A deep learning approach for automated quality control of iron ores

A.K. Gorai, B.C. Balusa

Department of Mining Engineering, National Institute of Technology, Rourkela-769008, India

U. Sameer

Department of Computer Science & Engineering, National Institute of Technology, Rourkela-769008, India

ABSTRACT: The present study aims to design a machine vision system using deep learning algorithm for quality monitoring of iron ores. A total of 53 image samples were used for model calibration and testing. The model was trained using 45 image samples and tested using 9 image samples. The model parameters like the number of nodes and number layers were optimized based on the root mean squared error (RMSE) values. It was observed that the RMSE was lowest for the network architecture having 5-nodes and 3-hidden layers. The performance of the optimized model was evaluated using four indices including RMSE, normalized mean square error (NMSE), R-squared, and bias. The RMSE, NMSE, R-squared, and bias of the optimized model were obtained as 8.77, 0.0026, 0.87, and -1.14 respectively. The results indicate that the model gives satisfactory performance in quality predictions of iron ores.

1 INTRODUCTION

The quality of ore in terms of physical (size, shape, strength, etc.) and chemical (composition, grade, etc.) properties specifically defined for the extraction of valuable material through particular treatment process (Ivanov, 1986). The underlying task of predicting a continuous value falls under the category of regression in the field of machine learning. It is the process of studying various characteristics of the data and fitting all the data points into one model. This was performed first by identifying various distinct characteristics of the data known as the features. Then, a regression algorithm such as linear regression (Neter, 1996) or a logistic regression (Hosmer, 2013) was applied, and the regression value of the test data was predicted. The success of these machine learning techniques is hugely dependent on the choice of the features extracted, thus limiting the effectiveness of the regression algorithms. Thus a deep learning algorithm has been designed for prediction of continuous data without any specific feature extraction from the image samples. Deep Learning is an emerging technique for performing representation learning. That is, information from the data is learned using artificial intelligence in the form of a deep neural network. Unlike machine learning techniques which are largely dependent on the features extracted from the data, deep learning techniques learn the features by themselves and perform the necessary prediction task using the learned features. It is the independent nature of deep learning that makes it more effective and superior to machine learning based models when identification and extraction of the feature are not easy. In this work, we perform the task of regression using a deep learning multi-layer network using the raw images of minerals as the input.

Deep learning approach of neural network has been used in many different applications for classification and regression analysis. Baccouche et al. (2011) proposed sequential deep learning for human action recognition. This study developed a fully automated deep learning model, which learns to classify human actions without using any prior knowledge. Yu and Deng (2011) reviewed deep learning and its applications in signal and information processing. Huang et al. (2014) developed a deep learning model for speech separation. This study explored the use of deep neural networks (DNN) and a recurrent neural network (RNN) for monaural speech separation in a supervised setting. Huval et al. (2015) proposed an empirical evaluation of deep learning on highway driving. In this study, a large data set of highway data was collected and apply deep learning and computer vision algorithms to problems such as car and lane detection. Aliper et al. (2016) demonstrated how deep neural networks trained on large transcriptional response data sets could classify various drugs to therapeutic categories solely based on their transcriptional profiles. Shen at al. (2017) reviewed the recent advances of deep learning to identify, classify, and quantify patterns in medical images. Levine et al. (2018) described a deep learning-based approach to hand-eye coordination for robotic grasping from monocular images. In this study, a large convolutional neural network was trained to learn hand-eye coordination for grasping.

Many researchers used traditional artificial intelligence approaches like artificial neural networks (ANN), support vector machines (SVM), genetic algorithms (GA) for classification ores in the mineral industry. Till date, deep learning neural networks are not used in ore classification in the mineral industry. The present study developed feed forward deep learning neural network model for classifications of Iron ore samples.

2 MATERIALS AND METHODOLOGY

The proposed deep learning algorithm for a machine vision system was developed in three major stages viz. image acquisition, model development, and model evaluation. The steps are represented in a flowchart (shown in Figure 1). The detailed description of each step is given below.

Figure 1. Broad Steps of the Proposed Work

2.1 Image acquisition

A small-scale conveyor-based transportation system was designed and fabricated in the laboratory for image acquisition of iron ore samples. A camera (Logitech HD Webcam c310) was fixed above the conveyor belt for image acquisition. Two sets of LED light were fixed for illumination. These were mounted on 45º inclination to reduce the reflectance, the material to the image capturing device. The conveyor system was powered by a 0.5 horsepower motor made of Crompton Greaves for driving the rollers. The iron ore samples collected from mine were continuously fed at the inlet point of the conveyor belt transportation system. The images of the ores were captured during transportation of ore from the inlet to the outlet point of the belt. A total of 53 images were captured for different types of ore samples (shown in Figure 2). The grade values of the iron ores corresponding to the captured images were estimated using the XRF analyses and summarized in Table 1.

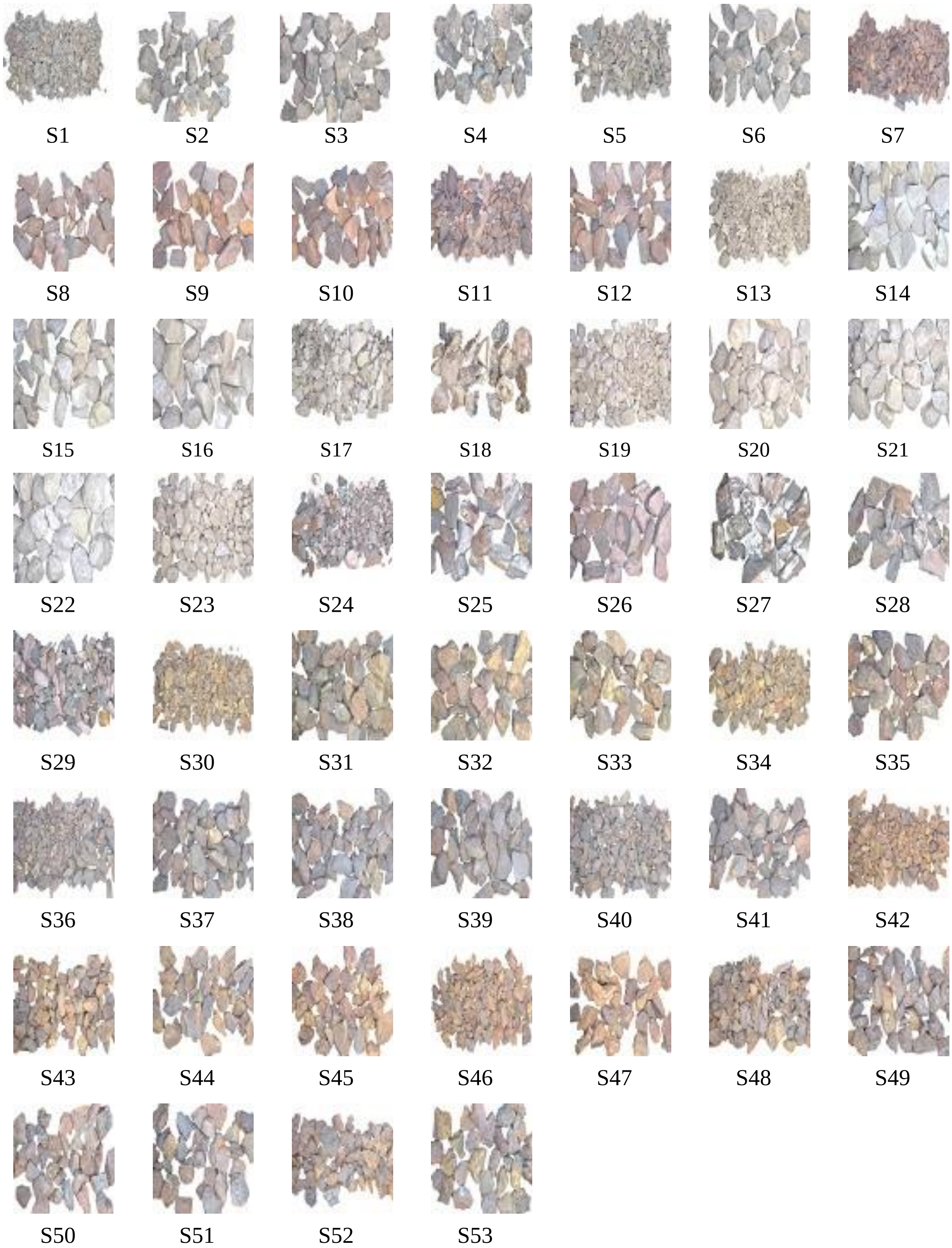

Figure 2. Images of different types of ore samples

Table 1. Observed grade values of iron ore samples corresponding to 53 images

Sample ID	Grade Value	Sample ID	Grade Value	Sample ID	Grade Value	Sample ID	Grade Value	Sample ID	Grade Value
S1	59.71	S13	53.73	S25	29.88	S37	28.47	S49	60.06
S2	64.36	S14	58.35	S26	29.49	S38	32.46	S50	65.87
S3	66.57	S15	68.57	S27	37.07	S39	30.06	S51	59.54
S4	62.32	S16	66.32	S28	36.66	S40	28.08	S52	65.96
S5	67.72	S17	68.00	S29	49.64	S41	30.77	S53	67.96
S6	68.03	S18	65.21	S30	30.89	S42	57.14		
S7	61.38	S19	66.20	S31	23.55	S43	68.47		
S8	57.76	S20	59.16	S32	15.75	S44	68.80		
S9	66.43	S21	63.19	S33	13.82	S45	63.79		
S10	67.95	S22	69.54	S34	25.64	S46	66.73		
S11	64.63	S23	68.54	S35	22.77	S47	64.67		
S12	68.74	S24	66.43	S36	48.51	S48	66.58		

2.2 Model development

The flowchart of the model development is shown in Figure 3. All the captured images were uniformly sized into 64 x 64 and store as a dataset for learning of the deep network. The deep network was trained with the 80% (44 image samples) of image samples and tested using 20% (=9) of the image samples.

In the current study, a multi-layer perceptron (MLP) (Ruck, 1990) based deep neural network was designed for prediction of ore grade. The architecture of the network is shown in Figure 4. The proposed network consists of one input layer, one output layer along with the hidden layers. The number of hidden layers in the network is changed from 1 to 5 for optimization. The input layer process the input data of image samples, the hidden layers are used for feature learning, and the output layer provides the predictions.

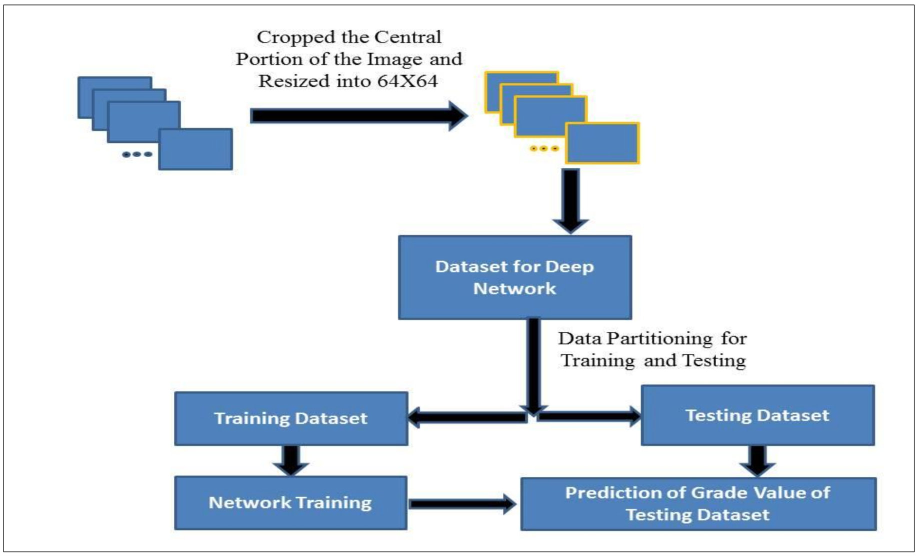

Figure 3. Flowchart of the Model Development

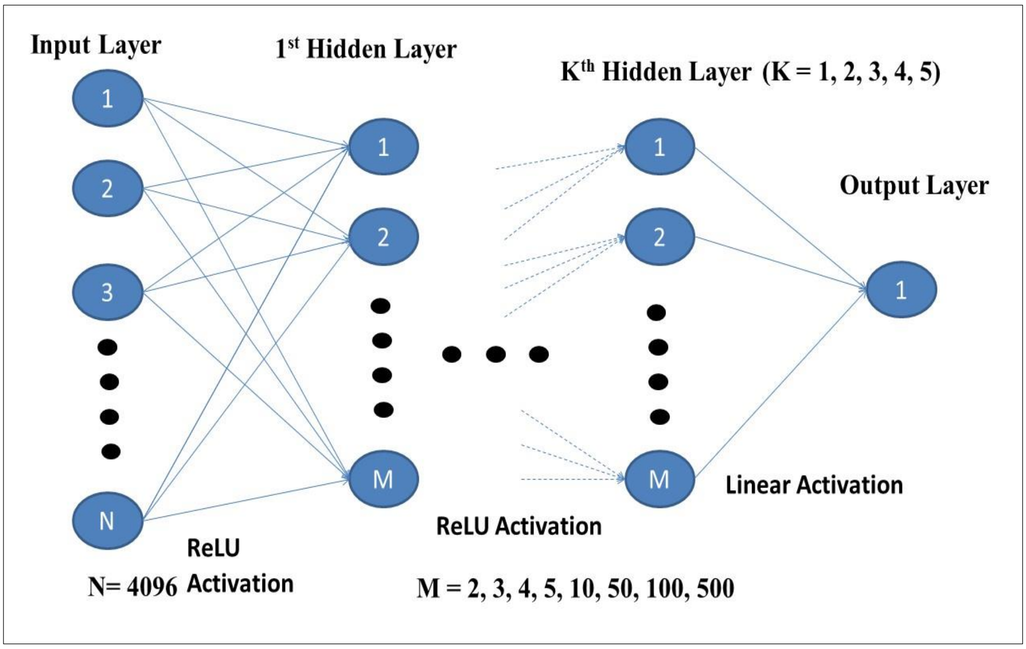

Figure 4. Proposed Network Architecture

The input layer consists of 4096 nodes representing each pixel of the 64×64 image. The network performances are examined for different numbers of nodes (2, 3, 4, 5, 10, 50, 100, and 500) in the hidden layers for optimization of the model. The final layer in the proposed network has a single node to predict the output value for the particular input. Two types of activation functions are used in the network, one with Rectified Linear Unit (ReLU) activation and the other with a linear activation. The ReLU activation (Dahl, 2013) is used for the input layer and all the hidden layers except last. The linear activation function is used for the last hidden layer. The role of an activation function is to produce a non-linear transformation of the input. The ReLU activation is defined as follows

$$ReLU(x) = \log(1 + e^{x}) \quad (1)$$

where 'x' represents the input value of a node

In the linear activation, the output is the same as the input. It can be defined as follows:

$$Linear(x) = x \quad (2)$$

The most important module in a deep learning architecture is the role of the optimizer. A deep neural network is initialized with random weights, and the weights are updated iteratively to achieve particular learning (in this case, regression). The weight updating is performed by finding a cost function J(Θ) where Θ is the parameter space of the network. The main function of an optimizer is finding the best values of Θ suitable for the task. Adam optimizer (Kingma, 2014) is used to perform learning. The Adam optimization algorithm is an extension to stochastic gradient descent (SGD) that has maintained a single learning rate for all weight updates, and the learning rate does not change during training. Adam also makes use of the average of the second moments of the gradients (the uncentered variance).

2.3 Model Evaluations

The model performance was evaluated with the testing datasets using four error indices viz. root mean square error (RMSE), normalized mean square error (NMSE), R-square, and bias. All the indices were determined from the predicted grade (represent by *o*) and observed grades (represented by *p*) of the testing samples using the following respective equations.

$$RMSE = \sqrt{\frac{1}{n_{ts}}\sum_i (o_i - p_i)^2} \tag{3}$$

$$NMSE = \frac{1}{n_{ts}} \frac{\sum_i (o_i - p_i)^2}{\overline{o_i}\,\overline{p_i}} \tag{4}$$

$$R^2 = \frac{\left(\sum\left(o_i - \overline{o_i}\right)\left(p_i - \overline{p_i}\right)\right)^2}{\sum\left(o_i - \overline{o_i}\right)^2 \sum\left(p_i - \overline{p_i}\right)^2} \tag{5}$$

$$Bias = \frac{1}{n_{ts}}\sum_i (o_i - p_i) \tag{6}$$

where o_i and p_i are the observed and predicted grade of i^{th} samples. n_{ts} represents the number of testing samples. The value of $\overline{o_i}$ and $\overline{p_i}$ can be determined as

$$\overline{o_i} = \frac{1}{n_{ts}}\sum_i o_i \quad and \quad \overline{p_i} = \frac{1}{n_{ts}}\sum_i p_i$$

3 RESULTS AND DISCUSSION

The proposed deep learning algorithm is based on multi-layer perceptron (MLP). In MLP, the network is initialized with random weights. The input image samples of iron ores are passed through the input layer forming the vector, x_i. The random weights at the input layer are multiplied and added to a bias, b. The model used Adam optimizer for learning rate optimization. The resultant vector then passed through an activation function to produce a non-linear value at the end of the layer. The deep learning algorithm is applied to develop a machine vision system for prediction of quality of iron ores. The iron ore samples were collected from mine using a stratified random sampling method. The samples are collected from the different part of the mine to represent the heterogeneity of the geology present in the specific mine. The images of the iron ore samples are captured in the laboratory by fabricating a pilot conveyor belt transportation system. A total of 53 image samples were captured for the training and testing of the deep network. The grade values of the iron ore samples are estimated using XRF analyzer correspond to each image sample. The grade values (Fe_2O_3 %) of the samples are ranged from 19.35% to 97.36%. A wide variation in the grades of iron ores indicates that the quality of ore are varied in the different part of the mine and thus a proper quality monitoring system in the mine is desired.

The deep network was trained using 44 image samples and tested using 9 image samples. The network used ReLU activation function for the input layer and all the hidden layers except last where a linear activation function is used. In the present study, the numbers of hidden layer and the number of nodes in each hidden layer are optimized based on the model performances results. The performance of the network is investigated separately for the different number of hidden layers (1, 2, 3, 4, and 5) with the different number of nodes in the hidden layers (2, 3, 4, 5, 10, 50, 100, and 500). In the present study, 40 Deep network models (M11, M12, …M15, M21, …M25, M31…M35, M41…M45, M51….M55, M61…M65, M71….M75, M81…M85) are trained and tested. The performance of the developed model was evaluated using the four indices (RMSE, NMSE, R_2, and bias) as explained in Section 2.4. The model performance indices for different combinations of layer and node are summarized in Table 2.

Table 2. Model performance results under different conditions

Model Name	Number of Nodes in each Hidden Layer	Factors	Number of layers				
			1	2	3	4	5
M11 to M15	2	R^2	0.8142	0.8037	0.7983	0.8017	0.8234
		RMSE	10.06	10.16	10.14	10.32	10.24
		Bias	-1.82	-1.27	-1.19	-1.33	-1.32
		NMSE	0.03	0.03	0.03	0.03	0.03
M21 to M25	3	R^2	0.8154	0.8173	0.8217	0.8206	0.8097
		RMSE	9.98	10.00	9.94	10.04	10.10
		Bias	-0.94	-0.84	-1.51	-1.00	-1.78
		NMSE	0.03	0.03	0.03	0.03	0.03
M31 to M35	4	R^2	0.8022	0.7876	0.7826	0.8140	0.8096
		RMSE	10.13	10.36	10.50	10.18	9.99
		Bias	-1.19	-1.47	-1.96	-1.14	-1.84
		NMSE	0.03	0.03	0.03	0.03	0.03
M41 to M45	5	R^2	0.7885	0.8064	0.8701	0.8145	0.8109
		RMSE	10.33	10.13	8.77	9.78	10.08
		Bias	-1.19	-1.55	-1.14	-1.81	-1.31
		NMSE	0.03	0.03	0.02	0.03	0.03
M51 to M55	10	R^2	0. 8758	0.8051	0.8134	0.8128	0.8164
		RMSE	9.87	10.15	10.06	10.05	10.21
		Bias	-3.47	-1.44	-1.20	-1.44	-1.77
		NMSE	0.03	0.03	0.03	0.03	0.03
M61 to M65	50	R^2	0.7940	0.8259	0.8542	0.8163	0.8261
		RMSE	10.27	10.21	9.60	10.23	9.78
		Bias	-1.67	-2.25	-0.86	-2.50	-1.53
		NMSE	0.03	0.03	0.03	0.03	0.03
M71 to M75	100	R^2	0.8113	0.8542	0.7662	0.7442	0.7445
		RMSE	10.11	9.69	10.38	10.74	10.52
		Bias	-1.33	-1.99	-1.49	-1.40	-0.98
		NMSE	0.03	0.03	0.03	0.04	0.03
M81 to M85	500	R^2	0.7881	0.8720	0.8546	0.8224	0.7432
		RMSE	10.81	9.59	9.43	10.22	10.30
		Bias	-1.68	-1.23	-1.69	-0.98	-0.37
		NMSE	0.04	0.03	0.03	0.03	0.03

All the indices are determined using the predicted grades and observed grades of the testing samples. The results indicated that the NMSE value of the model M43 is lowest for the 3-hidden layer and 5-nodes in each hidden layer. The lowest NMSE indicates that the observed grade values are more closely matched with the predicted values. The trend of RMSE and R^2 values are also represented in Figure 5 and Figure 6 respectively. It is observed that the RMSE value is lowest for model M43.The trend of RMSE clearly indicates that the performance of the M43 model is best out of all. Furthermore, the trend of R^2 shows that the regression coefficient of the M43 model (=0.8701) is nearly equal to the highest R^2 value observed for model M51 (=0.8758). The R^2 value indicates that the correlation between the observed and predicted values is good for the M43 model. The biases for all the models are found to be relatively high and negative. A negative bias indicates that the model performs with over prediction and thus the models need to be trained and tested using more number of image samples.

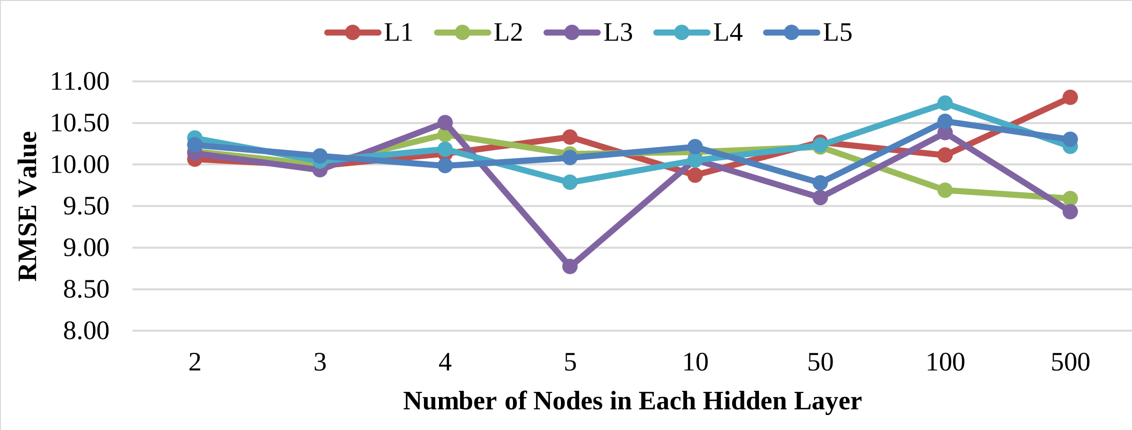

Figure 5. Trends of RMSE of testing samples for the different model condition

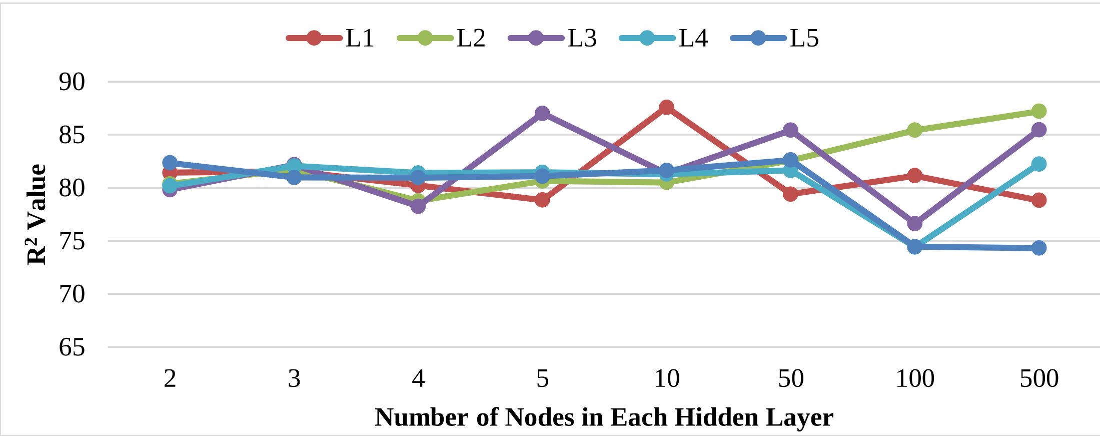

Figure 6. Trends of R^2 of testing samples for the different model condition

4 CONCLUSIONS

The proposed study attempts to develop an MLP-based deep learning network model for prediction of quality of iron ores. The performance of the network is investigated for the different number of hidden layers (1, 2, 3, 4, and 5) and the different number of nodes in each hidden layer (2, 3, 4, 5, 10, 50, 100, and 500). It is observed that the RMSE is lowest for the model having 5-nodes and 3-hidden layers. The performance of the models was evaluated using four indices including RMSE, normalized mean square error (NMSE), R-squared (R^2), and bias. The RMSE, NMSE, and R-squared of the optimized model were obtained as 8.77, 0.0026, and 0.87 respectively. The results indicate that the model gives satisfactory performance in quality predictions of iron ores.

The study was performed with less number of image samples, and the performance can be further improved by increasing the number of samples. The proposed technology of quality inspection may be more economical and robust. Thus, the system can play a significant role in automation in the mineral industries. The bias of the optimized model is found to be relatively high (1.14) and negative. A negative bias indicates that the model performs with over prediction and thus the models need to be trained and tested using more number of image samples.

REFERENCES

Aliper, A., Plis, S., Artemov, A., Ulloa, A., Mamoshina, P., & Zhavoronkov, A. (2016). Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. Molecular pharmaceutics, 13(7): 2524-2530.

Baccouche, M., Mamalet, F., Wolf, C., Garcia, C., & Baskurt, A. (2011). Sequential Deep Learning for Human Action Recognition. In: Salah A.A., Lepri B. (eds) Human Behavior Understanding. HBU 2011. Lecture Notes in Computer Science, vol. 7065. Springer, Berlin, Heidelberg.

Dahl, G. E., Sainath, T. N., & Hinton, G.E. (2013). Improving deep neural networks for LVCSR using rectified linear units and dropout. IEEE conference on Acoustics, Speech and Signal Processing (ICASSP), 26-31 May 2013, Vancouver, BC, Canada. Doi. 10.1109/ICASSP.2013.6639346

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression, Third Edition, John Wiley & Sons Inc. DOI:10.1002/9781118548387

Huang, P-S., Kim, M., Hasegawa-Johnson, M., & Smaragdis, P. (2014). Deep learning for monaural speech separation. IEEE conference on Acoustics, Speech and Signal Processing (ICASSP), 4-9 May 2014, Florence, Italy.

Huval, B., Wang, T., Tandon, S., Kiske, J., Song, W., Pazhayampallil, J., Andriluka, M., Rajpurkar, P., Migimatsu, T., Cheng-Yue, R., Mujica, F., Coates, A., & Ng, A.Y. (2015). An empirical evaluation of deep learning on highway driving. arXiv preprint arXiv:1504.01716.

Ivanov, O. P. (1986). Methodological aspects of ore quality control. Soviet Mining Science, 22(3): 207 - 212.

Kingma, D. P., & Ba, J. L. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kunter, M. H., Nachtsheim, C. J., Neter, John, Li, W. (2005). Applied linear statistical models, 5th Edition, McGraw Hill-Irwin, New York, NY, 10020.

Levine, S., Pastor, P., Krizhevsky, A., Ibarz, J., & Quillen, D. (2018). Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. The International Journal of Robotics Research, 37(4-5): 421-436.

Ruck, D. W., Rogers, S. K., Kabrisky, M., Oxley, M.E., Suter, B. W. (1990). The multilayer perceptron as an approximation to a Bayes optimal discriminant function. IEEE Transactions on Neural Networks 1(4): 296-298.

Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. Annual review of biomedical engineering, 19: 221-248.

Yu, D., & Deng, L. (2011). Deep learning and its applications to signal and information processing. IEEE Signal Processing Magazine, 28(1): 145-154.