

Received June 13, 2019; reviewed; accepted August 22, 2019

## Column flotation performance prediction: PCA, ANN and image analysis-based approaches

Fardis Nakhaei <sup>1</sup>, Mehdi Irannajad <sup>2</sup>, Sima Mohammadnejad <sup>3</sup>

<sup>1</sup> Department of Mining & Metallurgical Engineering, Amirkabir University of Technology, Tehran

<sup>2</sup> Department of Mining & Metallurgical Engineering, Amirkabir University of Technology, Tehran

<sup>3</sup> Department of Mining Engineering, Tarbiat Modares University, Tehran, Iran

Corresponding author: fardis\_nakhaei@aut.ac.ir

**Abstract:** The flotation froth surface appearance includes remarkable information, which can be employed as a helpful index for the qualitative evaluation of the process efficiency. Image analysis is a practical technology for the sake of achieving process related information that can be employed in expert controllers in order to amend flotation performance. In this paper, the intelligent modelling of relationship between froth characteristics and the metallurgical performance in a pilot column flotation of iron ore was established. Column flotation tests were carried out at a wide range of operating conditions and the froth features along with the metallurgical performances were specified for each run. The artificial intelligence models suggested for the performance parameters prediction include (1) multi-layer back propagation neural network (BPNN), (2) hybrid BPNN with principal component analysis (PCA). The hybrid network was on the basis of the PCA employment in order to decrease the number of variables to be given as input for BPNN. The relationships between the froth features and metallurgical performance factors were successfully modelled via the use of the two methods. The simulation results revealed that the prediction precision of BPNN model on the basis of all the data was relatively higher than the estimation precision of BPNN based on PCA algorithm. The Hybrid BPNN model that was trained by the pre-processed database of measurements achieved from the PCA can be considered a robust method when training time is of paramount importance in objectives of process control.

**Keywords:** image analysis, column flotation, froth features, performance, prediction, PCA, ANN

### 1. Introduction

Though the flotation process has been used for a long time, how to assess the mineral surface properties, the particle characteristics, the production conditions, and the control and modelling of process robustly and automatically are difficult tasks and still academic issues that have not been attended to (Gao et al., 2019; Li and Gao, 2017 and 2018; Shean and Cilliers, 2011). The weakness of adequately precise and trustworthy process measurements strengthens the difficulty. The experienced plant operators have traditionally controlled the flotation circuits through monitoring changes in the froth appearance and make modifications to process variables. However, because of the flotation process complexity, attaining optimal condition is almost impossible for the plant operators (Mehrabi et al., 2014). Furthermore, decision-making based on naked-eye observation of the bubble surface cannot skim all of cells in the flotation circuit timely (Liu et al., 2013). The process performance thus relies on the operator's experience and is confined by the lack of quantitative approaches for froth characterization analysis.

Since the accurate estimation of the exact moment for metallurgical factors (i.e. recovery and grade) changes leads to several issues in an effective economy, the diagnosis of a flotation process general conditions has a key role in the optimisation and control of process (Bergh and Yianatos, 1993).

The real-time prediction and measurement of performance parameters by X-ray analyser generally require sophisticated instruments which are costly to buy and keep and their field of application is commonly limited in terms of mineral size and types to be analysed (Nakhaei et al., 2012a). Machine vision is an automated, non-destructive and cost-effective alternative method to predict of performance process and was successfully employed for monitoring, analyzing and controlling purposes. Previous reports have displayed that the froth optical features involve significant information and show changes in the flotation process status and can be employed to estimate the metallurgical parameters (Bonifazi et al., 2001; Kaartinen et al., 2006; Barbian et al., 2007; Morar et al., 2012a,b). Therefore, techniques on the basis of image analysis have been produced for observation and interpretation of froth images, involving the extraction of bubble size, froth velocity and other characteristics (Holtham and Nguyen, 2002; Aldrich et al., 2010; Farrokhpay, 2011). These studies showed that there is no unique procedure to quantify froth characterization, but several features should be considered at the same time for a more beneficial explanation of froth behaviour.

In this paper, a machine vision system is formulated to assess and model the association between column flotation performance factors and the froth optical characteristics at different operating statuses. Data acquisition is conducted on the column flotation cell so as to desulfurize an iron concentrate. Froth images obtained during column flotation tests are applied for image data intelligent processing aimed at automatic and online metallurgical parameters assessment. Artificial Neural networks (ANN) are designed to create a mapping from froth image input data (froth characterization) to metallurgical performance parameters.

The combination of mathematical techniques provides more useful results than those obtained by the application of single method. In recent years, the techniques based on principal component analysis and artificial neural network (PCA-ANN) have been proposed as prediction tools when data were complex or when nonlinear and interaction interferences were present.

The selection of input data to NN is a critical issue to prevent from over fitting when a plenty of input variables offered. Since the ratio between samples and variables in the ANN should be kept as high as possible, PCA is widely used to reduce the number of variables in a data matrix (data pre-processing). The combined use of PCA and NN can usually improve the training speed (reduces computational overhead) and increase the robustness of the model. The first objective of this study is to illustrate if NN can estimate column flotation performance with high accuracy based on image analysis in the case study carried out for desulphurization of iron ore concentrate and also how the PCA can improve NN model.

In this work, a new approach based on PCA and NN was developed for the simultaneous determination of column flotation performance parameters based on froth characteristics. Performance of PCA-NN was compared with a simple NN to investigate the ability of the PCA in extracting the most useful information from the raw data set. In summary, the suggested artificial intelligence predicting systems for the performance parameters of column flotation comprise two models: (1) a multi-layer feed forward NN (FFNN) and (2) a hybrid FFNN with PCA. Performance of the two NN models were compared by generalization accuracy and convergence speed that are very important indexes for evaluation of models. Such examinations can supply noteworthy contributions towards the progress of control strategies based on real time machine vision system.

## 2. Experimental

### 2.1. Materials

The sample used in this study was taken from magnetic separators concentrate in hematite concentrate plant from Gole-Gohar iron ore complex, Iran. In this plant, the iron ore, after grinding, gravity and magnetic separation, enters flotation cells. As magnetic separators concentrates have a relatively high grade of sulphur (pyrite), the reverse flotation operation is used to reduce the amount of sulfur of the final product (Nakhaei and Irannajad, 2017). The flotation feed sample contained 63.3% Fe, 1.48% FeO and 0.48% sulfur. The particles size distribution of the sample showed that the 84% and 45% of particles were finer than 106 $\mu$ m and 44 $\mu$ m, respectively. The microscopic study and XRD analysis confirmed that hematite, magnetite and goethite were the primary minerals. Calcite, quartz and pyrite were also

commonly seen as gangue minerals. Based on the microscopic liberation results (point count technique), more than 90% of pyrite particles were liberated at  $-105\mu\text{m}$ .

## 2.2. Column flotation

In the present work, the column flotation operation was as the research object to set up soft-sensor model for the metallurgical performance estimation. Reverse flotation was applied for desulfurization of magnetic concentrate, whereby the pyrite was transferred to the froth. Flotation tests were conducted in a pilot flotation column (made of Plexiglas) with a 10 cm internal diameter and 4 m height as shown in Fig. 1. For reverse flotation, potassium amyl xanthate (PAX) was employed as the collector for pyrite. In addition, methyl isobutyl carbinol (MIBC) was employed as a frother in the flotation tests. After conditioning, the pulp was fed to the cell at a flowrate of 2-3 l/min with a certain air flow rate. Air was inserted via an air sparger for bubble generation. The air flowrate was measured by a flowmeter and manually adjusted by a valve. The froth pulp interface position was controlled by regulating the tailing flowrate and it was measured by direct observation and pressure measurements. To make sure that the process had reached steady state operation, the column cell was worked in a closed circuit for a further 15 minutes before sampling of the feed, tailing and concentrate.

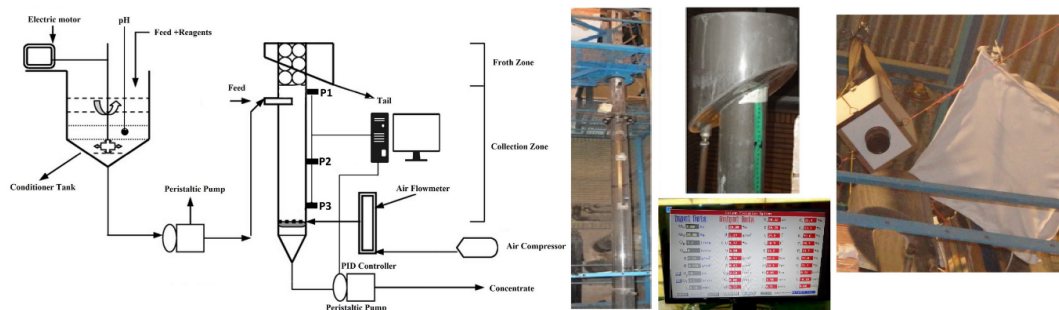


Fig. 1. Column flotation set-up

So as to establish the metallurgical parameters model based on the image analysis, a series of column flotation experiments (60 runs) have been performed in broad range of operation conditions as indicated in Table 1. The concentrate and tail sulphur grades ( $S_c$  and  $S_t$ ), separation efficiency (S.E), froth solid weight ( $S_w$ ), froth water weight ( $W_w$ ) as well as the froth image characteristics were calculated and described for each experiment.

Table 1. Variables ranges of column flotation experiments

Variable	Range
Collector dosage (g/ Mg)	70 - 190
Frother dosage (g/ Mg)	90 - 200
Froth depth (cm)	15 - 35
Gas flow rate (cm/s)	1.4 -1.8
Slurry solids (%)	15 - 35
pH	2.5 - 7

## 2.3. Images acquisition

Image acquisition was conducted using a digital colour camera (connected to a computer) placed in a wood structure with an adjustable arm (as depicted in Fig. 1). The distance from the camera lens to the top of cell was 30 cm. Lighting was supplied by a single 100W LED lamp placed in a soft box nearby the camera (Fig. 1). Also, all around the cell was completely isolated in a way that no external light could be inserted. The image and metallurgical data collections were taken 2.5 min during each set of experiments. Eight frames (with high resolution) for each run were selected and analysed exclusively and the average value of each feature was described for each test. In this study the most significant froth

characteristics consisting of bubble size, colour, texture, bubble collapse rate and froth velocity were extracted in each test.

### 3. Feature extraction of froth image

#### 3.1. Bubble structure

The objective of the structure analysis is to extract and measure morphological parameters, i.e., bubble size and shape, from the froth surface images. It is well recognized that the bubble size in both the froth and pulp phases plays a very important role in flotation efficiency (Aldrich et al., 2010). Various methods presented for bubble size measurement consist of watershed algorithm, texture spectrum and wavelet texture analysis (Mehrshad and Massinaei, 2011; Wang et al., 2003; Lin et al., 2008; Liu et al., 2005). For this work, a watershed algorithm (an intensity-based topographical representation) was used to compute the distribution of bubble size (Lin et al., 2008). Watershed transformation is a noise-sensitive operator so that it tends to over and under-segment the image. In this work, the pre-processing stages (contrast enhancement, noise removal, median and Gaussian filtering) would be applied to enhance quality of image and avoid the over and under-segmentation. After merging the multi-highlights and producing a smooth area for each bubble, the image was segmented. Fig. 2 displays segmentation results of a froth image sample. It can be observed that this approach is able to detect bubbles of different sizes.

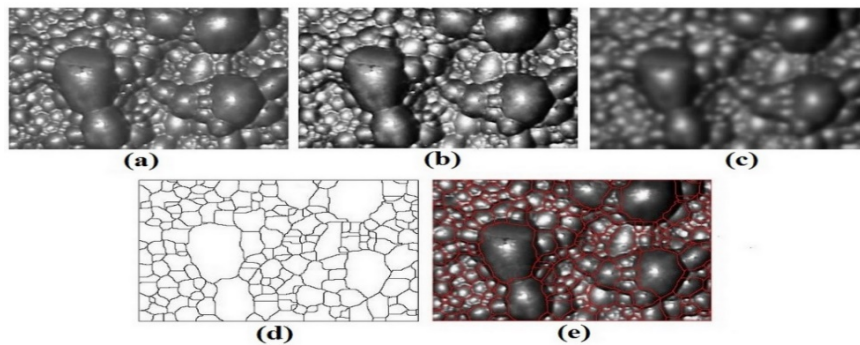


Fig. 2. (a) the initial image (b) contrast enhancement (c) after using the median and Gaussian filters and noise reduction, (d and e) after using Watershed algorithm

Also, after segmentation, morphological features such as roundness (circularity) and aspect ratio (AR) were computed for each froth image and then they were used in modelling. The AR is the ratio of width to length of the minimum enclosing rectangle. The circularity factor explains the closeness of a shape to the circular shape (Raadnu, 2005).

#### 3.2. Froth velocity

Many researchers have presented that the solid and water recovery in flotation process depends on the froth mobility (Cutting, 1986). Froth mobility can be computed by tracking the bubbles movement in successive frames. Block matching (Forbes, 2007), pixel tracing (Holtham and Nguyen, 2002) and bubble tracking (Botha et al., 1999) are the most frequently employed methods to determine the froth speed. In this study, the pixel tracing technique was used to calculate the mean froth speed between the successive frames (Fig. 3). The particular implementation of this technique has been presented by Nakhaei et al., (2018).

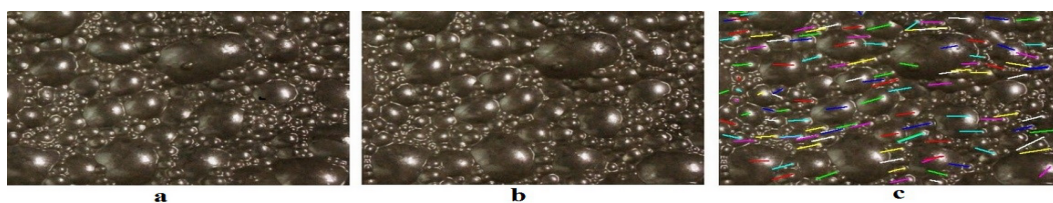


Fig. 3. Froth mobility measurement by the pixel tracing method

### 3.3. Bubble collapse rate

Froth stability is widely identified to play an important role in specifying mineral flotation selectivity and recovery (Ventura-Medina et al., 2003). In this paper, the bubble collapse rate was measured by analysing successive frames and finding the rate of change in the images on the basis of technique which has been presented by Nakhaei et al., (2018). Fig. 4 illustrates this process.

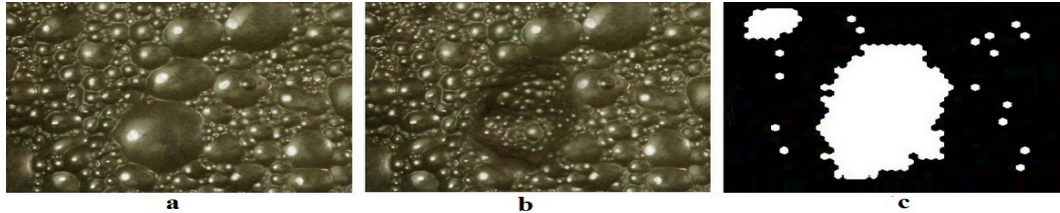


Fig. 4. Froth stability measurement in two successive frames

### 3.4. Froth colour

In this work, the average of the RGB, HSB (hue, saturation and brightness) and Lab values were extracted from images for quantifying the froth colour as model inputs to predict the metallurgical parameters. Both very bright and dark intensity values were omitted from the computation to avoid the effect of shadow and reflectance. Hue is the actual colour ranging from 0 to 360 (red = 0°, green = 120° and blue = 240°). Brightness refers to how much white (or black) is mixed in the colour (100% full saturation and 0% is a shade of grey) while saturation shows the purity of the colour (0% is black, 50% is normal and 100% is white).  $L$  is the luminance factor that varies from 0 to 100 and factors  $a$  (from green to red) and  $b$  (from blue to yellow) are the two chromatic factors that vary from -120 to 120 (Yam and Papadakis, 2004).

### 3.5. Froth image texture feature

It is believed that froth surface texture (roughness) is closely pertinent to the flotation operation condition. Thus, the extraction of texture feature as a useful index for assessment of the flotation performance can provide significant guidance for the optimal operation and control the process. Gray Level Co-occurrence Matrix (GLCM), which is the most widely used method for describing the second order statistical texture features of images, was suggested by Haralick et al. (1979). This approach is on the basis of the identification of the repeated occurrence of grey-level configurations in the texture. More details on GLCM matrix can be found in Haralick et al. (1979) and Pons et al. (2004).

In this paper, based on each GLCM, four of the most typically employed descriptors (the entropy, contrast, inverse different moment (IDM), angular second moment (ASM) or energy) are utilized to elicit textural characteristics of the froth surface greyscale image data set.

Entropy presents a degree of texture disorder or complexity of an image. The more complex texture is, the higher entropy will be, and the other way around. It has a minimum amount of zero once full pixels have the like intensity:

$$Entropy = - \sum_i \sum_j C(i, j) \log C(i, j) \quad (1)$$

Contrast is a measure of intensity or grey level variations between the reference pixel and its neighbour. It defines the texture groove depth degree. Deeper texture groove leads to a greater image contrast and a better optical outcome and vice versa:

$$Contrast = \sum_i \sum_j (i - j)^2 C(i, j) \quad (2)$$

Energy is a scale of the image homogeneity, thus it is a desirable value for finding of disorders in textures. For a heterogeneous image, the matrix will have a large number of small entries off the diagonal, and therefore the energy will be tiny:

$$Energy = \sum_i \sum_j C(i, j)^2 \quad (3)$$

IDM is a straightforward value of the local homogeneity in an image:

$$IDM = \sum_i \sum_j \frac{1}{1+(i-j)^2} C(i, j) . \quad (4)$$

Supposing  $(x, y)$  is a two-dimensional digital image, GLCM means the simultaneous occurrence probability  $C(i, j)$  of two pixels. They are the pixel with gray scale  $i$  from the image  $(x, y)$  and the pixel  $(x + \Delta x, y + \Delta y)$  with gray scale  $j$ , declination  $\theta$  and distance  $\delta$ . The concept of these descriptors can be explained by Bartolacci et al. (2006) and Elmasry and Nakauchi (2016). In this research, these features were acquired by Matlab software. The GLCM features were computed as the mean feature from eight frames of every sampling during different flotation times.

#### 4. Modelling strategy

After implementing the image analysis algorithms into the froth images, data collection campaign was established. In these regards, both image data and process performance parameters were gathered during the different operation conditions of the flotation column. The feature vector resulting from the froth image analysis was a set of 31 parameters established by a first set of colour components (ten parameters), a second set of morphological factors related to bubble size and shape features (three parameters), a third set of components inferred from the GLCM modelling (sixteen parameters) and a fourth set of dynamic parameters (two parameters). On the basis of the correlation analysis results, the following parameters were considered to be an important contributor towards the dependent variable forecast by neural network. The H and b colour features were not used in the parameter set, as their correlations with metallurgical parameters were low and inclusion of these parameters increased the error values of network significantly. In Table 2 the minimum, maximum and standard deviation for each variable used in the modelling procedure have been reported.

Table 2. The Statistical measures of froth features obtained different operating conditions and metallurgical factors used for NN

Input variables							
Features	Maximum	Minimum	<i>St.d</i>	Features	Maximum	Minimum	<i>St.d</i>
Grey level	112	74	9.5	Contrast*	241	45.7	44.6
Red	136	81	15.5	IDM *	0.34	0.14	0.04
Green	111	79.5	7.33	Entropy*	8.75	7.5	0.25
Blue	101	64	8.7	Bubble size	50.8	7.6	11.5
Saturation	0.4	0.19	0.06	Aspect ratio	1.7	1.14	0.2
Brightness	0.53	0.33	0.05	Circularity	0.8	0.36	0.1
L	45.4	33.66	2.9	Velocity	26.4	11.5	3.84
a	44.1	-6.3	5.06	Stability	85	55	6.1
Energy*	1×10 <sup>-3</sup>	2.5×10 <sup>-4</sup>	1.5×10 <sup>-4</sup>				
Output variables							
Metallurgical factor	Index	Maximum	Minimum	Average	<i>St.d</i>		
Solid weight of tail (g)	S <sub>w</sub>	190	16	83.93	39.43		
Water weight of tail (g)	W <sub>w</sub>	380	40	222.3	77		
S grade in conc. (%)	S <sub>c</sub>	0.44	0.08	0.25	0.1		
S grade in tail (%)	S <sub>t</sub>	34.4	13	25.6	5.2		
Separation efficiency (%)	S.E	82.5	6.5	47.3	20.8		

\*Calculated for (0°) and the others (45°, 90° and 135°) were not given in this table

NN is a non-linear computing technique including a large number of interconnected processing units (neurons), which simulates human brain learning. Among different kinds of NNs, the most commonly used one is the multi-layer feed forward NN. This kind of NN constructs a global function approximation and, even if the direct application of a single multi-layer feed forward NN to model a complex system has been proved to be better than conventional methods, there is a need for further improvement of its performance or generalization capability.

The performance of a multi-layer feed forward NN for modeling the column flotation process depends mainly on data representation, accuracy and processing time of the model. Generally speaking, the performance of a network can be improved by reducing the number of input variables correctly,

sometimes even at the cost of losing some useful input information. A useful tool for this is the PCA. PCA transforms the original data set into a set of uncorrelated variables that capture all of the variance of the original data set.

In this paper, the modeling strategy was led by the idea to detect the most effective NN model that will be able to predict metallurgical parameters with an adequate precision, but with a minimal number of predictors required as the input to the model. The aim of this paper was to investigate if the PCA will be effective in reducing training time and preserving or even improving the NN model accuracy for metallurgical performance prediction in a column flotation. Thus, the strategy involved the introduction of two models:

- Model 1: NN with all available input variables (29 variables).
- Model 2: NN with input variables extracted from principal component analysis.

Model 1 is supposed to be the costliest one, owing to the higher dimension of input space, which needs the longer training time. Model 2 could save the time of data collection and the training time.

#### 4.1. Neural network

ANNs are one of the best candidates for experimental data modelling, matching patterns, categorization, forecasting the future on the basis of past data, function approximation and real time data processing (Al-Thyabat, 2009). More details about neural networks as well as tutorials can be found elsewhere (Nakhaei et al., 2012b; 2013 a, b).

In this study, a Multi-Layer Perceptron (MLP) with Levenberg Marquardt back propagation (BP) learning method is employed for training and testing. Working principle of BPNN includes two components: forward propagation and error back propagation. This fast-iterative method is on the basis of gradient descent and it can reduce error function by changing the biases and weight of network in each epoch or iteration (Nakhaei et al., 2013c). A hyperbolic tangent sigmoid transfer function was employed between the hidden and output layers and a linear transfer function was used in the output layer. Training and testing were done using the Matlab software.

It should be pointed out that 75% of data (45 experimental data) was randomly chosen for training and the remaining data (15 experimental data) for testing the model. Input and output data used in the modelling process are shown in Table 2. Determination of the appropriate number of hidden layers and their neurons are very important in architecting NNs and was not determined priori. Research in this area has shown that one or two hidden layers with an adequate number of neurons are sufficient to model any solution surface of practical interest. An excessive number of hidden layer nodes have large number of associated undetermined parameters that leads to poor generalisation properties, such that the network tends to memorize rather than generalize and is unable to interpolate effectively between adjacent training data points. On the other hand, too few hidden neurons limit the competence of the net in locating adequate mapping between response and factors. In the last stage of the simulation, the hidden layer neurons number was specified by training different NNs with various number of hidden layer neurons and different transfer functions. The network doing best on a testing data collection was chosen as the most capable network architecture.

### 5. Results and discussion

#### 5.1. Modelling relationship between froth characteristics and metallurgical factors by NN

The MLP was trained by the Levenberg-Marquardt back propagation optimization algorithm using 140 iterations (epochs). In this study, number of layers and their neurons were determined by trial and error procedure. Several structures of the NN with the number of neurons varying from 5 to 20 were trained for different learning rates and momentum coefficient values. The best architecture of the network was determined by calculating the root mean squares error (RMSE) for all ANN models. The architecture 29-16-13-5 with momentum coefficient of 0.8 and learning rate of 0.2 gave the lowest estimation error among all the structures studied in this analysis. The architecture of this kind of NN is depicted in Fig. 5.

Fig. 6 shows the correlation between output and target values of the metallurgical performances predicted by the proposed NN for the testing data set. It appears that the NN model has estimated

values close to the measured ones. The correlation coefficient values for testing sets were 0.94, 0.93, 0.94, 0.94 and 0.93 in  $S_w$ ,  $W_w$ ,  $S_c$ ,  $S_t$  and S.E predictions, respectively. Since column flotation is a strong nonlinear process with changeable parameters, the correlation coefficient between estimated and measured values is very high.

Fig. 7 shows the comparison between actual and output values of the metallurgical performances predicted by the proposed NN for the testing data set. The results display that the NN can simulate the complex linking between the input and target vectors with high correlation coefficient and low error; thus, it can be successfully applied for real-time estimation of column flotation metallurgical parameters using froth images.

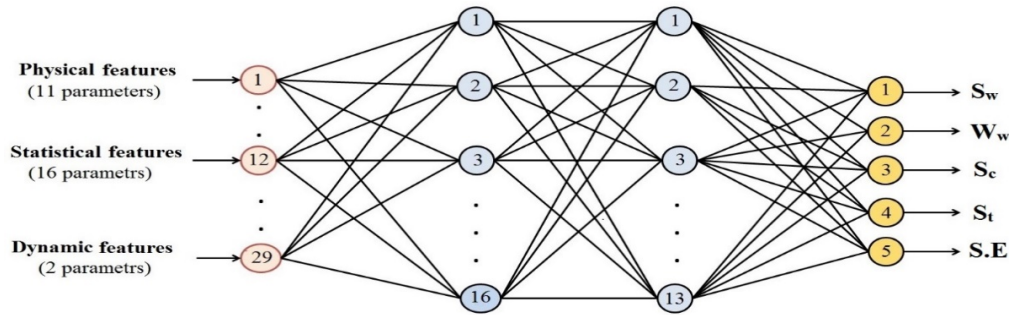


Fig. 5. NN architecture used in the simulation

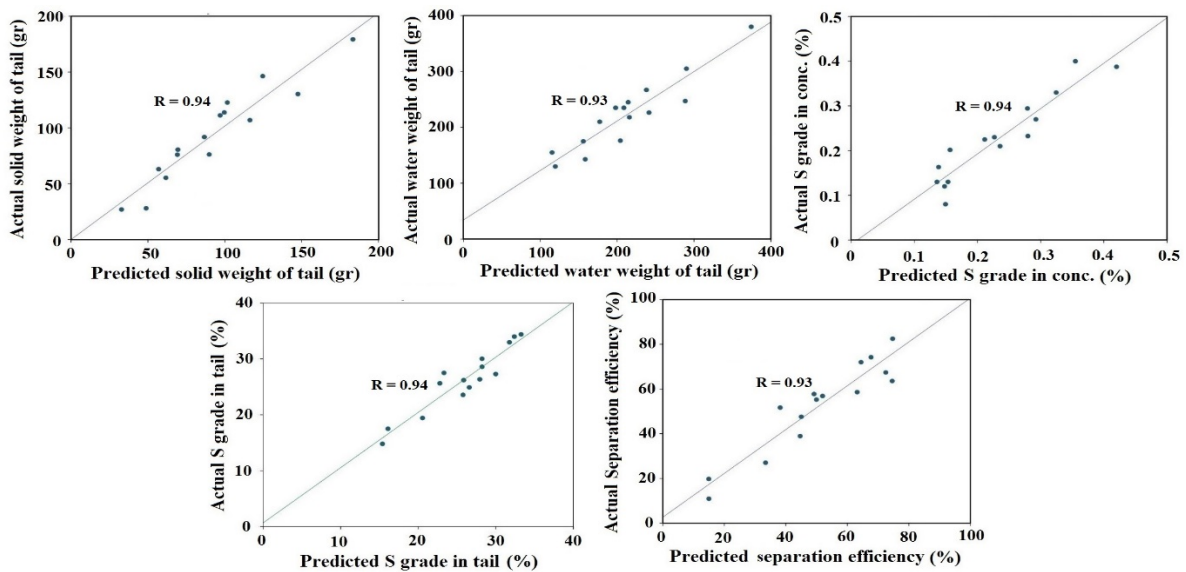


Fig. 6. Comparison of predicted results from the NN model with target values for the testing data set

## 5.2. Modelling relationship between froth characteristics and metallurgical factors by PCANN

For the proposed BPNN model, the input vector dimension ( $60 \times 29$ ) was too large, which makes the network geometry and training complicated. So, to simplify the model scale and decrease the NN processing time and to have very good performance, the best approach is to reduce dimensions of network variables. In this paper, Principal Component Analysis (PCA) technique was applied to reduce the dimensionality of high dimensional input vectors. PCA is used as a data reduction method to distinguish a small number of factors that describe most of the variance observed in a large number of manifest variables. In the image analysis area, PCA has become the most usual technique to decrease the dimension of the information sets by keeping the pertinent data and understanding the relationship among the variables employed to build a model. PCs are new uncorrelated and approximately normally distributed variables that supply reliable representations of the image, which can be employed later as input data for prediction, clustering and other targets (Kara and Direngali, 2007). The order of the PC's



relies on their degree of significance, like the first PC presents the most important relation with respect to the initial variables. The PCA method has been explained by other researchers (Prats-Montalbán et al., 2011; Balas et al., 2010).

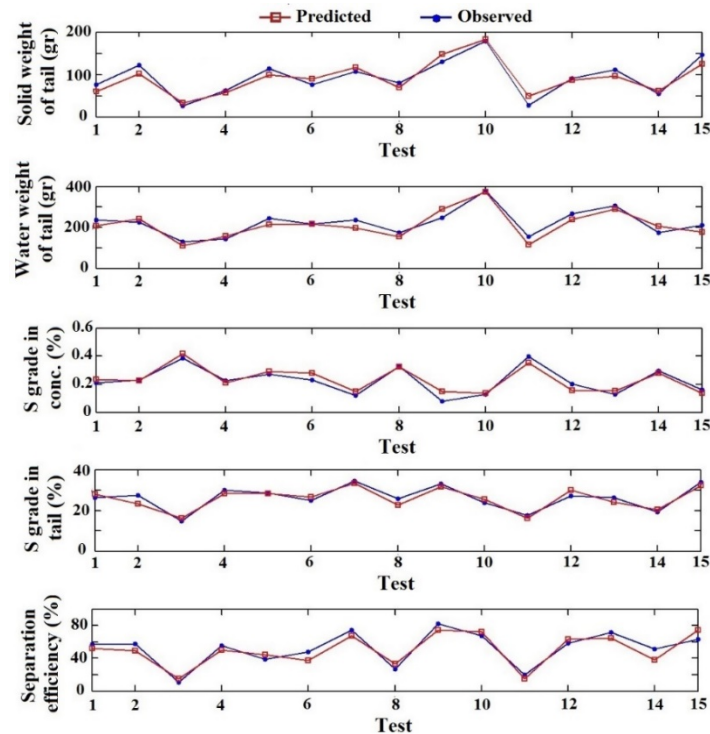


Fig. 7. Comparison of the predictive output and the actual output of the NN model for testing data set

In this work, PCA was employed for data pre-processing and for simplifying the NN structure, since the estimation capability of the NN is more dependent on data representation than on the choice of a learning rule. PCA was conducted by using SPSS software version 21. At first, the Bartlett's test of sphericity has been done on the correlation matrix, and the null hypothesis of correlation absence between the variables shows the appropriate use of PCA on this data. Kaiser criteria for determining the number of components were used in this study, based on which the PCs with eigenvalues bigger than 1 are kept for further analysis (Kaiser, 1958). Therefore, 4 was chosen as the optimum number of PCs. After extraction the PCs, the varimax rotation was done in order to get factors more linearly independent. The amount of total variance contributed by each component and the percentage of cumulative variance are given in Table 3. It can be realized that the first four PCs explained about 86 percent (%) of the variance in the data set.

Table 3. Eigenvalues and percentage of variance with and without varimax.

Component	Eigenvalues without varimax rotation			Eigenvalues with varimax rotation		
	Total	% of variance	Cumulative	Total	% of variance	Cumulative
1	19	63.8	63.8	10.86	37.45	37.45
2	3.67	12.68	76.48	6.69	23.1	60.54
3	1.5	5.24	81.73	4.14	14.27	74.81
4	1.2	4.17	85.89	3.21	11.08	85.89

The first PC showed a strong relation with characteristics of the texture; the second PC described the colour features and stability of the froth; the third PC had a relation with the movement of the froth; ultimately, the fourth PC showed a strong relation with the morphology of the froth.

After the implementation of PCA pre-processing, the size of the data set applied in the BPNN model (29 image variables extracted) was systematically reduced to four principal components. Then, this PCs

data set was divided into two subsets: training (75%) and testing (25%) sets. The topology of the proposed methodology for PCANN model is illustrated in Fig. 8. This figure shows a network with an input layer of four neurons representing the PCs, two hidden layers with 12 and 10 neurons, respectively, and finally the output layer with 5 neurons. The PCANN was trained using 60 epochs. The reduction of features reduces the number of epochs.

The correlation coefficient values in testing stage are shown in Fig. 9. The correlation coefficient values for testing sets were 0.91, 0.89, 0.91, 0.92 and 0.92 for  $S_w$ ,  $W_w$ ,  $S_c$ ,  $S_t$  and S.E predictions, respectively. Comparisons of predicted values by the PCANN model at the testing stage and the database of observed values are presented in Fig. 10. The testing set, which actually tests how good the model is, shows that the model could approximate the metallurgical parameters quite satisfactorily.

It can be observed that a well agreement was made between the predicted values of hybrid NN and real data set. The computed results display that PCA pre-process procedure to input data can also be used for prediction with much simpler network architecture. In other words, nearly the same accuracy of NN prediction can be gained by much fewer input data selected by PCA.

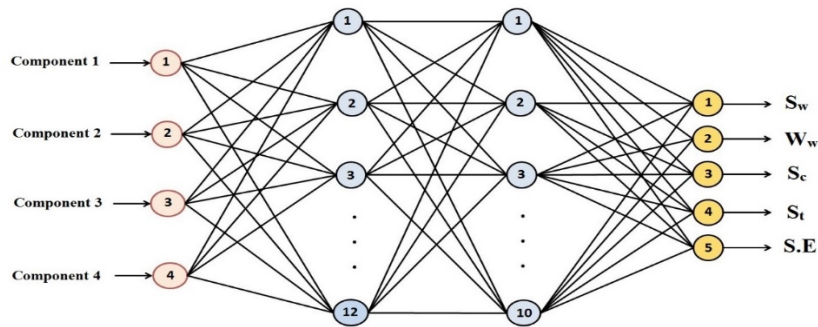


Fig. 8. PCANN architecture used in the simulation

Performance of the two NN models were compared by generalization accuracy and convergence speed that are very important indexes for evaluation of models. To display the accuracy of the proposed methods, root mean square error and other performance indexes of each model were calculated (Table 4). The statistical parameters showed that both the ANN models were reliable, but simple ANN demonstrated a bit better performance in the prediction of the validation set samples, showing lower residual errors.

On the other hand, convergence speed for PCA-NN (245 order) was shorter than NN (771 order), as

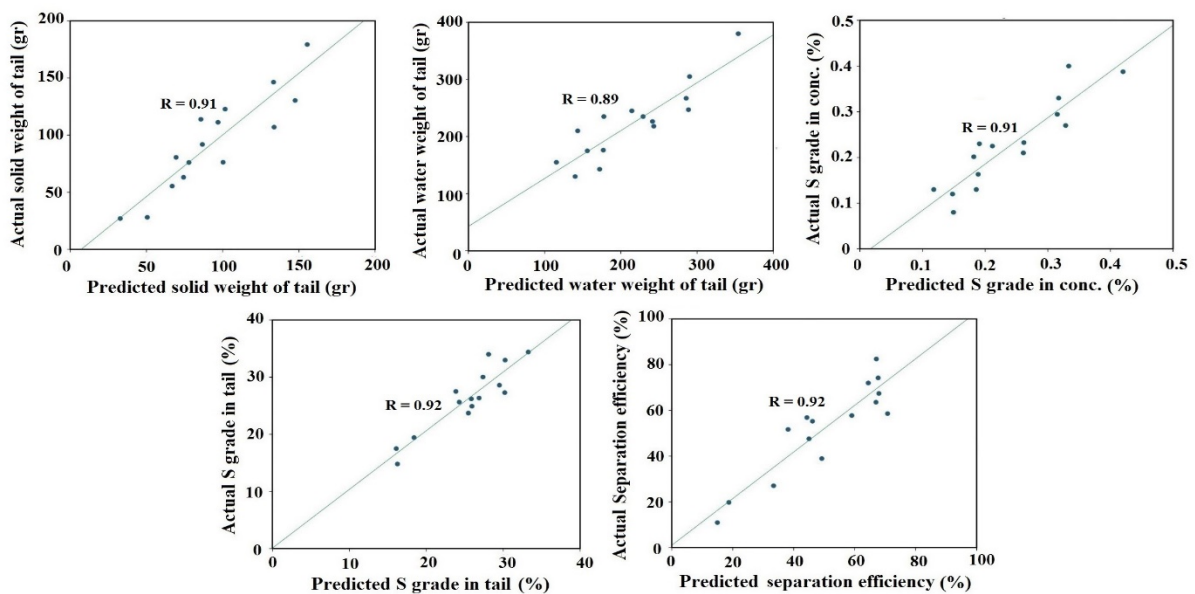


Fig. 9. Comparison of predicted results from the PCANN model with target values for the testing data set

verified in the model order and reducing the NN architecture that help effectively when training time is of paramount importance in objectives of process control. When the accuracy of model is not noticeably important, the time of data collection and training process should be regarded as standard for model selection. This study shows a potential use of new artificial intelligence techniques such as the hybrid NN and the PCA in engineering problems.

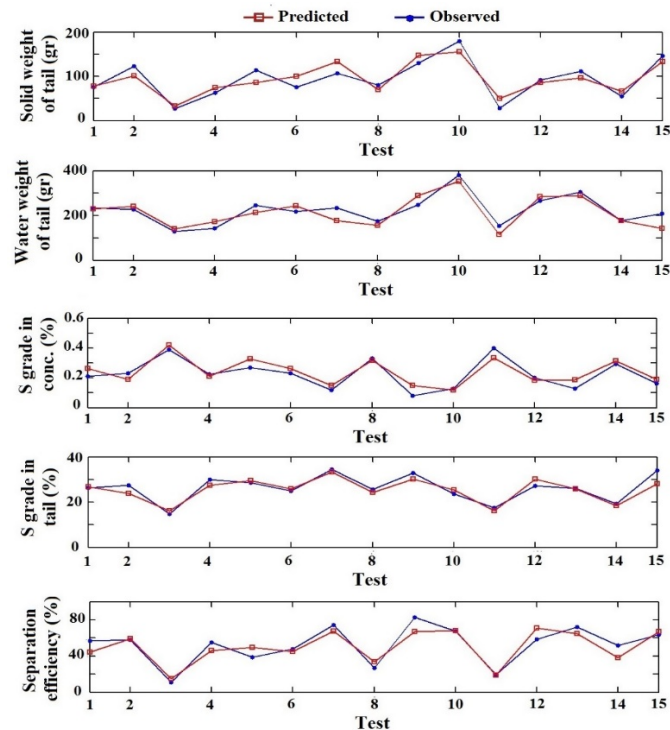


Fig. 10. Comparison of the predictive output and the actual output of the PCANN model for testing data set

Table 4. Performance of ANN and PCANN for test data

Predicted variable	Neural network (Order = 771)				Hybrid neural network (Order =245)			
	RMSE	MAX	MIN	Average	RMSE	MAX	MIN	Average
SW (gr)	13.76	20.74	-21.6	-2.13	17.72	26.8	-28.1	0.3
WW (gr)	27.03	41.26	-39.38	-10.44	32.09	41.65	-66.62	-7.9
SC (%)	0.03	0.07	-0.045	-0.007	0.04	0.07	-0.07	0.014
ST (%)	1.99	2.69	-4.18	-0.42	2.29	2.91	-5.9	-0.77
SE (%)	7.54	11	-13.5	-2.15	8.52	12.37	-15.4	-2

## 6. Conclusions

In this research, artificial intelligence models were successfully developed for the prediction of column flotation metallurgical parameters (solid and water weight of tailing, sulphur grades of concentrate and tail and separation efficiency) based on froth image characteristics. Bubble shape and size, texture features, froth colour, froth velocity and stability were elicited from the froth images at different operation conditions.

Two neural network models were applied to estimate the metallurgical parameters of the column flotation process using a wide range of experimental conditions. The first model was on the basis of all experimental data and the second model was on the basis of data simulated by applying the input optimization strategy according to principal component analysis (PCA). The aim was to lower the training time of the NN, which means to find the minimal number of items needed to estimate the output factor. The structure of the NN model was simplified using the corresponding important PCs as input variables instead of the original data so that model order decreased from 771 to 245.

The statistical parameters showed that both the ANN models were reliable so that the correlation coefficients between predicted and observed values in testing stages in two models were above 0.91, which was very satisfactory. This fact shows that NNs are long-range tools for studying and performance estimation of flotation process. NN model based on all data (without any pre-process procedure) demonstrated a bit better performance in the estimation of the metallurgical parameters and has higher precision and the error curve changes smoothly, which shows better generalization capability and robustness. However, convergence speed for PCA-NN was shorter than NN, as verified in the training stages. The suggested models can be employed to simulate and find the optimal operation conditions for the column flotation process. These procedures can also be generalized to industrial operations providing the progress of more suitable process control strategies because a perfectly trained NN can estimate target variables very expeditiously. The significance of such probes is that an important contribution is made towards the improvement of a control system on the basis of machine vision for industrial applications. In conclusion, our work may assist the following research of the overall coordinated control (froth height, reagent addition, etc.) and the total process optimization. As the majority of flotation processes exhibit non-linear characteristics, there will be a need to further research that would extend the potential of these models' application in real industrial conditions.

### Acknowledgments

The authors would like to acknowledge the support of the Gole-Gohar Iron Ore Complex and Steel Research Institute for this research.

### References

- ALDRICH, C., MARAIS, C., SHEAN, B. J., CILLIERS, J. J., 2010. *Online monitoring and control of froth flotation systems with machine vision: a review*. International Journal of Mineral Processing 96, 1-13.
- AL-THYABAT, S., 2009. *Investigating the effect of some operating parameters on phosphate flotation kinetics by neural network*, Advanced Powder Technology 20, 355-360
- BALAS, C. E., KOC, M. L., TUR, R., 2010. *Artificial neural networks based on principal component analysis, fuzzy systems and fuzzy neural networks for preliminary design of rubble mound breakwaters*, Applied Ocean Research, 32, 425-433.
- BARBIAN, N., CILLIERS, J. J., MORAR, S. H., BRADSHAW, D. J., 2007. *Froth imaging, air recovery and bubble loading to describe flotation bank performance*. International Journal of Mineral Processing 84 (1-4), 81-88.
- BARTOLACCI, G., PELLETIER, P., TESSIER, J., DUCHESNE, C., BOSSE, P. A., FOURNIER, J., 2006. *Application of numerical image analysis to process diagnosis and physical parameter measurement in mineral processes – part I: flotation control based on froth textural characteristics*. Minerals Engineering 19 (6-8), 734-747.
- BERGH, L. and YIANATOS, J., 1993. *Control Alternatives for Flotation Columns*. Minerals Engineering, 6(6), 631-642
- BONIFAZI, G., SERRANTI, S., VOLPE, F., ZUCO, R., 2001. *Characterisation of flotation froth colour and structure by machine vision*, Computers & Geosciences, 27, 1111-1117.
- BOTHA, C.P., WEBER, D.M., VAN OLST, M., MOOLMAN, D.W., 1999. *A practical system for real time on-plant flotation froth visual parameter extraction*. Proceedings of the 5th IEEE Africon, Cape Town, South Africa, 1, 103-106.
- CUTTING, G.W., BARBER, S.P., NEWTON, S., 1986. *Effects of froth structure and mobility on the performance and simulation of continuously operated flotation cells*. International Journal of Mineral Processing, 43-61.
- ELMASRY, G., NAKAUCHI, S., 2016. *Image analysis operations applied to hyper spectral images for non-invasive sensing of food quality: A comprehensive review*, Bio systems engineering, 142, 53-82.
- FARROKHPAY, S. 2011. *The significance of froth stability in mineral flotation – A review*. Adv. Colloid Interface Sci. 166, 1-7.
- FORBES, G., 2007. *Texture and bubble size measurements for modelling concentrate grade in flotation froth systems*, PhD Thesis University of Cape Town, South Africa.
- GAO, Z., FAN, R., RALSTON, J., SUN, W., HU, Y., 2019. *Surface broken bonds An efficient way to assess the surface behaviour of fluorite*. Miner. Eng. 130, 15-23.
- HARALICK, R. M. 1979. *Statistical and structural approaches to texture*, Proceedings of the IEEE 67 (5), 780-803.
- HOLTHAM, P. N., NGUYEN, L. K., 2002. *On-line analysis of froth surface in coal and mineral flotation using JK FrothCam*. Int. J. Miner. Process. 64, 163-180.

- KAARTINEN, J., HATONEN, J., HYOTYNIEMI, H., MIETTUNEN, J., 2006. *Machine vision based control of zinc flotation – a case study*. *Control. Eng. Pract.* 14, 1455–1466.
- KAISER, H. F. 1958. *The varimax criterion for analytic rotation in factor analysis*. *Psycho metrika*, 23, 187–200.
- KARA, S. and DIRENGALI, F., 2007. *A system to diagnose atherosclerosis via wavelet transforms, principal component analysis and artificial neural networks*, *Expert Systems with Applications*, 32, 632–640.
- LI, C., GAO, Z., 2018. *Tune surface physicochemical property of fluorite particles by regulating the exposure degree of crystal surfaces*. *Miner. Eng.* 128, 123–132.
- LI, C., GAO Z., 2017. *Effect of grinding media on the surface property and flotation behavior of scheelite particles*. *Powder Technol.* 322, 386–392.
- LIN, B., RECKE, B., KNUDSEN, J. K. H., JORGENSEN, S. B., 2008. *Bubble size estimation for flotation processes*. *Miner. Eng.* 21, 539–548.
- LIU, J. J., MACGREGOR, J. F., DUCHESNE, C., BARTOLACCI, G., 2005. *Flotation froth monitoring using multiresolutional multivariate image analysis*. *Minerals Engineering* 18, 65–76
- LIU, J., GUI, W., TANG, Z., YANG, C., ZHU, J., LI, J. 2013. *Recognition of the operational statuses of reagent addition using dynamic bubble size distribution in copper flotation process*. *Minerals Engineering*, 45, 128–141.
- MEHRABI, A., MEHRSHAD, N., MASSINAEI, M., 2014. *Machine vision based monitoring of an industrial flotation cell in an iron flotation plant*. *International Journal of Mineral Processing* 133, 60–66.
- MEHRSHAD, N., MASSINAEI, M., 2011. *New image processing algorithm for measurement of bubble size distribution from flotation froth images*. *Miner. Metall. Process. J.* 28 (3), 146–150.
- MORAR, S.H., HARRIS, M.C., BRADSHAW, D.J., 2012a. *The use of machine vision to predict flotation performance*. *Miner. Eng.* 36–38, 31–36.
- MORAR, S.H., BRADSHAW, D.J., HARRIS, M.C., 2012b. *The use of the froth surface lamellae burst rate as a flotation froth stability measurement*. *Miner. Eng.* 36–38, 152–159.
- NAKHAEI, F., MOSAVI, M. R., SAM, A., NAKHAEI, A., 2012b. *Prediction of XRF analyzers error for elements on-line assaying using Kalman Filter*. *International Journal of Mining Science and Technology* 22, 595–601.
- NAKHAEI, F., MOSAVI, M.R., SAM, A., VAGHEI, Y., 2012b. *Recovery and grade accurate prediction of pilot plant flotation column concentrate: Neural network and statistical techniques*, *International Journal of Mineral Processing* 110–111, 140–154.
- NAKHAEI, F., MOSAVI, M. R., SAM, A., 2013a. *Recovery and grade prediction of pilot plant flotation column concentrate by a hybrid neural genetic algorithm*, *International Journal of Mining Science and Technology* 23, 69–77
- NAKHAEI, F., IRANNAJAD, M., 2013b. *Comparison between neural networks and multiple regression methods in metallurgical performance modeling of flotation column*. *Physicochemical Problems Mineral Process.* 49 (1), 255–266.
- NAKHAEI, F., SAM, A., MOSAVI, M. R., 2013c. *Concentrate grade prediction in an industrial flotation column using the artificial neural network*. *Arab. J. Sci. Eng.* 38 (5), 1011–1023.
- NAKHAEI, F., IRANNAJAD, M., 2017. *Sulfur removal of iron ore tailings by flotation*, *Journal of Dispersion Science and Technology*, 38 (12), 1755–1763.
- NAKHAEI, F., IRANNAJAD, M., MOHAMMADNEJAD, S. 2018. *Evaluation of column flotation froth behaviour by image analysis: effects of operational factors*. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 40 (19), 1–21.
- PRATS-MONTALBAN, J. M., JUAN, A. FERRER, A. 2011. *Multivariate image analysis: a review with applications*, *Chem. Int. Lab. Syst.*, 107, 1–23.
- PONS, M.N., BELAROU, K., SIMER, N., 2004. *Characterisation of grinding media wear by visual texture*, *Int. J. Miner. Process.* 74, 255–265.
- RAADNUI, S., 2005. *Wear particle analysis–utilization of quantitative computer image analysis: A review*. *Tribology International* 38 (10), 871–878.
- SHEAN, B. J., CILLIERS, J. J., 2011. *A review of froth flotation control*. *International Journal of Mineral Processing* 100 (3–4), 57–71.
- VENTURA-MEDINA, E., BARBIAN, N., CILLIERS, J., 2003. *Froth stability and flotation performance*. In: *Proceedings of XXII International Mineral Processing Congress, Cape Town, South Africa*, pp. 937–945.
- WANG, W., BERGHOLM, F., YANG, B., 2003. *Froth delineation based on image classification*. *Miner. Eng.* 16, 1183–1192.
- YAM, K. L., and PAPADAKIS, S., 2004. *A simple digital imaging method for measuring and analyzing colour of food surfaces*. *Journal of Food Engineering*, 61, 137–142.