Prediction of Froth Flotation Performance Using Convolutional Neural Networks

A. Jahedsaravani[1] · M. Massinaei[2] · M. Zarie[1]

Abstract

Deep learning is a subset of machine learning that uses artificial neural networks for extracting high-level features from image data. In the present study, a soft sensor is proposed for the prediction of the flotation performance through froth features generated by the use of pre-trained convolutional neural networks. Several state-of-the-art convolutional neural networks (AlexNet, GoogLeNet, VGGNet, ResNet, and SqueezeNet) pre-trained on the ImageNet database are used to predict the metallurgical performance of two flotation systems. The first case study is a batch copper flotation system video-captured over a wide range of process conditions. The second case study is an industrial coal flotation column equipped with a continuous video recording system. The pre-trained networks are used to extract features from the froth images, and these features are subsequently used to predict the flotation conditions and performance. The prediction results by the pre-trained algorithms were compared with the traditional image processing algorithms. This demonstrates the ability of the pre-trained structures to generalize to images outside the ImageNet database. GoogLeNet outperforms other network architectures and provides more accurate predictions of the flotation process behavior and performance.

Keywords Froth flotation · Image analysis · Deep learning · Convolutional neural networks

1 Introduction

Froth flotation is a well-established beneficiation process that is widely used for the separation of the value from the unwanted gangue minerals in mineral processing plants [1]. Despite extensive research in modeling and simulation of the flotation process and recent developments in instrumentation and process control systems, optimizing control systems (model-based controllers and expert systems) has not been successfully implemented in flotation plants [2].

Previous studies have shown that the froth visual and textural features are closely related to working conditions and the performance of the flotation process [3–8]. Hence, several soft sensors were developed in the last few decades to measure the froth visual and textural characteristics in flotation plants. These sensors use machine vision learning algorithms and process data for modeling and predicting the flotation performance [3, 4, 9–15].

Advances in machine learning and computer systems led to the development of deep learning algorithms. Deep learning algorithms are capable of capturing relevant features from an image at different levels similar to a human brain [16, 17]. The convolutional neural network is a class of deep neural networks that have been successfully applied for image processing of froth flotation [18–22]. In a machine vision system for froth flotation, convolutional neural networks are typically used to directly and continuously extract features from froth images, which often yield better results than traditional approaches. The extracted features can be used for the classification of froth images or prediction of the flotation conditions and performance [18, 22–24].

The predominant advantage of convolutional neural networks over traditional fully connected neural networks is that they are specifically designed for image input. This is mainly due to a series of kernel filters employed to extract

✉ A. Jahedsaravani
  A_jahed@khadu.ac.ir

  M. Massinaei
  mmassinaei@birjand.ac.ir

  M. Zarie
  majidzarie@yahoo.com

1 Electrical Engineering Department, Khatam Al-Anbiya University, Tehran, Iran

2 Mining Engineering Department, University of Birjand, P.O. Box 97175-376, Birjand, Iran

fundamental features. The kernels in the first convolutional layer are used to extract the low-level features (edges and lines), while the kernels in the next layers capture the high-level features [25]. The classification performance of convolutional neural networks improves by increasing the number of network layers, but at the expense of more complexity [20].

Another approach is to use convolutional neural networks trained on a different image database (ImageNet) for the extraction of features from flotation froth images [26]. This would be beneficial because the pre-trained model requires less training and less effort to build a model of the process. Although the pre-trained model may be as accurate as or even more accurate than a custom-built convolutional neural network, it saves huge efforts and time required to make the model [15, 21, 24, 25].

To the author's knowledge, to date, a few studies have been published on using deep convolutional neural networks for monitoring and prediction of flotation performance. Horn et al. [27] compared the feature extraction quality of platinum froth flotation images by convolutional neural networks and some traditional feature extraction algorithms. Their findings indicated that the performance of convolutional neural networks was similar to the other feature extraction algorithms. Fu and Aldrich [20] used a pre-trained convolutional neural network (AlexNet) to extract features from froth images. These features were then used in a random forest model to predict the flotation performance and conditions.

In a study by Wang et al. [28], deep convolutional neural network algorithms were used for feature extraction and classification of froth images of a gold–antimony rougher flotation cell to identify the working conditions of the flotation process. Li et al. [29] used a combination of a convolution neural network (as a feature extractor) and a support vector machine (as a predictor) for fault detection and recognition of flotation conditions. Zarie et al. [22] developed a 16-layer deep convolutional neural network for the classification of froth images of an industrial coal flotation column. The convolutional neural network model achieved an accuracy of 93.1% for the classification of froth images in terms of concentrate ash content and combustible recovery, which was much higher than the artificial neural network model.

Zhang et al. [14] built a soft sensor model based on convolutional neural network and memory networks for the prediction of concentrate Zn grade in a zinc rougher flotation bank. Zhang and Gao [15] established a soft sensor based on a hybrid deep neural network for the classification of iron tailings froth images. The proposed algorithm achieved a classification accuracy of 97%. Wen et l. [24] proposed a soft sensor based on convolutional neural networks for the prediction of concentrate ash content of an industrial coal flotation cell. They used various convolutional neural networks for the classification of froth images with different concentrate ash content.

In this study, several convolutional neural networks of five different architectures (AlexNet, GoogLeNet, VGGNet, ResNet, and SqueezeNet) pre-trained on the ImageNet database including over 1.2 million images were used to predict the metallurgical performance of two flotation systems. The first database was the image and metallurgical data obtained from a batch flotation cell operated over a wide range of process conditions. The second database was obtained from an industrial flotation column in a coal preparation plant. The pre-trained networks were used to extract features from the images, and these features were used to construct a model for predicting the metallurgical parameters associated with the froth images.

2 Convolutional Neural Networks (CNNs)

Artificial neural networks (ANNs) are fast, reliable, and robust computational tools for modeling complex non-linear systems like froth flotation. In recent times, ANNs have been used extensively for classification, clustering, pattern recognition, and prediction in many applications [6]. A convolutional neural network (CNN) is a class of feed-forward artificial neural networks that have been successfully applied for image analysis [30]. CNN is a deep learning algorithm that can extract fundamental features from images. Several successive convolutional layers of CNN extract image features and finally learn to classify images. CNNs typically consist of the input layer, convolutional layers, pooling layers, and fully connected layers. In addition to these, rectified linear units (ReLU) can be used for faster training and convergence of the network. The image data are inputted to the first layer. This is followed by a series of convolutional layers. The convolutional layers extract high-level features from the image. The pooling layers are typically used to reduce the dimensions of the feature maps and also to make computation faster because of the reduced number of learning parameters. The extracted features from the previous layers are finally compiled in the fully connected layers to form the final output [20, 21]. A typical architecture of a CNN for the classification of froth images is shown in Fig. 1 [22].

3 Pre-trained Convolutional Neural Networks

The pre-trained CNN algorithms are robust feature extractors, which have been recently used for the analysis of the froth images in the froth flotation systems [20, 21]. The pre-trained CNNs can be used as transfer learning or feature extraction algorithms to speed up the learning
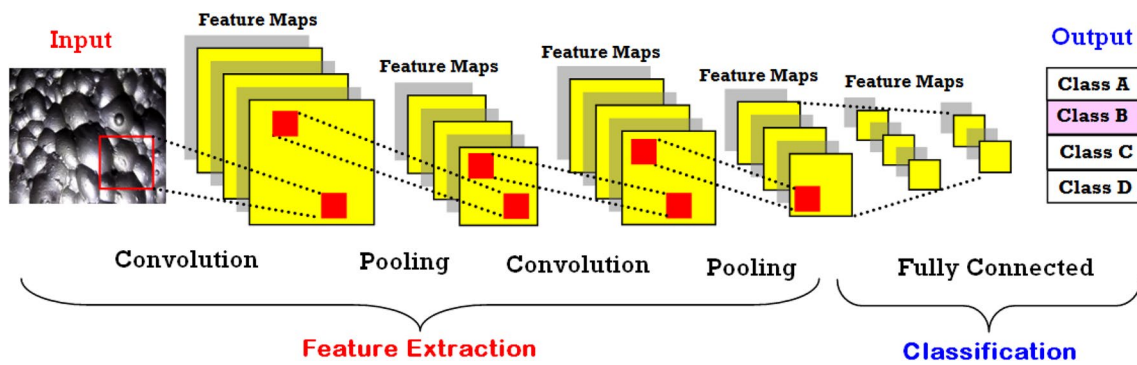
Fig. 1 Typical structure of a CNN for classification of forth images [22]

process and eliminate the need to design the network. The pre-trained CNN model is regarded as the starting point for transfer learning. In this approach, the input image is presented to the network, and high- and low-level image features are extracted by multiple layers of neurons. The first layers extract the low-level features while the last layers extract high-level features. The deep layer's parameters can be further tuned according to the new image database. Fine-tuning of parameters leads to accuracy improvement due to the learning of various sets of features. The principal challenge of this procedure is processing time which is greatly increased in the very large image dataset. Feature extraction is the next strategy benefiting from the potential of pre-trained CNNs. The new database is only passed through the pre-trained CNN, and image features are extracted based on adjusted parameters of the pre-trained CNN. In the froth flotation systems, these features can be subsequently used to predict the process conditions or performance [31].

The feature maps of the pre-trained networks are summarized in Table 1. A brief description of the pre-trained CNN algorithms utilized to extract features from the current froth image database is given in the following sections.

3.1 AlexNet (2012)

AlexNet is a deep CNN architecture designed by Krizhevsky, Sutskever [25]. AlexNet contains eight layers. The first five layers are convolutional layers, some of them followed by max-pooling layers, and the last three are fully connected layers. AlexNet uses the dropout technique to effectively prevent overfitting while training the network [32, 33].
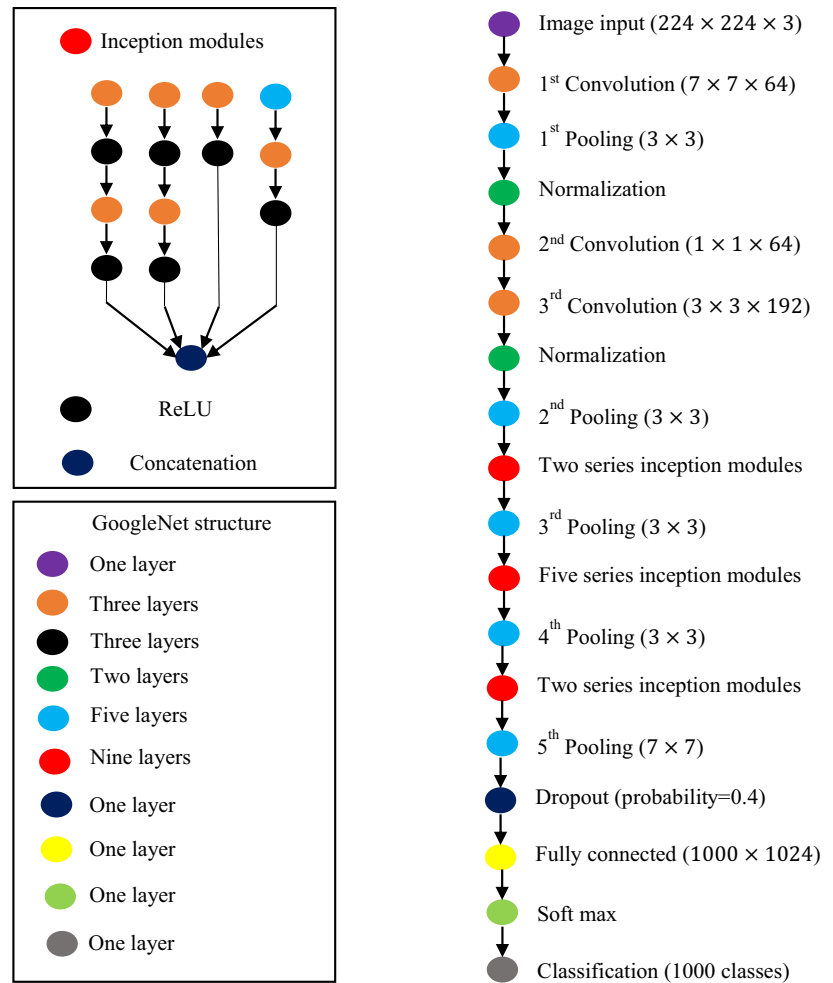
3.2 GoogLeNet (2014)

GoogLeNet is a deep CNN algorithm developed by researchers at Google. This network has 22 layers deep (27 layers, including 5 pooling layers), 9 inception modules, a global average pooling (GAP) layer at the end of the last inception module, and over 7 million parameters. GoogLeNet is a variant of the inception architecture. Convolutional layers with different sizes along with pooling layers are applied to the input in parallel, and then, the resulting output is created by staking convolutional and pooling layers. Thus, different types of features are extracted in an inception module. A dropout layer is also used during training to avoid overfitting the network. GoogLeNet showed a relatively lower error rate compared with AlexNet in the classification of the ImageNet

Table 1 Feature maps of pre-trained CNNs and number of used features in regression models

Pre-trained CNN	Feature map	No. of extracted features	No. of features that explain 95% of total variance	
			$t = 30s$	$t = 120s$
AlexNet	Latest fully connected layer	1000	10	10
GoogLeNet	Global average pooling layer	1024	14	11
VGGNet-16	Latest fully connected layer	1000	11	10
VGGNet-19	Latest fully connected layer	1000	10	10
ResNet-18	Global average pooling layer	512	15	13
ResNet-50	Global average pooling layer	2048	13	14
ResNet-101	Global average pooling layer	2048	12	14
SqueezeNet	Global average pooling layer	1000	6	5

Fig. 2 The architecture of GoogLeNet

* Three convolutional layers of GoogleNet are followed by ReLU layers which are not shown.

database [33]. The architecture of GoogleNet is shown in Fig. 2.

3.3 VGGNet (2014)

VGG16 is a deep CNN model proposed by K. Simonyan and A. Zisserman at the University of Oxford [19]. It is composed of 13 convolutional layers (with 5 max-pooling layers) and 3 fully connected layers. The architecture of VGG16 is similar to AlexNet, but with a few more multiple 3×3 kernel-sized filters. VGG16 is a relatively large network including a total of around 138 million parameters. VGG-19 architecture is the same as VGG-16 except that it has 3 more additional convolutional layers. [34].
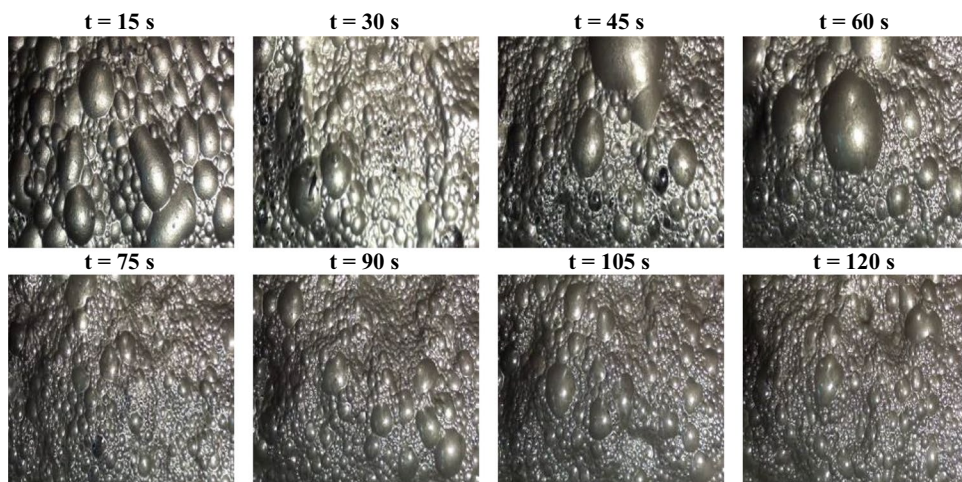
3.4 ResNet (2015)

ResNet is a residual neural network developed by He et al. [35]. Residual networks (ResNets) were introduced based on the key idea of identity connection which can skip one or more layers. ResNets use identity shortcut connections

between layers which turn the network more efficient compared to the widely used feedforward CNNs. Every ResNet block contains two (ResNet-18 and ResNet-34) or 3 layers deep (ResNet-50, ResNet-101, and ResNet-152). The ResNet
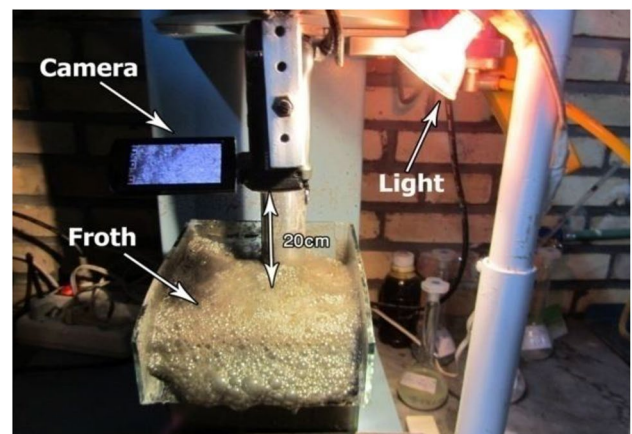
Fig. 3 Laboratory flotation cell equipped with a video camera [6]

Fig. 4 Typical froth images of laboratory flotation cell at different flotation times


model has lower complexity than VGGNet because of fewer filters employed. ResNet-18, ResNet-50, and ResNet-101 architects were used in the current study.

3.5 SqueezeNet (2016)

SqueezeNet is a deep CNN algorithm that was originally introduced by researchers at DeepScale, the University of California, Berkeley, and Stanford University in 2016. The main goal of the development of the SqueezeNet model was to build a smaller neural network with fewer parameters without losing network performance. A SqueezeNet model is composed of convolution layers, fire modules, and pooling layers. The number of connections is significantly reduced with the use of fire modules. An important feature of SqueezeNet is the lack of fully connected layers, which are responsible for extracting high-level features from the input images [36, 37].

Fig. 5 Comparison of pre-trained CNNs algorithms for prediction of concentrate grade (up) and copper recovery (down) at $t = 30$s
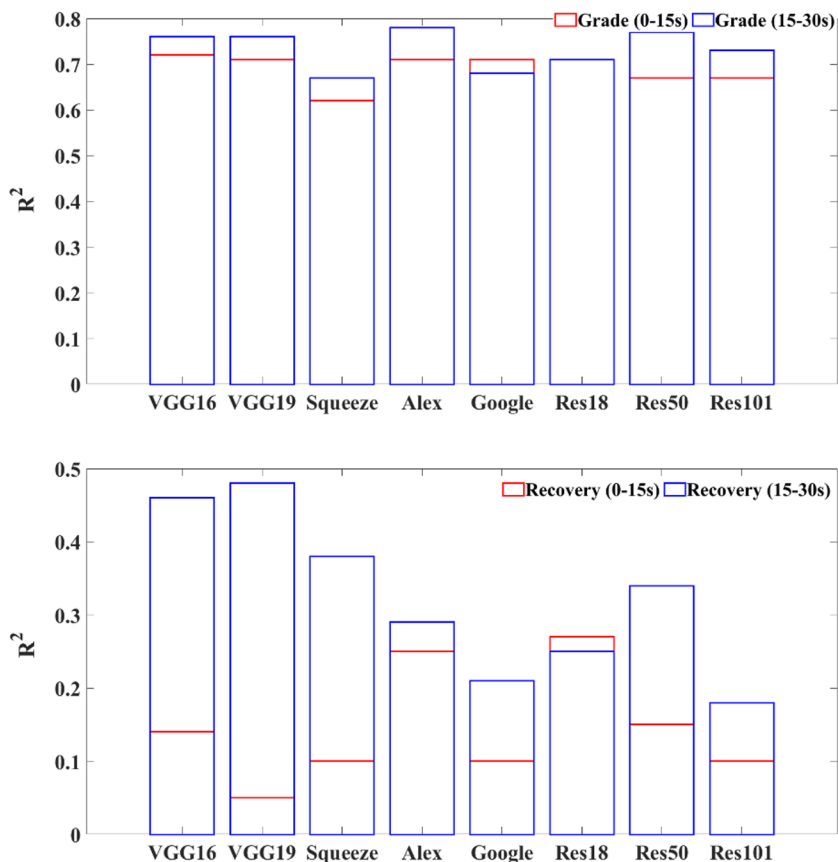
Table 2 Characteristics of developed ANN model

Developed model	No. of samples	No. of neuron					
		$t = 30s$			$t = 120s$		
		Input	Hidden	Output	Input	Hidden	Output
Cu grade	79	11	8	1	10	8	1
Cu recovery		10	7	1	9	9	1

Among the pre-trained neural networks, AlexNet, Goog-LeNet, and SqueezeNet have relatively similar architecture, which can be classified into the same category, while VGG-Net and ResNet have different structures.

4 Case Studies

The first database was the image and metallurgical data obtained from a batch flotation cell operated over a wide range of process conditions [6]. The second database was obtained from an industrial flotation column in a coal preparation plant [18].

4.1 Laboratory Flotation Tests

In the first case study, image and metallurgical data captured from a batch flotation cell treating a copper sulfide ore were considered [6, 38]. The flotation cell was operated at different process conditions, and the froth surface was continuously filmed using a digital video camera (Fig. 3).

A video camera and lighting system (a 50-W halogen lamp) were mounted at a distance of 20 cm above the froth surface. The froth images were captured at a rate of 25 frames per second and 3000 frames per test. The concentrate samples were collected at time intervals of 0.5, 2, and, 5 min. The copper recovery (R) at each time was calculated from the following equation:

$$R = \frac{C.c}{F.f} \tag{1}$$

where f and c are copper content of feed and concentrate samples (%), respectively, and F and C are the mass of feed and concentrate samples (g), respectively.

The metallurgical parameters and the froth images were analyzed in each run. The image and process data collected until 2 min were considered at different runs.

At first, each video recording was split into eight 15-s segments. Typical froth images at different flotation times are presented in Fig. 4. Then, features of the froth images were extracted using the pre-trained CNNs and the mean values of these features were computed for each 15-s segment. The
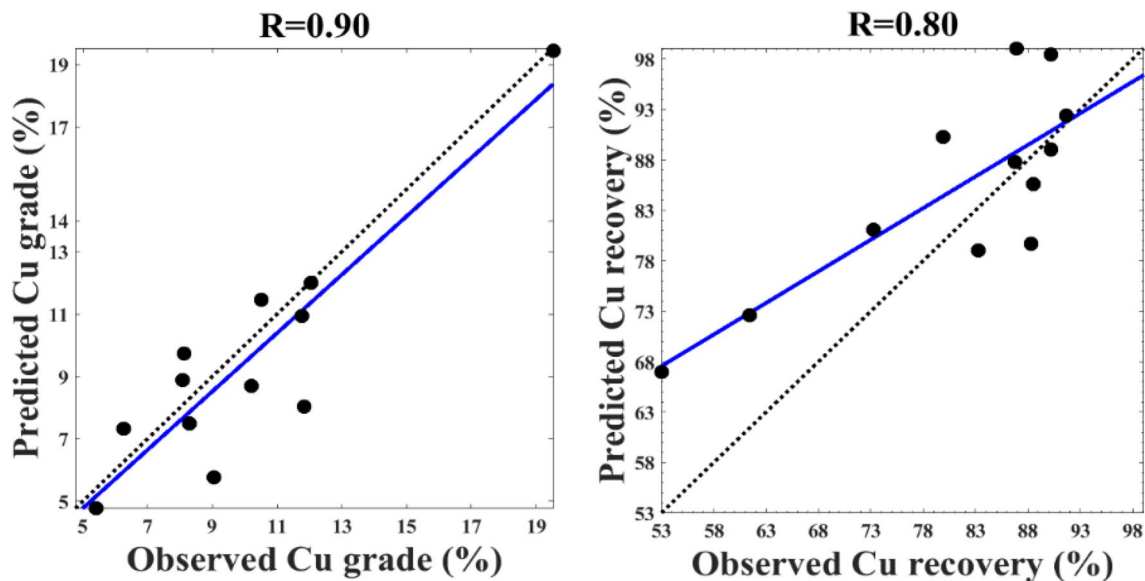
Fig. 6 Scatter plots of observed versus model-predicted values of copper recovery and concentrate grade at $t = 30$ obtained by GoogLeNet-ANN (testing data)

Table 3 Performance of GoogLeNet-ANN algorithms for predicting of the batch flotation performance at $t = 30$

| Metallurgical parameters	R(correlation coefficient)				RMSE			
	Training data	Validation data	Testing data	Total data	Training data	Validation data	Testing data	Total data
G_{Cu}	0.96	0.87	0.90	0.93	1.17	2.21	1.69	1.46
R_{Cu}	0.95	0.67	0.80	0.87	3.08	7.76	8.2	5.1

large number of features extracted from the froth images were reduced by principal component analysis (PCA) [18]. The principal components were subsequently used to build a model for the prediction of the metallurgical parameters (copper grade and recovery). The generalization capability of the regression models was assessed by the leave-one-out cross-validation technique. In this method, n data samples are divided into $n − 1$ training and 1 testing data. Each time, the model is trained by $n − 1$ data sample and then validated by the remaining 1 testing data. The hold-out process is repeated n times to test all the data samples, and then, the error of all testing data is computed and used to evaluate the model [39].

4.1.1 *Prediction of Copper Grade and Recovery at $t = 30$s Using Pre-trained CNNs*

The pre-trained CNN algorithms were applied to extract the froth features and these features were then used to make predictions of the copper recovery and the concentrate grade of the first batch run ($t = 30$ s) of the flotation system. Hence, the froth features were captured for the first two 15-s segments and the mean values were computed. Meanwhile, the copper content of the first-30 s concentrate was determined and the concentrate grade and the copper recovery were measured. Afterwards, two fitted linear
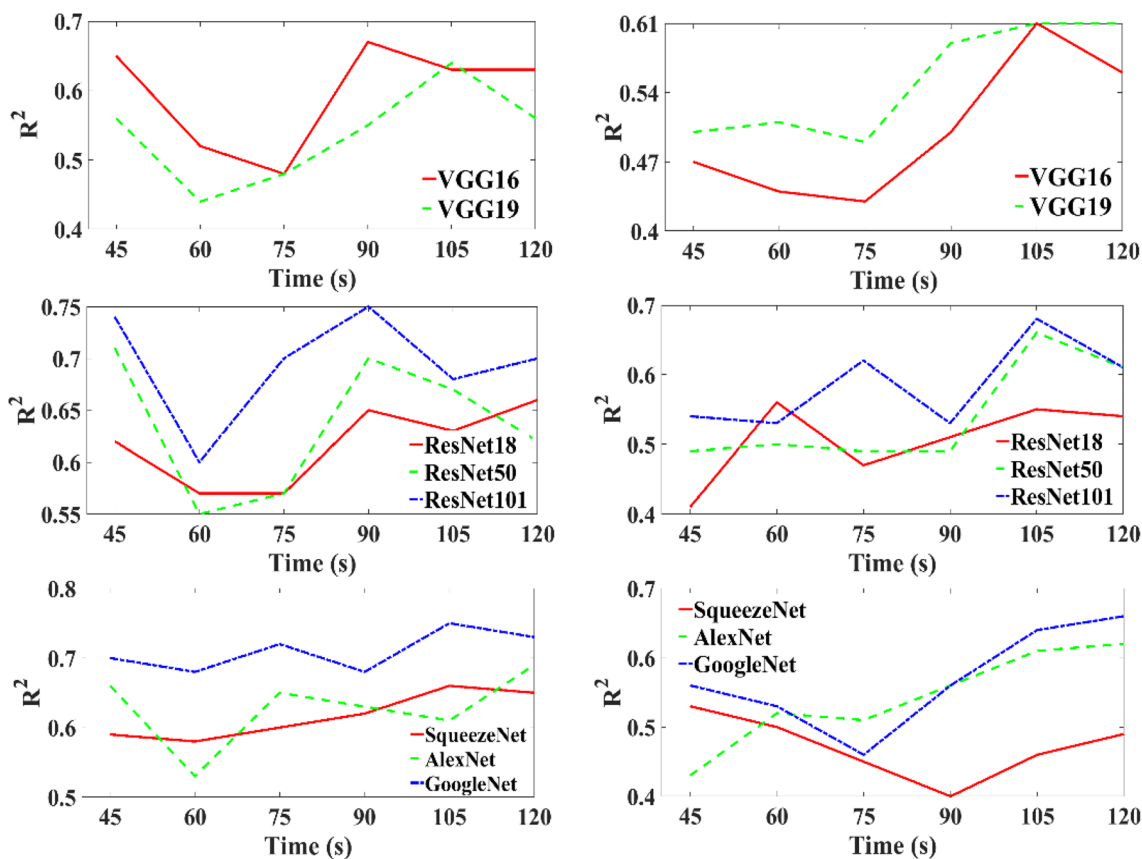
Fig. 7 Comparison of pre-trained CNNs algorithms for prediction of concentrate grade (left curves) and copper recovery (right curves) at $t = 120$s
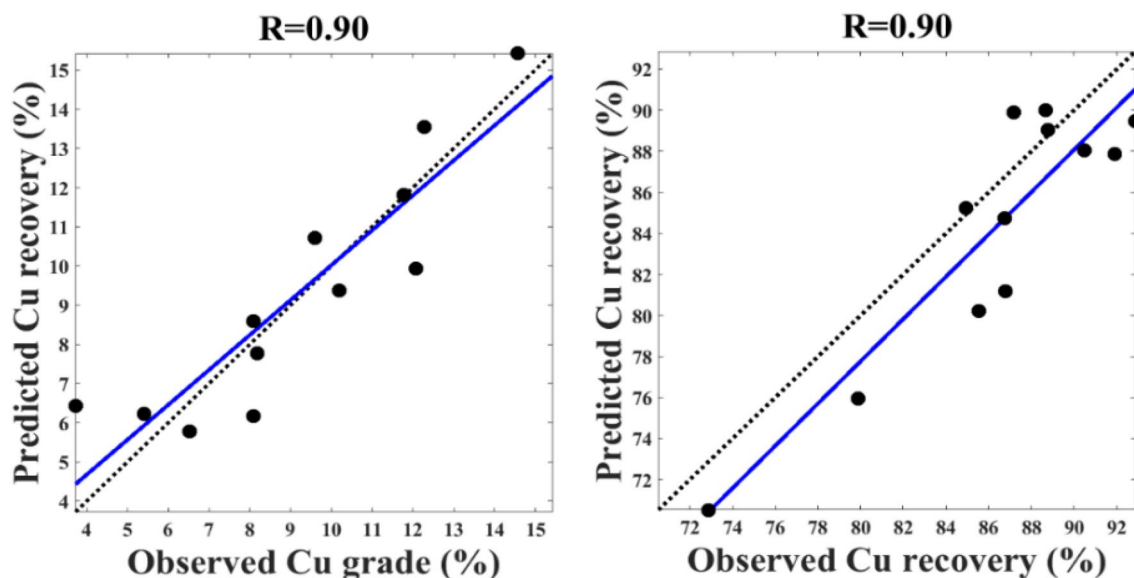
Fig. 8 Scatter plots of observed versus model-predicted values of copper recovery and concentrate grade at $t = 120$ obtained by GoogLeNet-ANN (testing data)

regression models were used to predict the copper recovery and the concentrate grade at $t = 30s$ by using the features extracted from the first two 15-s segments. The goodness-of-fit of the developed models was determined by the coefficients of determination (R^2), which were calculated from the following expression:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \widehat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} \quad (2)$$

where y_i and $\widehat{y}_i$ are the observed and the predicted values of the metallurgical parameters, respectively and $\overline{y}$ is the mean value of the observed values. R^2 values are between zero and unity. $R^2 = 0$ indicates that the regression model does not fit to the data points, and $R^2 = 1$ denotes a perfect fit.

Figure 5 shows R^2 values of the regression models obtained from the extracted features of the first ($t = 0 - 15s$) and the second ($t = 15 - 30s$) batch runs to predict the copper recovery and the concentrate grade at $t = 30s$. The results indicate that the froth visual features extracted at $t = 15 - 30s$ are more correlated with the metallurgical parameters obtained at $t = 30s$. In other words, the prediction accuracy of the algorithms improves as the

image capturing time approaches that of the froth sampling time.

The accuracy of VGG-19 in predicting the copper recovery and the concentrate grade is superior to VGG-16. GoogLeNet significantly outperforms both AlexNet and SqueezeNet models. Among ResNet models, ResNet-101 is the most accurate algorithm for estimating metallurgical performances. Thus, based on the experiments and analysis presented in this paper, it was concluded that the deeper ResNet algorithms with more neural layers exhibit better prediction results compared to the shallower networks. Deeper networks are able to learn more complex and non-linear functions. A large amount of training data involved enables them to more easily predict complicated systems. Compared with all tested algorithms, GoogLeNet achieved the best performance and the highest accuracy on the current dataset.

4.1.2 *Prediction of Copper Grade and Recovery at $t = 30s$ Using GoogLeNet-ANN*

Analysis of the first batch run data indicated that GoogLeNet was the most efficient algorithm for extracting features from

Table 4 Performance of GoogLeNet-ANN algorithms for predicting of the batch flotation performance at $t = 120$

Metallurgical performances	R				$RMSE$			
	Training data	Validation data	Testing data	Total data	Training data	Validation data	Testing data	Total data
G_{Cu}	0.95	0.91	0.90	0.94	1.29	1.56	1.34	1.34
R_{Cu}	0.95	0.90	0.90	0.92	1.97	2.73	3.25	2.33

Table 5 Performance of VFE-ANN algorithms for predicting the batch flotation performance [6]

Metallurgical performances	R				RMSE			
	Training data	Validation data	Testing data	Total data	Training data	Validation data	Testing data	Total data
G_{Cu}	0.97	0.92	0.92	0.96	1.09	1.39	1.07	1.13
R_{Cu}	0.90	0.88	0.9	0.89	2.52	2.9	2.9	2.64

the froth images. Furthermore, the image features extracted from the second 15-s segment showed a greater correlation with the metallurgical parameters. The froth features extracted by the GoogLeNet algorithm at $t = 15 - 30s$ were used to build an ANN model for predicting the flotation performance. The large number of features extracted from the froth images by the GoogLeNet algorithm was reduced by PCA (that explains 90% of total variance). The dataset was randomly partitioned into training (70%), validation (1.5%), and testing (15%) data. The number of neurons in the hidden layers were determined by trial and error. The structure of developed ANN models is listed in Table 2.

The correlation coefficient (R) and the root mean square error (RMSE) were used as predictors to explain the performance of the ANN model in predicting the metallurgical parameters from the image features.

$$R = \frac{\text{cov}(y_i, \widehat{y}_i)}{\sqrt{\text{var}(y_i) \times \text{var}(\widehat{y}_i)}} \tag{3}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2} \tag{4}$$

Figure 6 shows scatter plots of the observed versus model-predicted values of the copper recovery and the concentrate grade obtained by ANN. The solid lines are the best regression lines fitted to the scatter points. The results of the performance evaluation of the developed ANN for predicting the metallurgical parameters are summarized in Table 3. The results demonstrate that the developed GoogLeNet-ANN algorithm can successfully predict the conditions and performance of the batch flotation system from the froth visual features.

4.1.3 *Prediction of Copper Grade and Recovery at $t = 120s$ Using Pre-trained CNNs*

The second concentrate sample was collected at $t = 120s$ after start of the run. The froth features were extracted at $t = 30 - 45s$, $45 - 60$ s, $60 - 75$ s, $75 - 90$ s, $90 - 105$ s, and $105 - 120$ s, and the mean values for each 15-s segment was computed. The froth images were fed to the pre-trained CNNs algorithms, and the features were extracted. The extracted features were used as predictors in a model to estimate the copper recovery and the concentrate grade at $t = 120s$. Six linear regression models were fitted to the extracted features of each 15-s segment to make prediction of the flotation performance at $t = 120s$.

Figure 7 depicts R^2 values of the regression models built with the froth features extracted at different flotation times ($t = 45 - 120s$). The R^2 values exhibit ascending

Fig. 9 Coal flotation column and machine vision system [18]
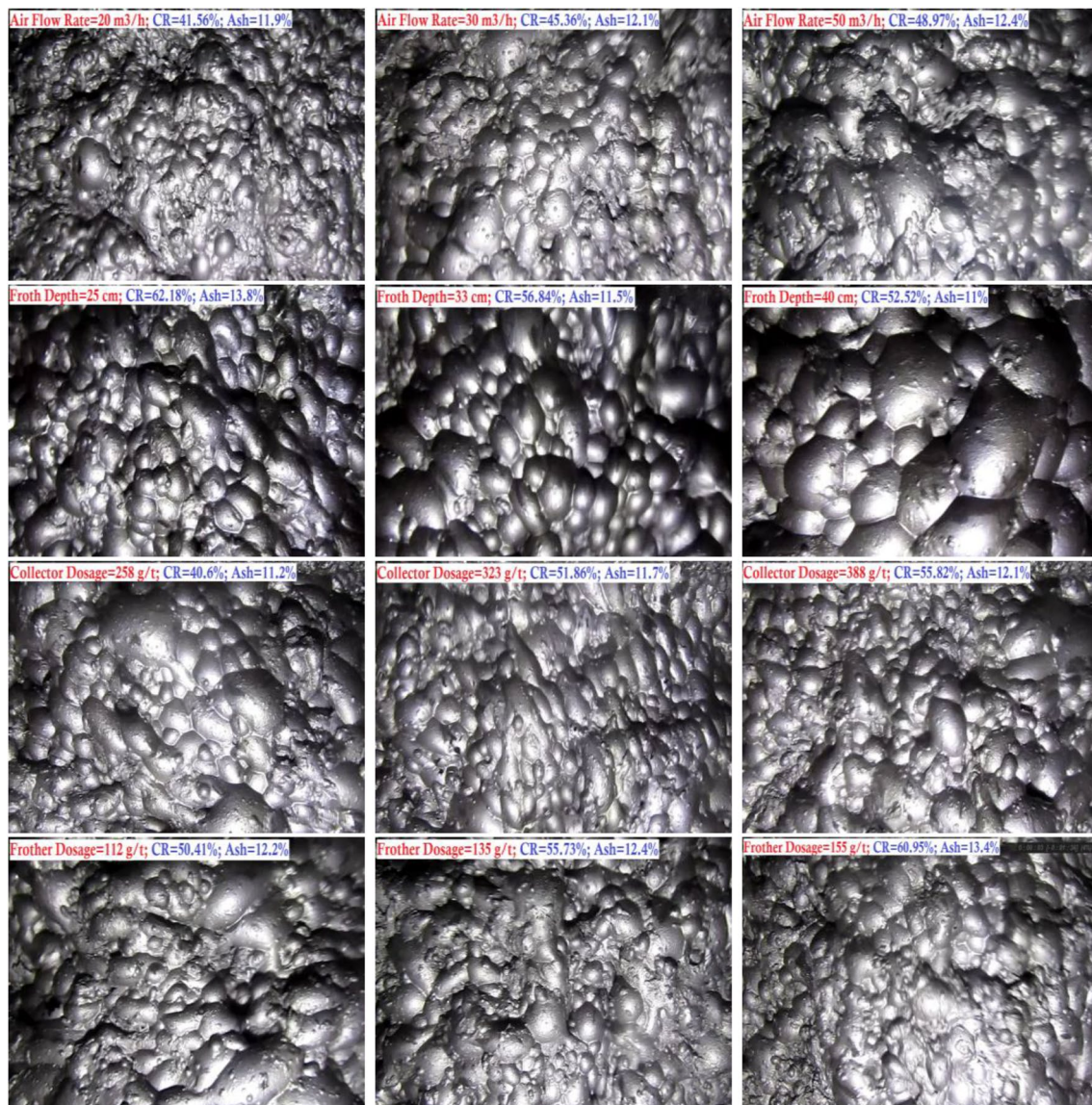
Fig. 10 Typical froth images of industrial flotation columns at different operating conditions

trends over time, indicating again that the last 15-s segment images are more representative of the flotation performance. The results as shown in Fig. 7 indicate that ResNet-110 and GoogLeNet algorithms outperform the other architectures in predicting the copper recovery and the concentrate grade at $t = 120$s. In general, it seems that the deeper networks with more convolutional layers are more successful feature extractors.

4.1.4 *Prediction of Copper Grade and Recovery at $t = 120$s Using GoogLeNet-ANN*

The results presented in the previous section showed that GoogLeNet was the best pre-trained CNN algorithm for extraction

of the visual feature from the froth images. The best results belonged to the froth images captured at $t = 105 - 120$s, which were close to the concentrate sampling time ($t = 120$s). An ANN was trained on the froth features extracted by GoogLeNet algorithm at $t = 105 - 1207$s to estimate the froth grade and recovery at $t = 120$s. The ANN was designed in the same way as the previous one as shown in Table 2.

The predictions of the copper recovery and the concentrate grade by the pre-trained CNN (GoogLeNet) and ANN models are shown in Fig. 8. The performance indicators of the proposed GoogLeNet-ANN algorithm, namely, R and RMSE, are summarized in Table 4. The high R and low RMSE values imply the good ability of the model to accurately predict the flotation performance.

Table 6 Performance of pre-trained CNNs for predicting the coal flotation column performance

Pre-trained CNN	Combustible recovery		Concentrate ash content	
	R^2	RMSE	R^2	RMSE
GoogLeNet	0.64	0.082	0.57	0.061
ResNet-18	0.65	0.081	0.49	0.066
AlexNet	0.60	0.086	0.46	0.067
SqueezeNet	0.52	0.092	0.38	0.071
ResNet-50	0.49	0.094	0.38	0.072
VGG-16	0.45	0.097	0.50	0.065
ResNet-101	0.21	0.11	0.44	0.068
VGG-19	0.18	0.116	0.27	0.077

In the earlier study, the authors developed visual feature extraction (VFE) algorithms for extraction of the froth visual properties including bubble size distribution, froth color, froth velocity, and bubble collapse rate from the froth images captured from this laboratory flotation cell. Afterwards, a three-layer feed-forward artificial neural network was designed to learn the relationship between the froth features and the metallurgical [6]. The results of predictive accuracy of the developed neural network are summarized in Table 5.

The results presented in Tables 3 and 4 suggest that the prediction accuracy of the developed ANN models through visual and GoogLeNet features are relatively similar. Thus, even though GoogLeNet is a convolutional neural network pre-trained on a different image database, it is able to accurately extract features from the froth images. Furthermore, there are some challenges facing the image processing algorithms for analysis of the froth images. (i) The designed algorithms should be tuned and their parameters should be adjusted at each flotation plant. (ii) The image processing algorithms are sensitive to imaging conditions [27]. (iii) The time required for processing of images may cause some delays in the process control system. However, the pre-trained CNN algorithms are less sensitive to changes in imaging conditions and can be applied to extract features from froth images of different flotation systems. These results suggest that the pre-trained CNN algorithms such as GoogLeNet can lead to significant improvements in analysis of froth images.

Table 7 Characteristics of developed ANN models

Developed model	No. of samples	No. of neuron		
		Input	Hidden	Output
Cu grade	20	2	8	1
Cu recovery		2	7	1

4.2 Industrial Flotation Tests

The second case study is based on image and process data collected from an industrial flotation column in flotation circuit of a coal preparation plant in Iran [18]. The camera equipped with a lighting system (a 50-W LED lamp) was installed at a distance of 40 cm above the froth surface (Fig. 9). The camera and lighting systems was placed in a hood to protect them from the ambient light interference. The feed, concentrate, and tailings of the flotation column were sampled simultaneously over a period of 2 h (at 15-min intervals), and their ash content were determined for each test. The combustible recovery (CR) at different experiments was calculated from the following expression [15]:

$$CR = \frac{C(100-c)}{F(100-f)} \times 100 = \left(\frac{t-f}{t-c}\right) \times \left(\frac{100-c}{100-f}\right) \times 100 \tag{5}$$

where f, c, and t are the ash content of feed, concentrate and tailings samples (%), respectively. At the time of sampling, continuous video sampling and analysis of the froth surface was performed at a rate of 30 frames per second. Typical froth images of industrial flotation columns at different operating conditions are shown in Fig. 10.

4.2.1 Prediction of Combustible Recovery and Concentrate Ash Content Using Pre-trained CNNs

The pre-trained networks as deep learning models were used to extract features from the froth images taken from the flotation column at different operating conditions, and the mean values of all the measured features were reported. The froth features extracted by the algorithms were reduced using PCA, which afterward used in a linear regression model to predict the metallurgical performance of the flotation column (combustible recovery and concentrate ash content). The predictive ability of the regression model was quantified by the leave-one-out cross-validation technique.

Table 6 shows the predictive accuracy of the regression models with different froth features (as predictors) extracted by the pre-trained CNNs. GoogleNet achieved the highest accuracy (R^2) and the lowest prediction error (RMSE) values. As a result, GoogleNet is the best pre-trained algorithm for extraction of features from the froth images of the coal flotation column.

4.2.2 Prediction of Combustible Recovery and Concentrate Ash Content Using GoogLeNet-ANN

The previous results revealed that GoogLeNet algorithm can extract the optimal features from the froth images of the coal flotation column. The extracted features by GoogLeNet were subsequently reduced using PCA (just two components) and

Table 8 Comparison of predictive capacity of GoogLeNet-ANN and VTFE-BRANN algorithms for predicting the coal flotation column performance

Algorithm	Combustible recovery		Concentrate ash content	
	R	RMSE	R	RMSE
VTFE-BRANN	0.92	0.045	0.86	0.041
GoogLeNet-ANN	0.90	0.055	0.83	0.047

then used to train an ANN in order to make predictions of the combustible recovery and the concentrate ash content of the coal flotation column at different operating conditions. k-fold cross-validation technique was used to evaluate the proposed ANN model because of the limited data sample. In this method, the dataset is split into k groups. Then, each of the k groups are used as a testing set, while the remaining groups are considered as the training sets. The ANN is trained on the training sets and evaluated on the testing set. This process is repeated for the k groups. Thus, k different

neural network models are trained and validated. In this approach, each sample has the opportunity to be used one time for testing and $k - 1$ times for training. To avoid overfitting, a term including sum of the squared weights ($\sum_{j=1}^{n} w_j^2$) was added to the cost function as follows [18]:

$$J(w) = (1 - \lambda)\frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2 + (\lambda)\frac{1}{n} \sum_{j=1}^{n} w_j^2 \qquad (6)$$

where y_i and $\hat{y}_i$ are the observed and model-predicted values, m is the number of training data, n is the number of weights, and λ is the regularization parameter ($\lambda = 0 - 1$). The structure of developed ANN models is shown in Table 7.

The results of prediction of the flotation column performance by use of a combination of GoogLeNet and ANN algorithms are summarized in Table 8. The results suggest that GoogLeNet is an efficient CNN algorithm that can be used in conjunction with different ANNs for predicting the metallurgical performance of flotation systems.

In the previous study, the authors proposed visual and textural feature extraction (VTFE) algorithms for measuring

Fig. 11 Scatter plots of observed versus model-predicted values of combustible recovery by GoogLeNet-ANN (top figure) and VTFE-BRANN (bottom figure) algorithms
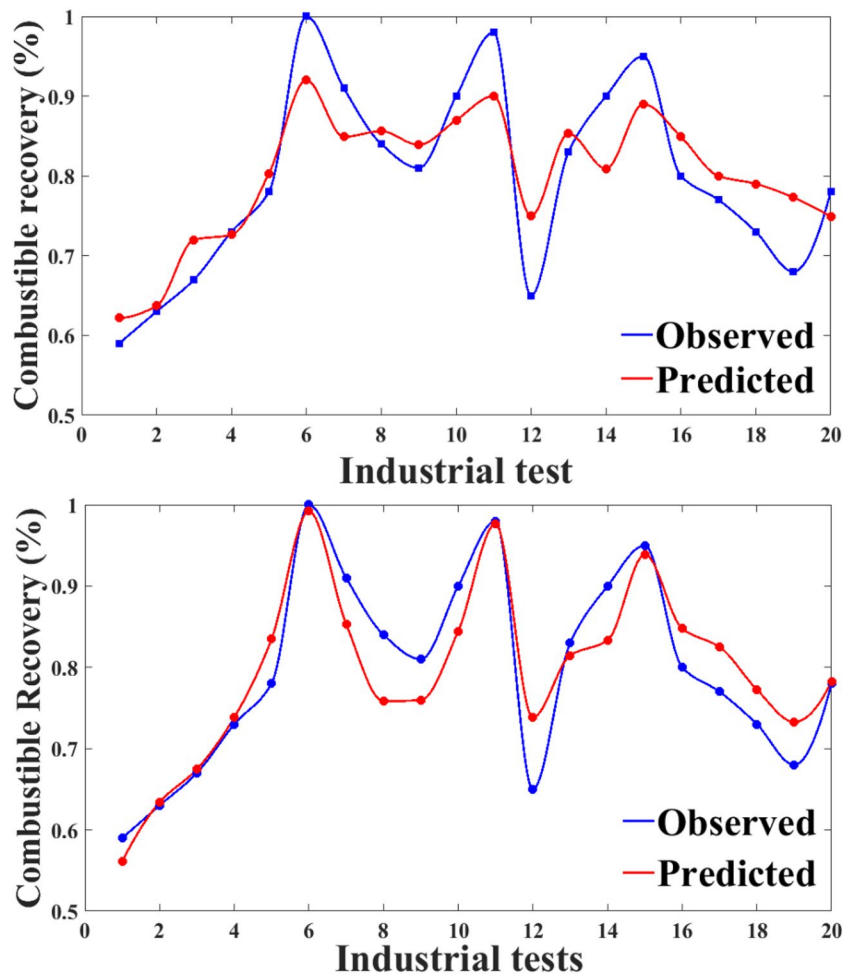
Fig. 12 Scatter plots of observed versus model-predicted values of concentrate ash content by GoogLeNet-ANN (top figure) and VTFE-BRANN (bottom figure) algorithms
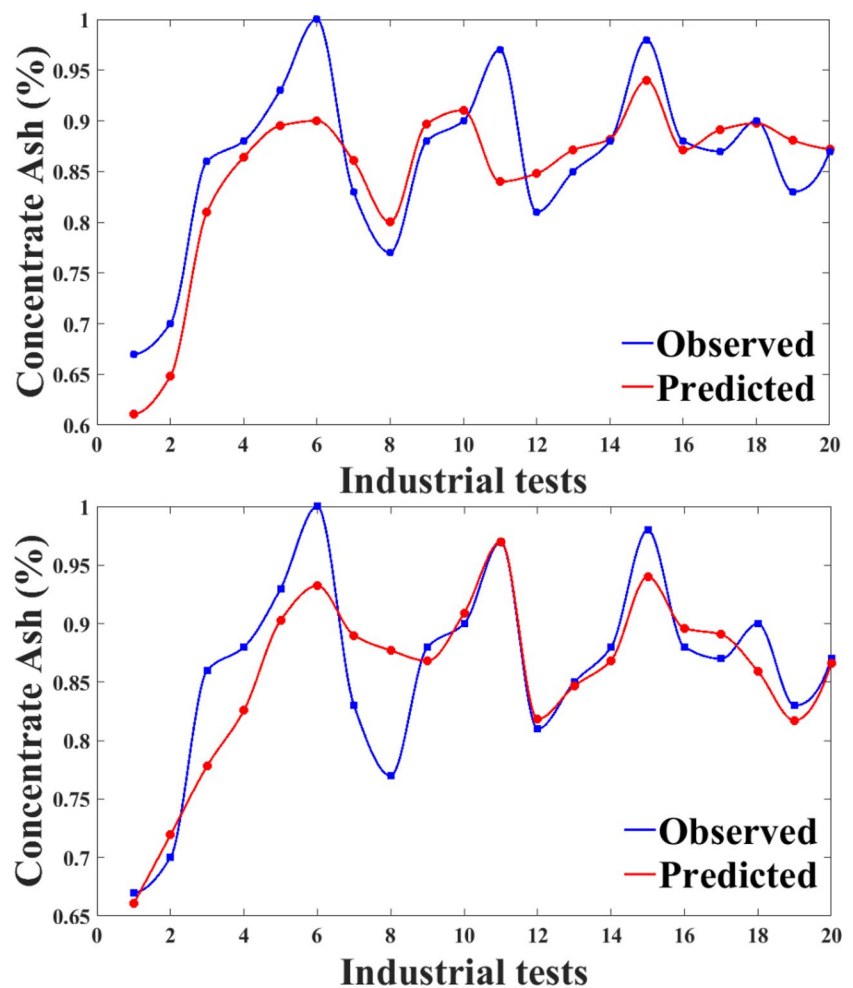
of the froth properties (bubble size, froth velocity, red color, green color, intensity, energy, entropy, contrast, homogeneity, and correlation) of the tested flotation column [18]. The extracted visual and textural features were subsequently used by different intelligent algorithms (BRANNs,[1] ANFIS,[2] and SVM[3]) to estimate the metallurgical performance of the flotation column. BRANN was found to be the best suited model for predicting the combustible recovery and the concentrate ash content. The results presented in Table 6 show that the both VTFE-BRANN and GoogLeNet-ANN algorithms are capable of providing accurate predictions of the flotation column performance at different operating conditions.

Figures 11 and 12 show scatter plots of the observed versus model-predicted values of the combustible recovery and the concentrate ash content obtained by GoogLeNet-ANN and VTFE-BRANN algorithms. The results show that the visual and textural features are more reliable predictors of the flotation conditions than the extracted features by GoogLeNet. It can be concluded that the pre-trained networks such as GoogLeNet can be applied to extract features from froth images and the extracted features may be used to predict working conditions and performance of flotation process.

Transfer learning can be used to improve the performance of pre-trained networks for analysis of industrial datasets. This can be done by tuning the structure of considered pre-trained neural network corresponding to froth image data base. Afterwards, adjusted network can be retrained on a large dataset. In other words, the pre-trained networks can be fine-tuned using the flotation plant datasets to learn higher-level features from the actual froth images [23, 28, 40].

Despite several advantages, deep learning algorithms have some drawbacks. The main practical limitations of deep learning algorithms include a large amount of labeled data required as well as a long training time. The problem of massive data required can be overcome by using active learning, core sets, and data augmentation. The computing time can be reduced by training the network with reduced data and using compact models.

---

[1] Bayesian regularized artificial neural networks.

[2] Adaptive network-based fuzzy inference system.

[3] Support vector machine.

5 Conclusion

In this investigation, a soft sensor based on pre-trained convolutional neural networks was developed for prediction of the metallurgical performance of the flotation systems. The following conclusions can be drawn from this research study.

- The pre-trained networks were effectively used to extract features from the flotation froth images in spite of their different training source (ImageNet). This would be beneficial because the pre-trained algorithms require less training and less effort to build a model of the process.
- The metallurgical parameters of a batch flotation system and an industrial flotation column were accurately predicted from the features generated by the pre-trained structures. The predictions were comparative to the previous results by image processing algorithms developed by the authors for extraction of visual and textural features of the froth images. In addition, pre-trained neural networks are more robust against several challenges which is often occurred in feature extraction by image processing algorithms.
- The difference between the froth sampling and the image capturing rates is the main challenge of froth image analysis in a batch flotation system. The sampling rate of froth images is much higher than the metallurgical data, which leads to several unlabeled froth images. The prediction accuracy of the pre-trained algorithms improves as the image capturing time approaches that of the froth sampling time.
- GoogleNet architecture outperformed all the other pre-trained networks and achieved more accurate estimations of the process conditions and performance on these particular databases.
- The promising results demonstrate that deep learning algorithms have great potential to be used as feature extractors in froth image analysis. Longer computation time because of massive amounts of labeled data is the main demerit of deep learning algorithms for practical applications.

Declarations

Competing Interests The authors declare no competing interests.

References

1. Fuerstenau MC, Jameson GJ, Yoon R-H (2007) Froth flotation: a century of innovation: SME
2. Shean B, Cilliers J (2011) A review of froth flotation control. Int J Miner Process 100(3):57–71
3. Aldrich C, Marais C, Shean B, Cilliers J (2010) Online monitoring and control of froth flotation systems with machine vision: a review. Int J Miner Process 96(1):1–13
4. Massinaei M, Jahedsaravani A, Taheri E, Khalilpour J (2019) Machine vision based monitoring and analysis of a coal column flotation circuit. Powder Technol 343:330–341
5. Hosseini M, Shirazi HHA, Massinaei M, Mehrshad N (2015) Modeling the relationship between froth bubble size and flotation performance using image analysis and neural networks. Chem Eng Commun 202(7):911–919
6. Jahedsaravani A, Marhaban M, Massinaei M (2014) Prediction of the metallurgical performances of a batch flotation system by image analysis and neural networks. Miner Eng 69:137–145
7. Kistner M, Jemwa GT, Aldrich C (2013) Monitoring of mineral processing systems by using textural image analysis. Miner Eng 52:169–177
8. Morar SH, Harris MC, Bradshaw DJ (2012) The use of machine vision to predict flotation performance. Miner Eng 36:31–36
9. Cipriano A, Guarini M, Vidal R, Soto A, Sepúlveda C, Mery D et al (1998) A real time visual sensor for supervision of flotation cells. Miner Eng 11(6):489–499
10. Holtham P, Nguyen K (2002) On-line analysis of froth surface in coal and mineral flotation using JKFrothCam. Int J Miner Process 64(2):163–180
11. Kaartinen J, Hätönen J, Hyötyniemi H, Miettunen J (2006) Machine-vision-based control of zinc flotation—a case study. Control Eng Pract 14(12):1455–1466
12. Mehrabi A, Mehrshad N, Massinaei M (2014) Machine vision based monitoring of an industrial flotation cell in an iron flotation plant. Int J Mineral Process
13. Jian H, Lihui C, Yongfang X. Design of soft sensor for industrial antimony flotation based on deep CNN. Conference Design of soft sensor for industrial antimony flotation based on deep CNN. IEEE, 2492–6
14. Zhang J, Tang Z, Xie Y, Ai M, Gui W (2020) Convolutional memory network-based flotation performance monitoring. Miner Eng 151:106332
15. Zhang D, Gao X (2021) Soft sensor of flotation froth grade classification based on hybrid deep neural network. Int J Prod Res 59(16):4794–4810
16. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117
17. Goodfellow I, Bengio Y, Courville A (2016) Deep learning: MIT press
18. Massinaei M, Jahedsaravani A, Mohseni H (2022) Recognition of process conditions of a coal column flotation circuit using computer vision and machine learning. Int J Coal Prep Util 42(7):2204–2218
19. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556
20. Fu Y, Aldrich C (2018) Froth image analysis by use of transfer learning and convolutional neural networks. Miner Eng 115:68–78
21. Fu Y, Aldrich C (2019) Flotation froth image recognition with convolutional neural networks. Miner Eng 132:183–190
22. Zarie M, Jahedsaravani A, Massinaei M (2020) Flotation froth image classification using convolutional neural networks. Miner Eng 155:106443
23. Liu X, Aldrich C. Monitoring of froth flotation with transfer learning and principal component models. Conference Monitoring of Froth Flotation with Transfer Learning and Principal Component Models. IEEE, p. 110–5.
24. Wen Z, Zhou C, Pan J, Nie T, Zhou C, Lu Z (2021) Deep learning-based ash content prediction of coal flotation concentrate using convolutional neural network. Miner Eng 174:107251
25. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks

26. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. ImageNet: a large-scale hierarchical image database. Conference ImageNet: A large-scale hierarchical image database. Ieee, 248–55

27. Horn Z, Auret L, McCoy J, Aldrich C, Herbst B (2017) Performance of convolutional neural networks for feature extraction in froth flotation sensing. IFAC-PapersOnLine 50(2):13–18

28. Wang X, Song C, Yang C, Xie Y (2018) Process working condition recognition based on the fusion of morphological and pixel set features of froth for froth flotation. Miner Eng 128:17–26

29. Li Z-m, Gui W-h, Zhu J-y (2019) Fault detection in flotation processes based on deep learning and support vector machine. J Cent South Univ 26(9):2504–2515

30. Valueva MV, Nagornov N, Lyakhov PA, Valuev GV, Chervyakov NI (2020) Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. Math Comput Simul 177:232–243

31. López-Saucedo F, Uribe-Salas A, Dávila-Pulido G, Magallanes-Hernández L, Batista Rodriguez J (2017) Effects of solids concentration on gas dispersion in industrial mechanical cells. Miner Metall Process 34(1)

32. Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov RR (2012) Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:12070580

33. Baldi P, Sadowski P (2014) The dropout learning algorithm. Artif Intell 210:78–122

34. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S et al (2015) Imagenet large scale visual recognition challenge. Int J Comput Vision 115(3):211–252

35. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Conf Deep Residual Learn Image Recog 770–8

36. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K (2016) SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. arXiv preprint arXiv:160207360

37. Özyurt F, Sert E, Avcı D (2020) An expert system for brain tumor detection: fuzzy C-means with super resolution and convolutional neural network with extreme learning machine. Med Hypotheses 134:109433

38. Jahedsaravani A, Massinaei M, Marhaban M (2017) Development of a machine vision system for real-time monitoring and control of batch flotation process. Int J Miner Process

39. Molinaro AM, Simon R, Pfeiffer RM (2005) Prediction error estimation: a comparison of resampling methods. Bioinformatics 21(15):3301–3307

40. Giveki D (2021) Improving the performance of convolutional neural networks for image classification. Opt Mem Neural Netw 30(1):51–66