system, the visual characteristics of the Coal/Gangue have been added to the mechanical system to increase the efficiency of the separation system and to minimize pollution disadvantages by replacing the recognition methods which lead to environmental pollution by vision sensors which give the ability of making visual recognition of the Coal/Gangue [10], [14], [15], many work effort have been done in the last years to improve the separation systems in this direction using the computer vision techniques as the visual sensor decision maker for the separation systems by using the image processing techniques to recognize the Coal/Gangue using either the gray scale images [5], [7], [9], [14], [16], using wavelet transform to improve the images quality [10], [12], [17], [18] and also using the neural network algorithms as the main decision unit which proves that it has great abilities in pattern

The associate editor coordinating the review of this manuscript and approving it for publication was Sudhakar Radhakrishnan¹.

recognition and objects classifications [3], [6], [11], [13], [19]–[23]. Yaqun *et al.* [20] addressed the use of 17 characteristic parameters of the gray-scale histogram and gray level co-occurrence matrix (GLCM) selected according to their differences in gray scale and texture then using the (PCA) algorithm to get the principle component from the selected parameters to be inputs for (GA-ANN) to make the Coal/Gangue identification, according to the simulation of the experiment it achieved (100%) of recognition accuracy but the data set was too small around (32) rock images and (32) coal images were used in the experiment. Gao *et al.* [16] introduce (RNRA) as an image processing technique was developed to recognize the Coal/Gangue based on gray scale feature using Bayesian Decision Theory and using (70) gangues and (23) coals for building the system analysis and testing, experiment result comes with (96.8%) recognition accuracy, but the samples for analyzing the system are the same for making the testing phase and they didn't use new samples to verifying the testing results. Li and Sun [13] proposed the use of the LS-SVM as based and gray scale with texture as features, the experiment used (500) images for 4 kind of Coal/Gangue(Lean Coal, Shale, Coking Coal, Sandstone) from two different mines and achieved around (98.7%, 96.6%, 98.6%, 96.6%) recognition accuracy. Eshaq *et al.* [24] a recognition system has been developed using thermal images of coal and gangue which perform feature extraction based on YCbCr and use SVM to classify the coal gangue, the proposed method achieved high classification accuracy (98.1%) for gangue and (96.6%) for coal this work well be mention frequently later as SVM-YCbCr.

The Convolutional Neural Network has lately become widely known in the field of object recognition and classifications tasks, looking for the use of the convolutional neural network in the field of Coal/Gangue recognition systems some previous works have been done, Hong *et al.* [11] proposed an algorithm to recognize the Coal/Gangue by using the (CNN), development has been done based on the AlexNet model using (2012) images for three kinds Coal(Matt), Coal(Gloss) and Gangue in the experiment and achieved around (96.6%)recognition accuracy. Su *et al.* [19] proposed an algorithm to recognize the Coal/Gangue by using the (CNN) and development has been done based on the Yann LeCun's LeNet-5 model, the developed algorithm achieved around (95.88%) using (20000) images of Coal/Gangue and training epoics around (10000) epoic. Pu *et al.* [6] proposed an algorithm to recognize the Coal/Gangue by using the (CNN) and development has been done based on the VGG16 model, transfer learning used to overcome the over-fitting problem because of the shortage of the training samples which was around (240) images for training and testing, the proposed algorithm achieved around (82.5%) recognition accuracy.

In the case of coal and gangue images a problem of heterogeneous of the Coal/Gangue sources usually lead to differences in the Gray scale images leading to mis-recognition of the same object [6], this problem could be avoidable using

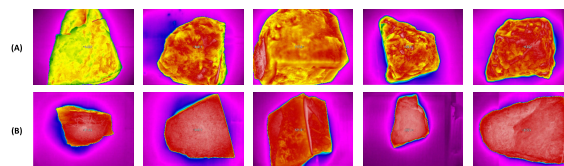


FIGURE 1. Thermals Images (A) Five Thermal Coal Images. (B) Five Thermal Gangue Images.

the thermal images in Eshaq *et al.* [24] which tested the coal and gangue in different heat degrees and analyzed the best conditions for thermal images to perform a classification test on coal and gangue, but using SVM algorithm need to perform feature extraction to prepare the inputs of the SVM which require deep analysis and usually limited to the set of analyzed factors, on the other hand using the convolutional neural network shorten the analysis process because CNN perform features extraction by itself which will not be limited to pre_analyzed factors and extract more features, based on that in this paper a CNN model has been designed to perform a recognition of coal and gangue using thermal images dataset from the Eshaq *et al.* [24] for training the model Figure 1, the ability to operate in medium-term operational requirements (Memory size, Training time) has been taken in account so that the new model should be able to run in medium size GPU such as 4GB GPU and does not take along time for learning the model, while maintaining its ability to effectively obtain high accuracy comparing with the related work. The developed model was able to achieve better coal recognition reached 100% compared to 96.6 % for coal in Eshaq *et al.* [24] also achieved a near gangue recognition reached 97.5% compared to 98.1% for gangue in Eshaq *et al.* [24], so the implementing of the new model takes the following steps:

- Design the model to run in normal moderate hardware and perform learning in an acceptable time with high recognition accuracy, it will be called CGR-CNN in the rest of the paper.
- Preparing the dataset and using augmentation processing to increase the dataset.
- Comparing the performance of the CGR-CNN which come in three steps:
 - Comparing CGR-CNN with the main research work SVM-YCbCr [24] and demonstrate the achieved improvement (sect. III-A).
 - Comparing CGR-CNN with related CNN previous works(LeNet-5 [25], LeNet-5_improved [19], Alexnet [26], VGG_A and VGG_B [27])(sect. III-B)
 - Comparing CGR-CNN with Alexnet [26] as the best CNN models that respond well with the thermal images dataset(sect. III-C)

The rest of the paper is organized as follows: the second section (Model Development Stage) briefly explains the reasons of using the convolution neural network CNN and explain the CGR-CNN structure. The experiment setup explaining the hardware and software platforms, Data setup

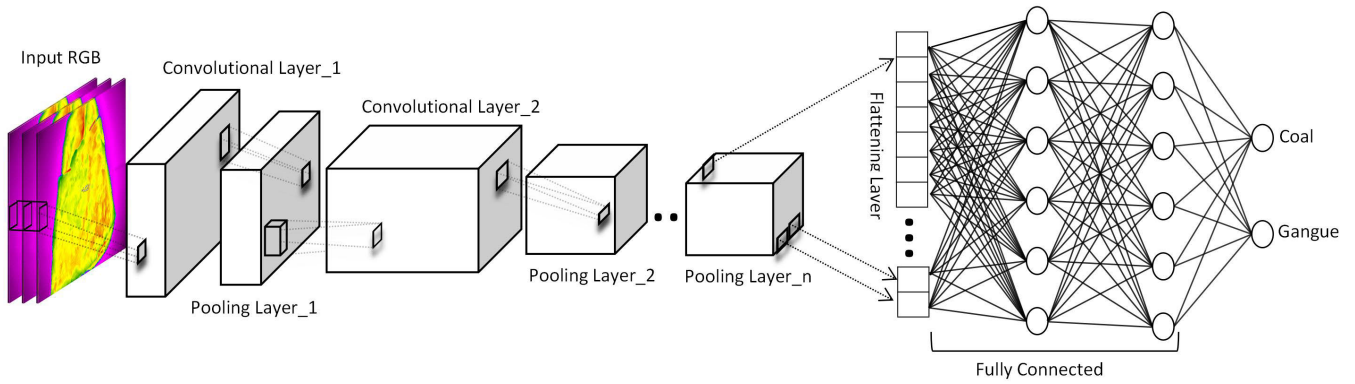


FIGURE 2. The construction of the CNN model.

(data collection, data preprocessing and data augmentation), discuss CGR-CNN learning results with comparison to related work respectively in the third section. Finally, conclusion of the research and clarify it's importance with notes for future work.

I. MODEL DEVELOPMENT STAGE

A. CONVOLUTIONAL NEURAL NETWORK (CNN)

A Convolutional neural network (CNN) advantage over other ordinary fully connected neural networks comes from the unique architecture whereas CNN consist of four different layers convolution layer, pooling layer, flattening layer and fully connected layer, this architecture gives the CNN the ability of reducing the trainable parameters in compared to the ordinary fully connected neural network which could be a crucial factor in training models with high resolution input images, the structure of the convolutional neural network model is shown in **Figure 2**. Visual perception needs good vision which needs a high quality images or we can say high resolution images, using the ordinary neural networks in learning with high resolution images need high requirements to learn the huge number of training parameters in such neural network, to explain that suppose a fully connected neural network consist of three layers take an image with (224 × 224x3) pixels as input which will be converted into vector of (150,528) value holding the whole pixels values and passing into the first layer, the number of the training parameters can be calculated by:

$$TP = \sum_{i=1}^{n-1} 1(L_i + 1) * (L_{i+1}) \tag{1}$$

where (TP) is the number of trainable parameters, (n) total number of layers, (L) is the size of the respected layer, so if the second layer for example have size of (1024) and the output layer has (2) classes, this shallow neural network need to learn (154,143,746) training parameters, this is just with shallow network if we get deeper in the network and increase the hidden layers the training parameters will increase which mean the need for high requirement to execute it, actually this size of the input image has been used in the CGR-CNN shown

in **Figure 3** which consist of 8 convolutional layers, 8 max pooling layers and 4 fully connected layer, to calculate the training parameters in CGR-CNN it take three steps first calculating the training parameters for every convolution layer by:

$$TP_c = \sum_{i=1}^n (K_i^2 * C_i * N_i) + N_i \tag{2}$$

where (TP_c) is the number of trainable parameters for the Conv layer, (K) Size of kernels used in the Conv Layer, (C) Number of channels in conv layer, (N) Number of kernels, (n) Number of Conv layers, second step is to calculate the training parameters for the first fully connected layer connected to the last convolution or max pooling layer by:

$$TP_{cf} = (O^2 * F * N) + F \tag{3}$$

where (TP_{cf}) is the number of trainable parameters for the first fully connected layer, (O) Size (width) of the output image of the previous Conv/maxpool Layer, (F) Number of neurons in the fully connected Layer, (N) Number of kernels in the previous Conv/maxpool Layer, third step to calculate the training parameters of the rest of fully connected layers by:

$$TP_{ff} = \sum_{i=1}^{n-1} (F_i * F_{i+1}) + F_{i+1} \tag{4}$$

where (TP_{ff}) is the number of trainable parameters for the rest of the fully connected layer, (F) Number of neurons in the fully connected Layer, (i) layers index, (n) number of fully connected layers, so the total number of parameters (TP) in the CNN equal to:

$$TP = TP_c + TP_{cf} + TP_{ff} \tag{5}$$

so the CGR-CNN comes with (5,232,750) trainable parameters in total (sect. I-B), the big difference in the number of the trainable parameters makes the convolutional neural network more practical, on the other hand with convolutional neural network there is no need for manually features extraction, which need extensive analysis for the input data to determine

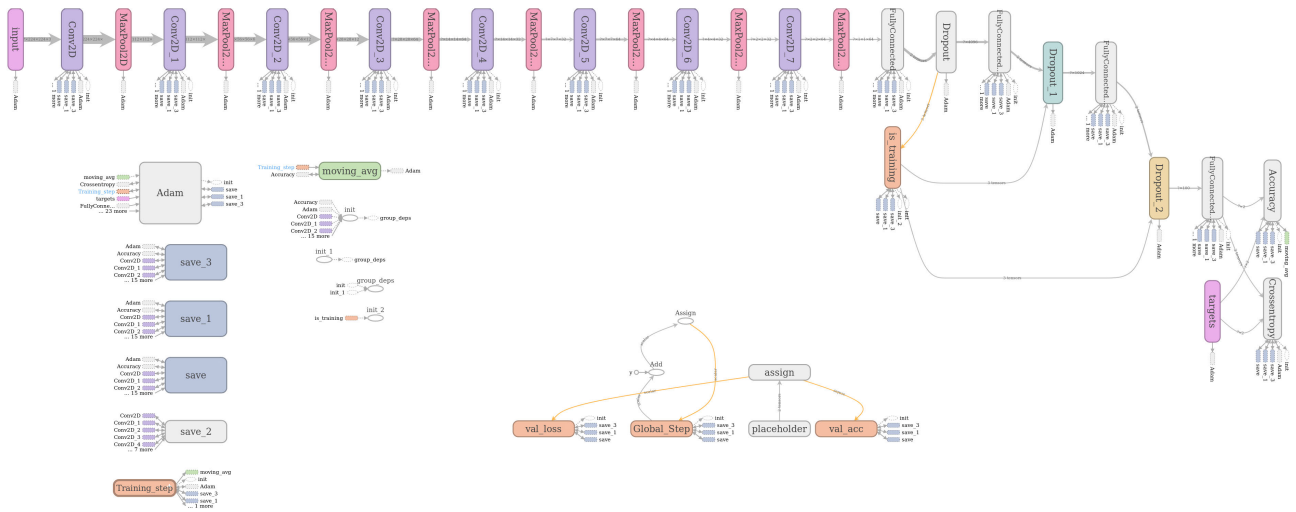


FIGURE 3. The model graph produced by Tensorboard.

the effective features to be inputs for other kinds of neural networks such as (SVM), because the convolution methods in the convolutional neural networks perform features extraction by itself this advantage make it easier to work with CNN in classification tasks more than the other neural networks models.

B. THE DEVELOPED MODEL

Based on the CNN advantages the CGR-CNN has been built on the base of the classical LeNet-5 but more deeper, it consists of 8 convolutional layers, 8 max pooling layers, 3 dropout layer set to (0.8), three fully connected layer and one fully connected out put layer, the input layer comes with (224 × 224 × 3) and no learning parameters, Conv2D is the first convolution layer has kernel of (5 × 5) with (32) kernels giving (2,432) trainable parameters with Relu as activation function, followed by MaxPool2d with kernel size (2 × 2) which reduce the size into (112 × 112 × 32) with no trainable parameters, Conv2D_1 is the second convolution layer has kernel of (5 × 5) with (64) kernels giving (51,264) trainable parameters with Relu as activation function followed by MaxPool2d_1 with kernel size (2 × 2) which reduce the size into (56 × 56 × 64) with no trainable parameters, Conv2D_2 is the third convolution layer has kernel of (5 × 5) with (128) kernels giving (204,928) trainable parameters with Relu as activation function followed by MaxPool2d_2 with kernel size (2 × 2) which reduce the size into (28 × 28 × 128) with no trainable parameters, Conv2D_3 is the fourth convolution layer has kernel of (5 × 5) with (64) kernels giving (204,864) trainable parameters with Relu as activation function followed by MaxPool2d_3 with kernel size (2 × 2) which reduce the size into (14 × 14 × 64) with no trainable parameters, Conv2D_4 is the fifth convolution layer has kernel of (5 × 5) with (32) kernels giving (51,232) trainable parameters with Relu as activation function followed by MaxPool2d_4 with kernel size (2 × 2) which reduce the size into (7 × 7 × 32) with no trainable parameters, Conv2D_5 is the sixth

convolution layer has kernel of (5 × 5) with (64) kernels giving (51,264) trainable parameters with Relu as activation function followed by MaxPool2d_5 with kernel size (2 × 2) which reduce the size into (4 × 4 × 64) with no trainable parameters, Conv2D_6 is the seventh convolution layer has kernel of (5 × 5) with (32) kernels giving (51,232) trainable parameters with Relu as activation function followed by MaxPool2d_6 with kernel size (2 × 2) which reduce the size into (2 × 2 × 32) with no trainable parameters, Conv2D_7 is the convolution layer number 8 has kernel of (5 × 5) with (64) kernels giving (51,264) trainable parameters with Relu as activation function followed by MaxPool2d_7 with kernel size (2 × 2) which reduce the size into (1 × 1x64) with no trainable parameters. After this sequence of convolutions and max polling layers a fully connected layer comes with (4,096) node using Relu activation function and receiving an output from the last pooling layer equal to (1 × 1 × 64) so the training parameters equal to (266,240), to reduce the over fitting problem a dropout layer comes with a rate (0.8), second fully connected layer comes with (1,024) connected to the previous layer giving (4,195,328) training parameter with Relu activation function, followed by dropout layer set to (0.8), followed by third fully connected layer with (100) node connected with the previous layer giving (102,500) training parameters with Relu activation function, finally the last fully connected layer comes with softmax activation function at the end with (2) classes to represent the coal and gangue and training parameters equal to (202), a regression layer comes with Adam optimization and using categorical_crossentropy, the total training parameters in the whole model equal to (5,232,750). In this design of CGR-CNN a good extraction of the features has been done to get high accuracy with respect to minimize the GPU RAM used during an acceptable execution time, so to measure the performance of the CGR-CNN these three factors (Execution Time, Execution RAM, Learning Accuracy) are the main comparison factors with the related work.

TABLE 1. The experimental dataset for coal and gangue.

Phase	Coal	Gangue
Training	46	45
validation	14	14
testing	10	10

II. EXPERIMENT

A. EXPERIMENT PLATFORM

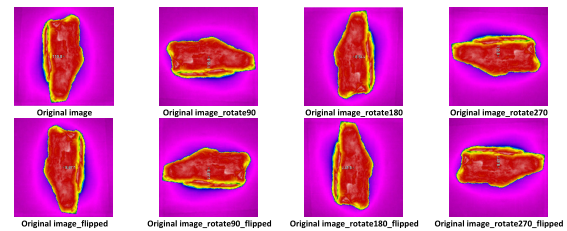
The experiment was done in two hardware platform, the first one was a server platform with 8 core CPU and 16GB RAM, GeForce RTX 2080ti with 12GB and Ubuntu 16.04.5 LTS (Xenial Xerus), this platform was used specially for making the comparison with the related works which need high requirements **Table 3**, the second hardware platform is PC comes with CPU A10 PRO-7800B R7, 12 Compute Cores 4C+8G 3.50 GHz has 4.00 GB installed memory (RAM) and graphic card NVIDIA Quadro K2000 with 2GB memory data rate (Total available graphics memory 4060MB, Dedicated video memory 2048 MB GDDR5, Shared system memory 2012 MB), Windows 10 Enterprise 64-bit Operating System, x64-based processor, this platform to demonstrate that the CGR-CNN works well in ordinary requirements. The software platform used anaconda environment to install python 3.5 with the Tflern library for building the models and performing the learning and testing of the CNN models, the TensorBoard has been used for drawing the models graphs and statical curves of the training accuracy, validation accuracy, training loss and validation loss. For capturing the thermal images FLUKE Ti32 IR FUSION TECHNOLOGY was used to generate fluck thermal image (*.IS2) and SmartView 3.1 application came with it to open and save the thermal images as (*.PNG) files.

B. THE SAMPLES SETS OF THERMAL IMAGES

Thermal images was captured for the Coal/Gangue in certain conditions to increase the difference characters between the two types because of the variability of heat factor taking in account that the Coal/Gangue react by different degrees for the heating environment, a hypothesis has been proposed to say that putting the Coal/Gangue in hot environment and capture the thermal images of the surface will make the classification of the Coal/Gangue more efficient, this hypothesis has been discussed extensively in (SVM-YCbCr) [24], the Coal/Gangue samples have been collected from Bituminous coal, produced in Shanxi Province, western of China, they were put in thermal container until they reach 50 Celsius, after that 139 thermal images (70 coal, 69 gangue) have been captured using the thermal camera which generate thermal images of (.IS2) extension then using the Fluck SmartView 3.1 application the captured thermal images (.IS2) have been converted into PNG images with (680×480) pixels resolution as shown in **Figure 1**, in the experiment the dataset has been divided into three categorizes training (91), validation(28) and testing(20), **Table 1** shows the data set divided between the three phases of the experiment. But number of images still not enough and inevitably will case over fitting problem,

TABLE 2. The experimental dataset for coal and gangue after augmentation process.

Phase	Coal	Gangue
Training	368	360
Validation	112	112
Testing	80	80

**FIGURE 4.** Dataset Augmentation Transformation Processes.

this problem has been addressed by many researchers in there works and to solve the scarcity of image resources an augmentation process performed in the small dataset in order to increase it with respect to generate different pixel values in the same position to make sure that a different image are generated beside the original image, Krizhevsky *et al.* [26] in the **Alexnet** used the augmentation principles to increase the data set in there work, so in order to increase the dataset samples here the augmentation principle has been applied, first the images have been centered and cropped into (480 × 480) pixels resolution to be suitable for augmentation process, then three rotation processes with degrees (90,180,270) have been done and increased the data set from 139 into 556 after that a horizontal inverting has been done to create 1112 images divided in the three categories as explained in **Table 2**, **Figure 4** shows the transformation done to an image and the new generated images and the differences between them.

C. TRAINING AND LEARNING OF THE CGR-CNN

After the designing and implementation of the CGR-CNN two phases have been performed (Learning phase and Testing Phase), in the learning phase two connected steps were performed (Training and Validation), as mentioned before the dataset divided into three separate files (Training 91 images, Validation 28 images and Testing 20 images) **Table 1**, this distribution has been done before the augmentation process, the separation of dataset made the validation phase more accurate and improved the validation loss, during the learning a validation process is important to ensure that the training accuracy getting better result, based on the validation results the training updates weights and biases, so it is important to notify that the images in the three files are different from each other to make sure prevent overlapping of the samples and getting more accurate validation accuracy and testing results, to measure the affect of separation the training and validation dataset, an experiment has been done and **Figure 5** shows the validation loss of two situations (a) represent the validation loss during learning using data samples for learning

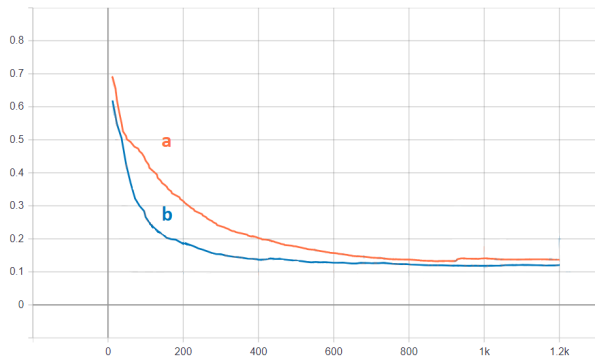


FIGURE 5. Effect of separating dataset before augmentation, (a) validation loss of non-separated dataset. (b) validation loss of separated dataset.

and validation without separation before augmentation process, where in (b) represent the validation loss during learning where the data samples for learning and validation have been separated before the augmentation process so the images in the two files are different, it is noticeable that curve (b) has lower loss values at faster time and tends to stabilize faster than (a), which mean that the separation of the dataset gives a good result in the loss validation of the learning, this result supports the dataset distribution methodology used in this model.

After the preparation of dataset the learning phase starts and Coal/Gangue images input into CGR-CNN for training and validation, the learning rate sets to (0.0001) and the epoch_no sets to (170) with snapshot_step sets to (100) giving iteration equal to (2040). By using accuracy and loss rates of training and validation during the learning to observe the training status and determine whether the network structure is stable and weather the training parameters are appropriate for learning. **Figure 6** shows the learning curves of the CGR-CNN where (a, b) show the accuracy curves of the training and validation respectively and show that the training accuracy at (a) get improved and start to be stable after (328) step and get the best training accuracy of (99.93%) at step (1861) also in (b) validation accuracy varies between (86.16% and 97.77%) and get best validation accuracy about (97.77%) at (1000) step, (c, d) show the loss curves of training and validation respectively also show that learning loss has best value (0.0101) at step(772) and validation loss around (0.0711) at step (996).

After the learning of the CGR-CNN the testing phase has been done using the (160) images in the test file (80 coal, 80 gangue) which as the learning and validation images, these images also have been separated before the augmentation process to ensure accurate prediction. The testing results show that CGR-CNN was able to predict the coal images with no false giving (100%) coal prediction accuracy, on the other hand it predicted the gangue images with (2 out of 80) false predictions which gave (97.5%) gangue prediction accuracy.

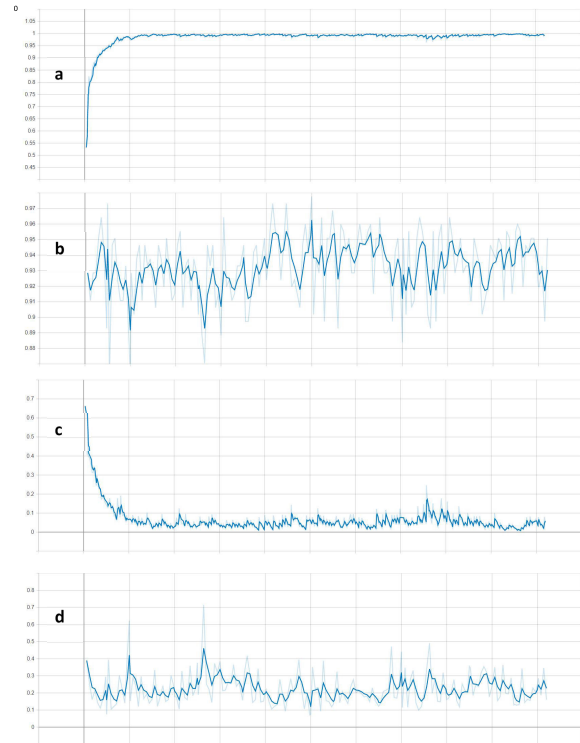


FIGURE 6. CGR-CNN performance:(a) training accuracy, (b) validation accuracy, (c) training loss, (d) validation loss.

III. COMPARISON WITH RELATED WORK

A. COMPARING THE DEVELOPMENT ON CGR-CNN AGAINST SVM-YCbCr

This paper present a development for the work titled “**Separation between Coal and Gangue based on Infrared Radiation and Visual Extraction of the YCbCr Color Space**” by Eshaq *et al.* [24] mentioned here as SVM-YCbCr, by developing a Convolutional Neural Network model for recognition rather than using SVM and feature extraction processes, in the matter of recognition accuracy CGR-CNN raise up the coal recognition accuracy to (100%) compared with (96.6%) in SVM-YCbCr, also achieve a near gangue recognition reached (97.5%) compared to (98.1%) which lead to a recognition accuracy (98.75%) for CGN-CNN compared to (97.83%) in SVM-YCbCr, also in the matter of reducing the execution time using CNN lead to reduce a lot of pre-processing steps which will reduce the latency during the operation time and increase the production efficiency, based on the experiment time(TABLE 7: THE TIME RESULTS OF ACQUISITION, READING, PROCESSING, TRAINING AND PREDICTION OF INFRARED IMAGES) in [24] the acquisition time is for a samples preparation phase which take different amount of coal and gangue in the same time and can be done in a scenario that doesn’t affect the production real time also it will be the same amount of time in both CGR-CNN and SVM-YCbCr, the Training time is a consumed time in the initialization of the system before start using the system in the production real time and it is different on both, but it varies based on different situations

which does not affect the production real time, although if there is a need for retraining the model during the production to raise the efficiency of the system based on feedback situations, there are different fine tuning techniques in the case of CNN making it faster and easier than SVM-YCbCr also this retraining time considered an exceptional case that can be executed while the production line does not work, so it will not affect the real time production. Looking to (Image reading and Processing) and (Prediction time) these two time factors are involving in the real time of production. CGR- CNN have been tested in similar operation environment such as SVM-YCbCr environment, and CGR-CNN achieved for (Image reading and Processing) around (0.0001875)seconds compared to (5.8)seconds in SVM-YCbCr, this huge difference in the time comes from the fact that CGN-CNN actually doesn't do any preprocessing steps that could consume the time like what's happening with CGR-CNN for feature extracting rather than that CNN already learned the features during the training time, also for (Prediction time) CGR- CNN achieved better time (0.0001875) second compared to (0.00097) for Gaussian SVM which had the best prediction Accuracy equal to (97.83%) in SVM-YCbCr [24], it's clear that CGN-CNN demonstrate a good abilities in both recognition accuracy and execution time, on the other hand the ability of improving the output classes in CGN-CNN by adding more classes using techniques such as transfer learning and fine tuning gives it an additional advantage to keep improving with work needs easily and smoothly more than the SVM-YCbCr witch will need a new feature analysis.

B. COMPARED THE PERFORMANCE OF CGR-CNN PREVIOUS RELATED WORK

In comparison with the previous related work and some state-of-art works, the design had attention for three points (speed of learning, the execution environment requirement and the performance accuracy), developing CGR-CNN is taking into account the operational ability in normal PC's with no need for high performance equipments, from the previous related work in Coal/Gangue recognition with CNN and some of the state-of-art (LeNet-5 [25], LeNet-5_improved [19], Alexnet [26], VGG_A and VGG_B [27]) have been implemented and tested with dataset, optimization and learning rate have been set according to the original work, these algorithms chosen according to some aspects, LeNet-5 for classical CNN network simple structure and fast in execution, LeNet-5_improved previous related Coal/Gangue work using CNN, Alexnet one of previous state-of-art work which had the best results with the thermal images dataset, VGG_A and VGG_B state-of-art CNN with high structure, **Table 3** shows the comparison results for training experiment with epoch number equal to (170) epoch and (2040) steps. It's clear that performance varies from model to another, some models had a poor reaction with the thermal images dataset (Lnet5_improved, VGG_A and VGG_B), **Figure 7** shows the poor reaction of them at (a, b) they got the lowest overlapped three curves at (50%) rate for leaning accuracy and validation accuracy

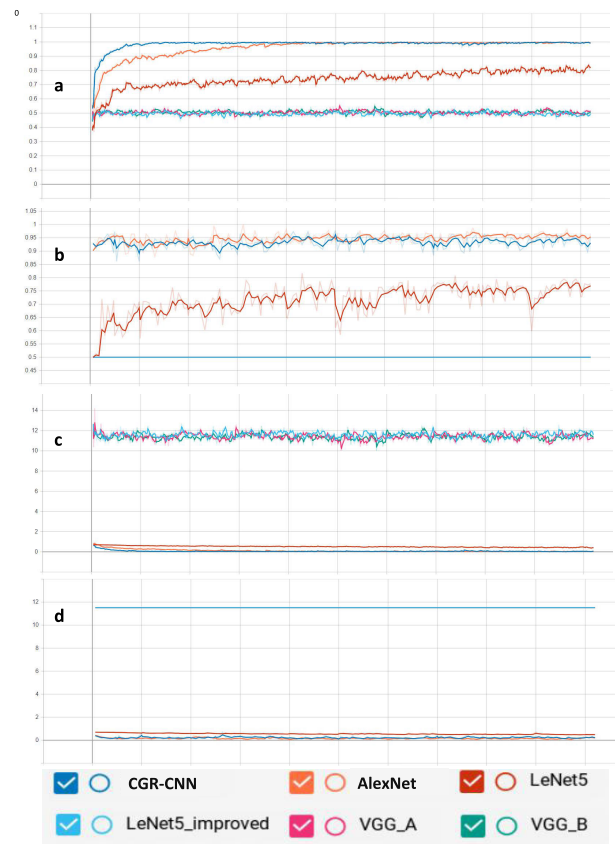


FIGURE 7. CNN models comparison:(a)training accuracy. (b)validation accuracy.(c)training loss. (d)validation loss.

also in (c, d) they got highest overlapped three curves around the (12) for learning loss and validation loss, it's clear from results that the models behave like when they just had initial values and didn't get any updates for the weights and biases values so in the testing phase the models intend to classify all the test samples from one class but with poor prediction values getting C_{test} with 100% recognition which is faulty results, this is because these models have been developed and tested to perform with high number of epoch which is out of the time scope in the comparison here because it will consume more time.

C. PERFORMANCE OF CGR-CNN VS ALEXNET

The other models had good reaction with the thermal images dataset, the LeNet-5 gives good reaction with the data set but still not enough to gives excellent results because the simplicity of the structure, it had the best timing of execution in the comparison but the learning accuracy did not get more than (83%) and the testing get around (66.3%) even with increasing the epoch numbers to (600) getting learning time near to the best result of Alexnet and CGR-CNN, however LeNet-5 still in low learning rate this is because the input in LeNet-5 so small ($28 \times 28 \times 1$) therefore extracting features from thermal images with this size was not good enough to get good results. The comparison shows that Alexnet and

TABLE 3. Performance Comparison where **L_acc** learning accuracy, **V_acc** validation accuracy, **L_loss** learning loss, **V_loss** validation loss, **C_test** prediction of coal test samples, **G_test** prediction of gangue test samples, **P** overall prediction of test samples, **T_s** learning time in seconds.

model	L_acc	V_acc	L_loss	V_loss	C_test	G_test	P	T_s
Alexnet [26]	99.56%	96.43%	0.037	0.16	100%	98.75%	99.37%	8379
CGR-CNN%	99.36%	95.09%	0.049	0.16	100%	97.5%	98.75%	5831
LeNet-5 [25]	82.84%	77.68%	0.408	0.469	80%	52.5%	66.25%	297
LeNet-5_improved [19]	50.27%	50%	11.45	11.51	100%	0.0%	50%	16552
VGG_A [27]	51.4%	50%	11.32	11.51	100%	0.0%	50%	13822
VGG_B [27]	50.2%	50%	11.47	11.51	100%	0.0%	50%	16720

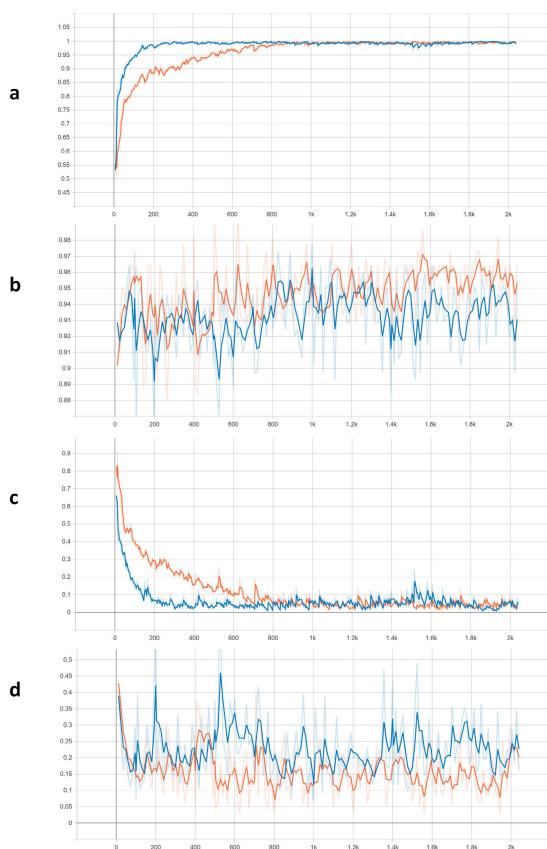


FIGURE 8. Performance Comparison between CGR-CNN and Alexnet (a) training accuracy. (b) validation accuracy. (c) training loss. (d) validation loss. The blue curve represent the CGR-CNN and the orange curve represent the Alexnet.

the CGR-CNN gave an excellent performance and they were able to recognize the Coal/Gangue with more than (98%), but Alexnet still consume (1.5) learning time more than the CGR-CNN, also consume more GPU RAM for execution with (62,378,344) training parameters which around (12 times) of the CGR-CNN with (5,232,750) training parameters, it means that the CGR-CNN will be more flexible and executable in more less equipments than what Alexnet model need. **Figure 8** shows the accuracy and loss of the learning of the two models blue curves represents CGR-CNN and orange curves represent Alexnet, in (a) learning accuracy curves show that CGR-CNN learning rates get improved faster than Alexnet and stabilized earlier (b) validation accuracy curves show that both CGR-CNN and Alexnet approximately same with better

result with Alexnet, (c) learning loss curves also show that CGR-CNN get improved quickly, (d) validation loss curves also approximately same in the two models.

IV. CONCLUSION

This paper addressed the topic of using the convolution neural network in the matter of Coal/Gangue recognition and built new deep convolutional neural network model called CGR-CNN for recognition Coal/Gangue using dataset of Coal/Gangue thermal images based on previous work [24] which are considered suitable for industries with hot environment such as power plants because it will easily provide suitable hot containers to be used in the heating phase of the samples.

The experimental result shows that the model during the learning phase achieved an excellent training accuracy near to(99.93%) and validation accuracy reached (97.77%), in the prediction phase the model was able to predict the coal images with no false predictions giving (100%) accuracy and the gangue images had (97.5%) with (2 out of 80) images false prediction with overall prediction around (98.75%), these false predictions was attributed to the presence of samples with mixed materials of coal and gangue so it is recommended to add one more class for the mixed samples and this will be future work. CGR-CNN demonstrate high abilities against the SVM-YCbCr in the recognition accuracy CGR-CNN with coal recognition (100%) compared with (96.6%), gangue recognition(97.5%) compared to (98.1%) making recognition accuracy (98.75%) for CGN-CNN vs(97.83%) in SVM-YCbCr, also CGR-CNN demonstrate superior performance in the (Image reading and Processing) and (Prediction time) with time around (0.00038s) against (5.80097s) in SVM-YCbCr.

It's clear that using large models in classifying particular objects doesn't lead to good results in the medium-term working environments and using smaller and lighter models could give a good and more efficient work results. Also in the case of using the augmentation process to increase the learning dataset, separation the input data for training and validation before augmentation shows a better results.

REFERENCES

[1] B. Wang, C.-Q. Cui, Y.-X. Zhao, B. Yang, and Q.-Z. Yang, "Carbon emissions accounting for China's coal mining sector: Invisible sources of climate change," *Natural Hazards*, vol. 99, no. 3, pp. 1345–1364, Dec. 2019.

- [2] Z. Ze-Lin, Y. Jian-Guo, W. Yu-Ling, X. Wen-Cheng, L. Xiang-Yang, and W. Xing-Xing, "Fast predicting the washability of coal using digital image processing method," in *Proc. Int. Conf. Informat., Cybern., Comput. Eng. (ICCE)*, Melbourne, VIC, Australia, L. Jiang, Ed. Berlin, Germany: Springer, 2012, pp. 383–390.
- [3] W. Li, Y. Wang, B. Fu, and Y. Lin, "Coal and coal gangue separation based on computer vision," in *Proc. 5th Int. Conf. Frontier Comput. Sci. Technol.*, Aug. 2010, pp. 467–472.
- [4] L. Li, Y. Zhang, Y. Zhang, J. Sun, and Z. Hao, "The thermal activation process of coal gangue selected from zhungeer in China," *J. Thermal Anal. Calorimetry*, vol. 126, no. 3, pp. 1559–1566, Dec. 2016, doi: 10.1007/s10973-016-5711-4.
- [5] D. M. Hobson, R. M. Carter, Y. Yan, and Z. Lv, "Differentiation between coal and stone through image analysis of texture features," in *Proc. IEEE Int. Workshop Imag. Syst. Techn.*, May 2007, pp. 1–4.
- [6] Y. Pu, D. B. Apel, A. Szmigiel, and J. Chen, "Image recognition of coal and coal gangue using a convolutional neural network and transfer learning," *Energies*, vol. 12, no. 9, p. 1735, 2019.
- [7] R. Wang and Z. Liang, "Automatic separation system of coal gangue based on DSP and digital image processing," in *Proc. Symp. Photon. Optoelectron. (SOPO)*, May 2011, pp. 1–3.
- [8] X.-M. Ma, "Improved EZW compression coding algorithm for coal gangue image," in *Proc. 2nd Int. Congr. Image Signal Process.*, Oct. 2009, pp. 1–4.
- [9] X.-M. Ma and C. Liang, "Application of rough set theory in coal gangue image process," in *Proc. 5th Int. Conf. Inf. Assurance Secur.*, vol. 1, Aug. 2009, pp. 87–90.
- [10] X. Ma and J. Zang, "Coal gangue image process approaches with wavelet analysis," in *Proc. Congr. Image Signal Process.*, vol. 4, May 2008, pp. 352–356.
- [11] H. Hong, L. Zheng, J. Zhu, S. Pan, and K. Zhou, "Automatic recognition of coal and gangue based on convolution neural network," *CoRR*, 2017. [Online]. Available: <http://arxiv.org/abs/1712.00720>
- [12] X.-M. Ma, "A revised edge detection algorithm based on wavelet transform for coal gangue image," in *Proc. Int. Conf. Mach. Learn. Cybern.*, vol. 3, Aug. 2007, pp. 1639–1642.
- [13] M. Li and K. Sun, "An image recognition approach for coal and gangue used in pick-up robot," in *Proc. IEEE Int. Conf. Real-Time Comput. Robot. (RCAR)*, Aug. 2018, pp. 501–507.
- [14] M.-H. Zhao, S. Ma, and D.-X. Zhao, "Image processing based on gray information in sorting system of coal gangue," in *Proc. 10th Int. Conf. Intell. Hum.-Mach. Syst. Cybern. (IHMSC)*, vol. 2, Aug. 2018, pp. 81–83.
- [15] X.-R. Song and F.-J. Wang, "Research on coal gangue on-line automatic separation system based on the improved BP algorithm and ARM," in *Proc. Int. Conf. Mach. Learn. Cybern.*, vol. 5, Aug. 2007, pp. 2897–2900.
- [16] K. Gao, C. Du, H. Wang, and S. Zhang, "An efficient of coal and gangue recognition algorithm," *Int. J. Signal Process., Image Process. Pattern Recognit.*, vol. 6, no. 4, pp. 345–354, Aug. 2013.
- [17] X.-M. Ma, "Coal gangue image identification and classification with wavelet transform," in *Proc. 2nd Int. Conf. Intell. Comput. Technol. Automat.*, vol. 1, Oct. 2009, pp. 562–565.
- [18] Q. Mu and J.-X. Dong, "The application of coal cleaning detection system based on embedded real-time image processing," in *Proc. 5th Int. Conf. Measuring Technol. Mechatronics Automat.*, Jan. 2013, pp. 1125–1127.
- [19] L. Su, X. Cao, H. Ma, and Y. Li, "Research on coal gangue identification by using convolutional neural network," in *Proc. 2nd IEEE Adv. Inf. Manag., Commun., Electron. Automat. Control Conf. (IMCEC)*, May 2018, pp. 810–814.
- [20] H. Yaqun, H. Jingfeng, Z. Nianxin, C. Bo, and L. Haonan, "Notice of retraction: Research on identification of coal and waste rock based on PCA and GA-ANN," in *Proc. 3rd Int. Conf. Comput. Sci. Inf. Technol.*, vol. 6, Jul. 2010, pp. 579–584.
- [21] H. Liang, H. Cheng, T. Ma, Z. Pang, and Y. Zhong, "Identification of coal and gangue by self-organizing competitive neural network and SVM," in *Proc. 2nd Int. Conf. Intell. Hum.-Mach. Syst. Cybern.*, vol. 2, Aug. 2010, pp. 41–45.
- [22] D. P. Tripathy and K. G. R. Reddy, "Novel methods for separation of gangue from limestone and coal using multispectral and joint color-texture features," *J. Inst. Eng. India, D*, vol. 98, no. 1, pp. 109–117, Apr. 2017, doi: 10.1007/s40033-015-0106-4.
- [23] Y. Li, G. Cheng, X. Chen, and C. Liu, "Coal-Rock interface recognition based on permutation entropy of LMD and supervised Kohonen neural network," *Current Sci.*, vol. 116, no. 1, pp. 96–103, 2019.
- [24] R. M. A. Eshaq, E. Hu, M. Li, and M. S. Alfarzaei, "Separation between coal and gangue based on infrared radiation and visual extraction of the YCbCr color space," *IEEE Access*, vol. 8, pp. 55204–55220, 2020.
- [25] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
- [26] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Red Hook, NY, USA: Curran Associates, 2012, pp. 1097–1105. [Online]. Available: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>
- [27] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: <http://arxiv.org/abs/1409.1556>



China. His research interests are in neural networks and deep learning.



artificial intelligence and pattern recognition, machine learning, and data mining and other aspects of research.



His current research interests include multiobjective optimization, machine learning deep learning, and image processing.



REFAT MOHAMMED ABDULLAH ESHAQ was born in Sana'a, Yemen, in 1993. He received the B.S. degree in mechanical engineering from the Arab Academy for Science, Technology, and Maritime Transport, Alexandria, Egypt, in 2018. In same year, he has been awarded scholarship by China Scholarship Council (CSC) to resume postgraduate. He is currently pursuing the M.S. degree in major: mechanical and electronic engineering with the School of Mechanical and Electrical Engineering, China University of Mining and Technology, Xuzhou, China. The field of his interests include the heat transfer, gas dynamics, and fluid mechanics application in intelligent industry.



ERYI HU received the Ph.D. degree from the School of Mechanical Engineering, Huazhong University of Science and Technology, Wuhan, China. His current position is an Associate Professor and the Director of the Mine Robot Research Institute, School of Mechatronic Engineering, China University of Mining and Technology, Xuzhou, China. His research interests include mining robot design, machine vision, and optical measurement technique.

...