*i*Journals: International Journal of Software & Hardware Research in Engineering (IJSHRE)
ISSN-2347-4890
Volume 10 Issue 10 October 2022

# Intelligent System to Detect the Extent of Chemical Reaction by Monitoring the Colour Change of the Solution

**Authors: Ranbir Minocha[a], Reetu Jain[b]**

(a) Dhirubhai Ambani Int. School, Email-id: ranbirminocha@gmail.com
(b) Supervisor, IEEE member, *On My Own Technology Pvt. Ltd., 2A, Pace House, 7 Swastik Society, Gulmohar Road, Vile Parle West*, Mumbai - 400049, India.
Email-id: reetu.jain@onmyowntechnology.com
ORCID: 0000-0002-7199-2823

***Abstract:***

Titration is a quantitative volumetric technique used for assessing the amount of the titrant required for balancing the chemical. In titration, the level of balancing is identified by the degree of change in color. The present research study focusses on developing of an intelligent system to detect the end point of the titration process. For achieving the aims and objectives of the present research an intelligent system is developed based on the Convolutional Neural Network (CNN). The CNN model is trained with images of different level of titration between Potassium Permanganate ($KMnO_4$) and Oxalic acid ($COOH - COOH$). The CNN model is used to detect three classes namely no dilution, initial dilution and final dilution. 200 images from each class is used to develop the CNN model. The dataset is divided into training and testing as 70:30. The training and testing accuracy of the developed model is 0.9201 and 0.7188 respectively whereas the training and testing loss of the model is 0.8806 and 1.2601 respectively. The strength of the proposed model is its ability to detect the level of dilution for the titration process. The identification and detection of the end point of the titration process could be automated by the usage of this intelligent system.

***Keywords:*** *Digital image, Convolutional Neural Network, Titration, End-point detection, Color change detection*

## 1. Introduction

Titration is the defined as the process of determining the quantity of one chemical by adding measured quantity of second chemical often called as titrant with which it reacts until exact chemical equivalence is observed.

### 1.1. Literature review

In the literature, there are a number of instrumental methods that have been extensively used for carrying out the quantitative chemical assessment based on titration [1]. In the conventional method, for determining the end point of the titration mostly plots are employed that are developed by using the measured signal after the addition of each increment of titrant. The shape of the curves are dependent on two set of prime factors namely indicator, titrant, analyte and formed product and second set of factor is the instrumental method chosen for the titration process [2].

In the recent years, the advancement made in the field of the digital imaging and application of deep learning (DL) for its classification has paved the way for determining the end point of titration by image processing of digital images. In the recent researches the RGB color system is used as the primary color to identify the intensities with values varying in the range 0–255 (8 bits) per color [3, 4].

With the high resolution images taken by the digital camera containing charge-coupled device integrated chip is used for studying the multivariate image regression [5], classification of inhomogeneous food matrices [6], spectral image analysis for measuring ripeness of tomatoes [7], dental plaque [8], oral diseases [9] etc. However, there exist little application of image processing of digital images and DL for its classification in the literatures. There are very few researches that involves its application for determining the end point of titration.

Some of the research paper that involves the application of image processing in titration are summarized in this section. In paper [10], the authors used digital camera for detecting the Al(III) and Fe(III) alloys with chrome azurol S (CAS) as chromogenic reagent. The RGB values of the pixel for the digital images were used as input for the artificial neural network (ANN) and the output is the end point of the reaction. The reason for employing ANN in the paper is to capture the complex relation of the pixels' RGB values and analyte concentrations. Very less work is done in using image processing and DL classification for determining the end-point of a titration process.

### 1.2. Motivation and Novelty

The aim of the present study is to develop an intelligent system to detect the end point of a titration process. In this regard, an intelligent system is developed in the study that integrates image processing with DL technique to determine the end point of a titration process. The DL technique that is used in this paper is the Convolutional Neural Network (CNN). The CNN model classify the titration process into three class viz. no dilution, initial dilution and final dilution. The model takes into account the pixel value of the digital images and maps it to the three classes. The image dataset needed to train and test the CNN model is digital pictures of the titration process between Potassium Permanganate ($KMnO_4$) and Oxalic acid ($COOH - COOH$).

The remainder of the paper is drafted as such. Section 2 briefly describes the dataset and methodology used in the research paper. Section 3 is the result and discussion that discusses the performance of the intelligent system. Finally, section 4 concludes the research paper.

## 2. Material and Methodology

This section of the paper briefly describes about the dataset and preliminary concept required to analyze the images and subsequent development of the DL model.

### 2.1. The dataset

The dataset comprises of pictures of change in color of the titration process between Potassium Permanganate ($KMnO_4$) and Oxalic acid ($COOH - COOH$). Initially, when $KMnO_4$ is added into a conical flask containing $COOH - COOH$, it gets discharged and the solution remains colorless. After complete consumption of oxalic acid ions, the endpoint is indicated by a pink colour due to excess of unreacted $KMnO_4$. Further addition of $KMnO_4$ darkens the color indicating presence of more unreacted $KMnO_4$ and after some point the there is no evident change in the color. Depending on the change of color the titration process is classified as no dilution, initial dilution and final dilution. No dilution is the phase when the solution is colorless after adding $KMnO_4$ into a conical flask containing $COOH - COOH$. Initial dilution is the phase when the color of the solution starts to change to pink due to excess of unreacted $KMnO_4$ and the final dilution is that phase of the reaction when no further change in the color is evident. Figure 1 shows the diagrammatic representation of the titration setup. For each class 200 images were taken by a digital camera from different angles.

### 2.3. Methodology

Deep learning (DL) is a family of techniques to recognize the hidden pattern in the input and output data by adding multiple layers of computational processing units [11]. In the past decade, DL is widely applied to a large set of intricate problem areas such as speech recognition, object recognition, facial recognition and many other areas. The DL techniques mostly works on the back-propagation algorithms where the error computed by comparing the predicted and target value is propagated backward thereby updating the parameter of the network. CNN is a DL technique that is mostly used for categorizing the images into different classes [12].

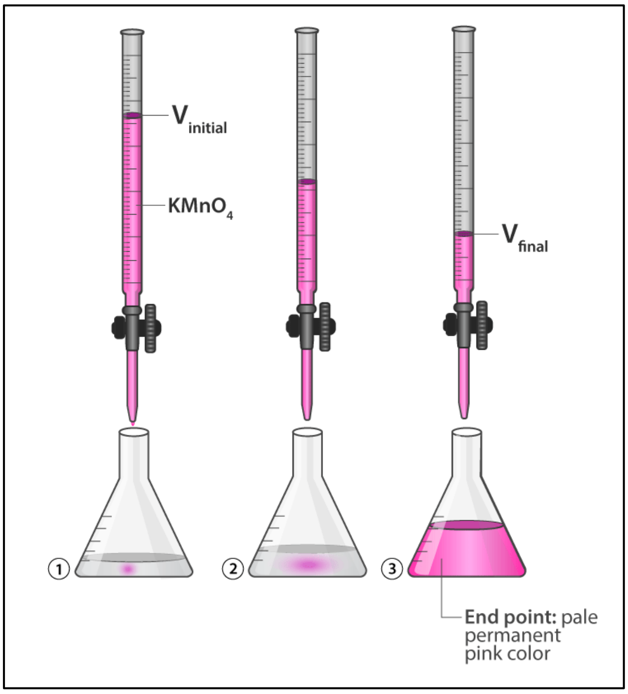
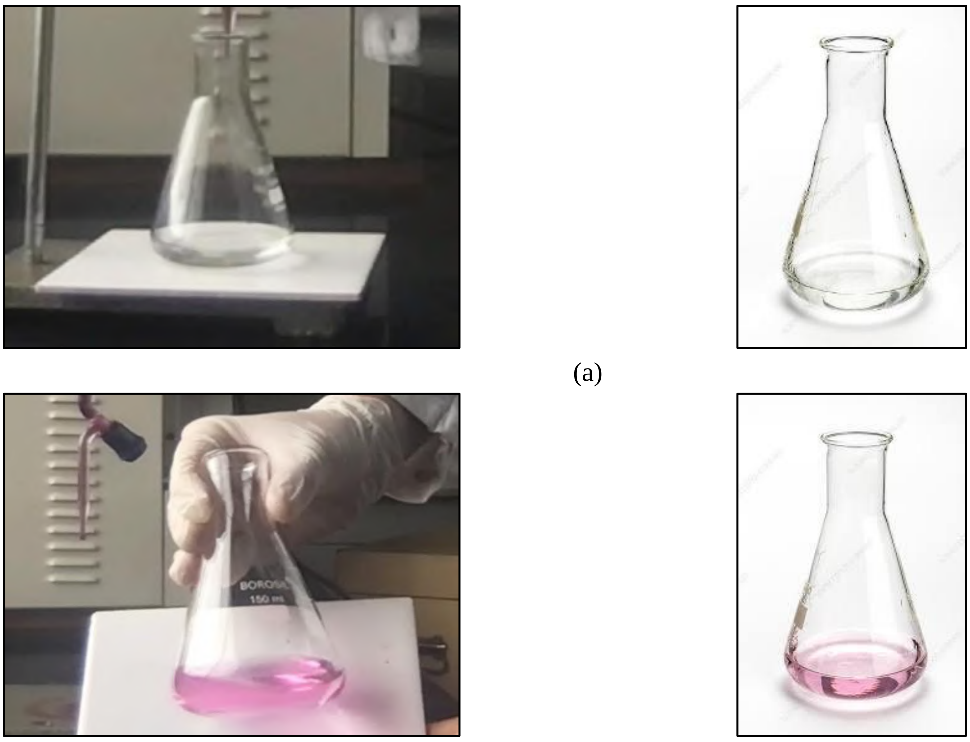



**Figure (1)**: Diagrammatic representation of the titration setup



(a)

(b)

(c)

**Figure (2)**: Images of three classes (a) no dilution, (b) initial dilution and (c) final dilution from different angles.

In the study, the CNN model is applied to detect and categorize the Titration procedure into three levels namely viz. No Titration, Initial Dilution and Final Dilution. The categorization is done on the basis of the color and based on the color it will give the approximate value of $KMnO_4$ value. CNN comprises of two parts namely training and testing. The training part determines the best value for the set of CNN parameters whereas the testing part validate the model. The CNN model is developed with the help of the 70% of the dataset. The CNN model thus built is tested with the remaining 30% dataset and then compared with the actual label determine the accuracy and preciseness of the model. The different operations involved for CNN prediction are as follows:

**Convolution operation**

The convolution operation is the first step of creating CNN model. In this step, the features are extracted from the training images and the irrelevant noises are dropped off. For extracting the features, the training images are converted into small fragments often of the size of pixel which is termed as an image matrix $(Im)$. Another user defined matrix called as the filter matrix $(Fi)$ is passed over the $Im$ throughout its width and height. The $Fi$ filters out the features from the $Im$ leaving behind the noises. Matrix multiplication of the $Im$ and $Fi$ give the resultant matrix $(Re)$ termed as convolution matrix. Mathematically,

$$Im \times Fi = Re \tag{1}$$

**Activation operation**

The second step of CNN is the activation operation. In this step, rectified linear unit (ReLU) is used as the activation operator. The ReLU is mathematically expressed as:

$$y = 0, if\ x < 0 \tag{2a}$$

$$y = x, if\ x \geq 0 \tag{2b}$$

In general ReLU is used as it allows effective as well as efficient training of the DL networks for large and complex problems [13].

**Pooling operation**

The pooling operation is the third step in the CNN modelling. In this step, the number of features required to mapped to certain classes are reduced so that an efficient model is developed. The pooling operation optimizes the number of features required for categorizing different classes. There are two ways of performing the pooling operation viz. 1) Average pooling and 2) Maximum pooling. In the average pooling method the average value of the feature map is determined whereas for maximum pooling the maximum value of the patches of the feature map is determined. In average pooling, the disadvantage is that the result is largely effected by the presence of outliers. This disadvantage can tackled by the usage of maximum pooling operation [14].

**Layer stacking**

The fourth step of developing a CNN model is the layer stacking operation. In this step, the previous three steps are repeated until the error is reduced or starts to increase.

**Fully connected layer**

This is the final operation of developing a CNN model. In this step, the matrix formed after stacking is flatten and connected to the neurons of the fully connected $(FC)$ layer.

**Classification and Prediction**

Classification is the process of categorizing the input images to one of the classes. Each neuron of $FC$ layer is connected to a class. Depending on the activation of the neuron the class is predicted. The SOFTMAX activation function is used in the present research paper to predict the classes of the images. The SOFTMAX activation function computes the probability for predicting using a multinomial probability distribution [15]. The mathematical expression for SOFTMAX activation function is:

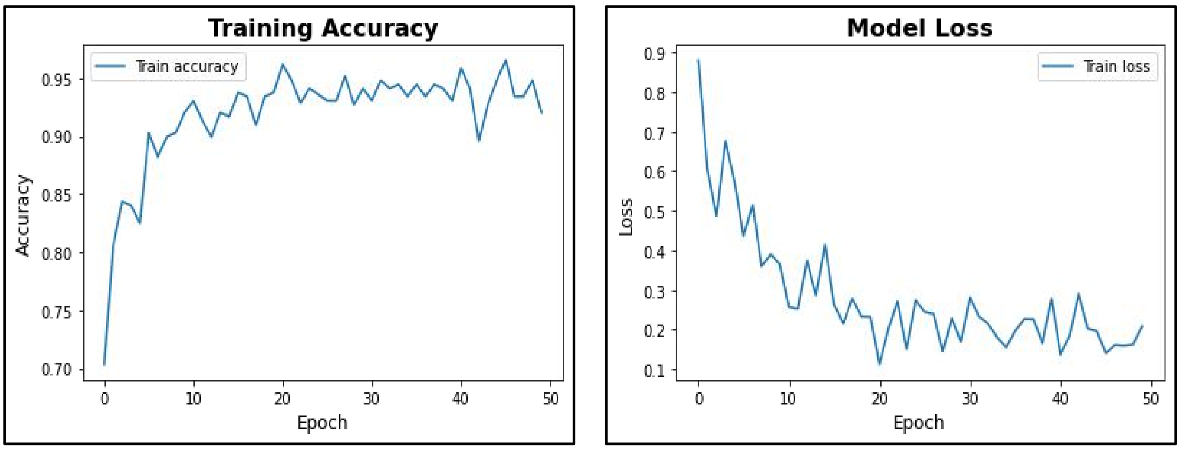
$$\sigma(z_i) = \frac{e^{(z_i)}}{\sum_{j=1}^{K}(e^{z_i})} \tag{3}$$

## 3. Results and Discussion

In this section of the paper the performance and the validation of the CNN model is discussed in brief.

### 3.1 Performance of the CNN model

The CNN model is built with the help of Tensorflow and Keras libraries which is coded in Python 3.8. For updating the parameters of the CNN model Adaptive Moment Estimation (ADAM) algorithm is used. The ADAM method is the integration of two well-known parameter optimizers namely the Gradient Descent with Momentum (GDM) and Root Mean Square Propagation (RMSE) algorithms. The error computed by the RMSE is propagated in the backward direction thereby updating the parameters by the GDM algorithm. The developed CNN model is ran for 50 epochs with 14 steps per epoch and 1 validation step. The graph for training accuracy and training loss computed for the CNN model for detecting and categorizing classes of dilution for the titration process is shown in figure 3.



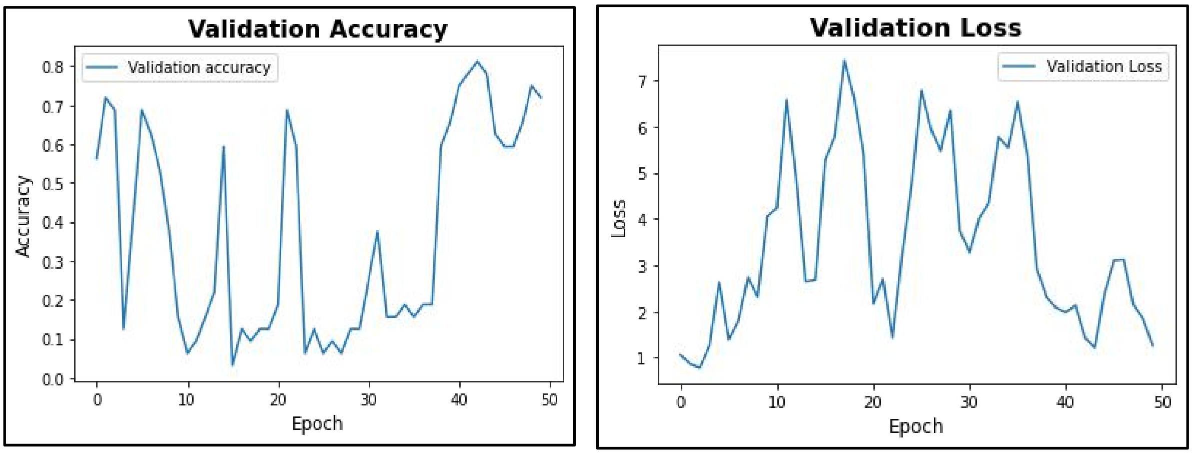

(i) (ii)

**Figure (3):** Training (i) Accuracy and (ii) loss curve of the CNN model

The training accuracy for the 1st epoch of the CNN model is 0.7036 which was increased to 0.9201 in the 50th epoch. On the other hand, the training loss for the 1st epoch of the CNN model is 0.8806 which was reduced to 0.2089 in the 50th epoch.

### 3.2 Validation of the CNN model

The CNN model developed is validated with testing images of dilution of the titration process for the different classes. The validation accuracy and validation loss graphs are shown in figure 4.




(i) (ii)

**Figure (4):** Validation (i) Accuracy and (ii) loss curve of the CNN model

The validation accuracy for the 1st epoch of the CNN model is 0.5625 which was increased to 0.7188 in the 50th epoch. Whereas, the validation loss for the 1st epoch of the CNN model is 1.059 which was reduced to 1.2601 in the 50th epoch. Although the validation loss increased but the trend is in the downward direction which may result in lower validation loss if the model is ran for more epochs.

### 4. CONCLUSIONS

The intention of the study is to develop an intelligent system that could detect the end point of a titration process. In order to achieve the aim of the present study, a system is developed that couple image processing with CNN. The developed system is trained and tested with images of titration process between $KMnO_4$ and $COOH - COOH$. The classes are the levels of dilution at three different stages namely as No Titration, Initial Dilution and Final Dilution. The developed system is trained and tested with 200 images of each class. The dataset is divided in the ratio of 70:30 where the 70% data is used for training and the remaining 30% data is used for testing the developed model. The CNN model developed employed Adam optimizer to optimize the learning rate and cross entropy to calculate the loss values. The activation of the CNN model is done using the ReLU and SOFTMAX function. The developed model is coded in Python and ran for 50 epochs with 9 steps per epoch and 1 validation step. The trained and testing accuracy of the developed model is 0.9201 and 0.7188 respectively whereas the trained and testing loss of the model is 0.8806 and 1.2601 respectively. Comprehensively, the present study is able to set the substratum for developing an intelligent system in identifying the change in chemical reaction and detecting the end point of the titration process by detecting the change in color. Hence it can be concluded that can be applied for detecting and identifying the dilution level of titration process along with the detecting the end point of the titration process.

This research work could be further extended for identifying chemical change by detecting the changes in the color.

### REFERENCES

1. Skoog, D. A., Holler, F. J., & Crouch, S. R. (2017). *Principles of instrumental analysis*. Cengage learning.
2. da Nobrega Gaiao, E., Martins, V. L., da Silva Lyra, W., de Almeida, L. F., da Silva, E. C., & Araújo, M. C. U. (2006). Digital image-based titrations. *Analytica Chimica Acta*, *570*(2), 283-290.

3. Geladi, P., & Grahn, H. F. (2006). Multivariate image analysis. *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation*.
4. Berns, R. S. (2000). Principles of Color Technology, Ch. 4 Measuring Colors.
5. Lied, T. T., & Esbensen, K. H. (2001). Principles of MIR, multivariate image regression: I: Regression typology and representative application studies. *Chemometrics and intelligent laboratory systems*, *58*(2), 213-226.
6. Antonelli, A., Cocchi, M., Fava, P., Foca, G., Franchini, G. C., Manzini, D., & Ulrici, A. (2004). Automated evaluation of food colour by means of multivariate image analysis coupled to a wavelet-based classification algorithm. *Analytica Chimica Acta*, *515*(1), 3-13.
7. Polder, G., van der Heijden, G. W., & Young, I. T. (2002). Spectral image analysis for measuring ripeness of tomatoes. *Transactions of the ASAE*, *45*(4), 1155.
8. Singh, I., & Jain, R. (2022). (2022). Statistical Analysis and Machine Learning Amalgamated Convolution Neural Network (CNN) Approach on Correlating the Dental Plaque with Cardiac Illness. *iJournals:International Journal of Software & Hardware Research in Engineering ISSN:2347-4890*, *10*(4). Retrieved from https://ijournals.in/journal/index.php/ijshre/article/view/107
9. Sehgal, I., & Jain, R. (2022). Application of deep learning neural network to create a multi-class identifier to detect and categorize periodontal diseases. *iJournals:International Journal of Software & Hardware Research in Engineering ISSN:2347-4890*, *10*(4). Retrieved from https://ijournals.in/journal/index.php/ijshre/article/view/110
10. Maleki, N., Safavi, A., & Sedaghatpour, F. (2004). Single-step calibration, prediction and real samples data acquisition for artificial neural network using a CCD camera. *Talanta*, *64*(4), 830-835.
11. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.
12. Wlodarczak, P., Soar, J., & Ally, M. (2015, October). Multimedia data mining using deep learning. In 2015 Fifth International Conference on Digital Information Processing and Communications (ICDIPC) (pp. 190-196). IEEE.
13. Behnke, S. (2003). Hierarchical neural networks for image interpretation (Vol. 2766). Springer.
14. D. Song, Z., Liu, Y., Song, R., Chen, Z., Yang, J., Zhang, C., & Jiang, Q. (2018). A sparsity-based stochastic pooling mechanism for deep convolutional neural networks. Neural Networks, 105, 340-345.
15. E. Gao, B., & Pavel, L. (2017). On the properties of the softmax function with application in game theory and reinforcement learning. arXiv preprint arXiv:1704.00805.