

Development of Segmentation and Classification Algorithms for Computed Tomography Images of Human Kidney Stone

Orobosa L. Joseph^{1*}, Waliu O. Apena^{1,2*}

¹Department of Electrical and Electronic Engineering, The Federal University of Technology, Akure, Nigeria

²Biomedical Computing and Engineering Technology, Applied Research Group, Coventry University, Coventry, United Kingdom

*Corresponding author: Orobosa L. Joseph, josephorobosa@yahoo.com; Waliu O. Apena, woapena@futa.edu.ng

Abstract: Computed tomography (CT) scan diagnostics procedures adopt the use of image information retrieval system with the help of radiographer's expertise. However, this technique is prone to errors. Significant height of accuracy is required in healthcare decision support, as 20% of CT scans are associated with error. The application of artificial intelligence (AI) can improve performance level, mitigate human error, and enhance clinical decision support in the context of time and accuracy. The study introduced machine learning algorithm to analyze stream of anonymous CT scans of kidney. The research adopted deep learning approach for segmentation and classification of kidney stone (renal calculi) images in Python (with Keras and TensorFlow) environment. A control volume of data along with 336 kidney stone images were used to train the deep learning network with 10 testing images. The training images were divided into two sets (folders) as follows; one was labeled as STONE (containing 167 images) and the other as NO-STONE (containing 169 images); 10 iterations were performed for model training. The network layers were structured as input layer in the following with 2-D convolutional neural network machine learning (CNN-ML), ReLU activation, Maxpooling, and fully connected (dense) layer including the sigmoid activation layer. The training adopted a batch size of 8 with 10% validation. The output result, upon testing the model, has an accuracy of 90%, sensitivity value of 80% and effectiveness of 89%. The segmentation and classification algorithm model could be embedded in future CT diagnostic procedure to enhance medical decision support and accuracy.

Keywords: Computed tomography scan; Diagnostics; Convolutional neural network; Accuracy; Kidney Stone

Publication date: September 2021; **Online publication:** September 30, 2021

1. Introduction

Following the discovery of X-ray imaging by Wilhelm Conrad Röntgen in the late 19th century, scientists have striven to find new ways of gaining in-vivo insight into the human anatomy and further improving existing imaging techniques. Nowadays, clinicians have a plethora of imaging modalities available for all aspects of their work, including diagnosis, treatment planning and interventional usage. Anatomical imaging modalities such as X-ray, computed tomography (CT), magnetic resonance imaging (MRI) and ultrasound (US) provide support in clinical decision and detailed examination of the tissue structure, while functional imaging modalities such as nuclear imaging or contrast agent-enhanced imaging provide insight into the metabolism and other (patho-) physiological processes within the body ^[1].

In recent years, great advances have been made in automated systems for detecting kidney diseases using CT. This involves the use of feature information of quality patients' image extracted during imaging and clinical management (procedures). The use of computational expertise such as feature extraction, image

analysis and pattern recognition techniques for classification, patients' management and clinical decision support system is attracting more interest.

Data mining techniques and validity have significant impact on clinical decision output and patients' management; hence, there is a need for the application of healthcare informatics and machine learning. Data mining plays a vital role in patients' healthcare management and promotes research in the learning curve context. Studies revealed that need for an efficient analytical methodology to detect unknown and valuable information beyond stochastic traditional empirical view of healthcare data; this could produce better information to support clinical output in diagnosis and trials. Applying data mining techniques and expertise learning techniques in healthcare domain could enhance positive patients' management and medical diagnosis. In healthcare industry, the data mining and expert system are used for prediction models in public health [2].

The process of extracting useful knowledge from stream cumbersome information is known as Data Mining. The domains of data mining could be classified as follows: image mining, opinion mining, web mining, text mining and graph mining. Some of its applications include features detection, financial data analysis, medical data analysis, social network analysis and market analysis [3].

Biomedical professional adopts the use of image feature extraction to support clinical decision. Nevertheless, there are numerous adopted machine leaning techniques in healthcare sector which depends on the nature of data and mining techniques. The study developed machine learning algorithm to analyze stream of anonymous CT scan images of kidney stone. The research adopted deep learning approach for segmentation and classification of kidney stone (renal calculi) images in Python (TensorFlow) environment to enhance clinical decision accuracy and patients' time management.

2. Literature review

The CT scan imaging is one of the most common schemes to detect diseases in the clinical practice. There are many advantages of CT scan imaging, such as safety, convenience, and low cost. However, reading CT scan imaging could be difficult during diagnosis with respect to patients' time management; hence, computer-aided diagnosis (CAD) expertise (artificial intelligence) systems are proposed towards effective patients' management. The adoption of deep learning in image segmentation and classification led to knowledge management exposure [4]. **Figures 1a** and **1b** show samples of CT scan images portraying various stone sizes.

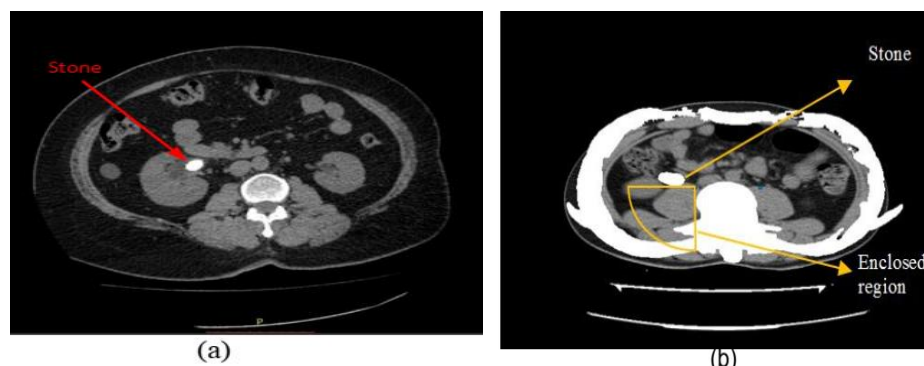


Figure 1. CT scan images of kidney with stones. (a) Small stone [5]. (b) Large stone [6].

The heights of accuracy of CT scan in medical decision has proven worthy in patient management and decision support towards diagnostic accuracy. CT scan and conventional x-rays could be adopted in anatomy decision support and clinical procedure. In conventional X-rays, the structures (organs and body framework – bones) could overlap; this could distort clinical decision accuracy and diagnosis time.

In the context of adapting expert system in clinical decision and patients' management imagery [7], four key segmentation-related challenges were identified: (i) the objects of interest are often diffuse and lack strong edges; (ii) there are often many objects, both of interest and not of interest, within a small volume; (iii) many objects have fairly similar intensity profiles, but this effect cannot be removed by simple pre-processing such as histogram equalization; and (iv) many of the objects are of roughly the same shape [7,8].

3. Methods

3.1. Image dataset acquisition

The dataset images supplied were subject to preprocessing steps to enhance its contrast as pure gray scale [9]. The principle of adopting of the image dataset into the designed method is illustrated in **Figure 2**.

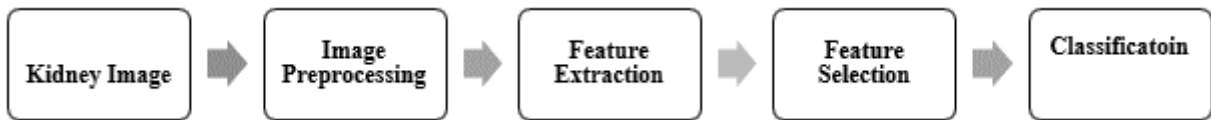


Figure 2. Block diagram for image operation

3.2. Group operations (template convolution)

Group operations calculate new pixel values from a pixel's neighborhood by using a "grouping" process. The group operation is usually expressed in terms of template convolution, where the template is a set of weighting coefficients. The template is usually square, and its size is usually odd to ensure that it can be positioned appropriately. The size is usually used to describe the template; a 3×3 template is 3 pixels wide by 3 pixels long. New pixel values are calculated by placing the template at the point of interest as shown in equation 1.

$$N_{x,y} = w_0 X_{x-1,y-1} + w_1 X_{x,y-1} + w_2 X_{x+1,y-1} + w_3 X_{x-1,y} + w_4 X_{x,y} + w_5 X_{x+1,y} + w_6 X_{x-1,y+1} + w_7 X_{x,y+1} + w_8 X_{x+1,y+1} \quad \forall x,y \in 2, N-1 \quad (1)$$

Pixel values are multiplied by the corresponding weighting coefficient and added to an overall sum. The sum (usually) evaluates a new value for the center pixel (where the template is centered) and this becomes the pixel in a new output image. If the template's position has not yet reached the end of a line, the template is then moved horizontally by one pixel and the process repeats [10]. The new image **N**, at point with coordinates x,y , the template in **Figure 3** operates on an original image **O** according to equation (2). A graphical representation of the operation is shown in **Figure 4** [10].

w_0	w_1	w_2
w_3	w_4	w_5
w_6	w_7	w_8

Figure 3. 3×3 template and weighting coefficients

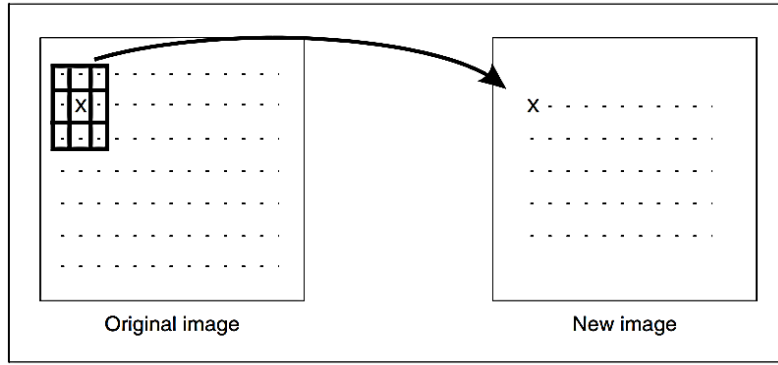


Figure 4. Template convolution process ^[10]

An alternative representation for this process is given by using the convolution notation as

$$\mathbf{N} = \mathbf{W} * \mathbf{O} \quad (2)$$

where \mathbf{N} is the new image which results from convolving the template \mathbf{W} (of weighting coefficients) with the image \mathbf{O} .

3.3. Neural network model training phase

The artificial neural network training for the image operation is expressed with the following set of equations below:

For a given set of input vector \vec{x}_k (which represents the image matrix), with its corresponding transpose, is expressed in equations (3) and (4):

$$\vec{x}_k = (x_{k1}x_{k2}x_{k3} \dots x_{km}) \quad (3)$$

$$\vec{x}_k^T = [x_{k1}x_{k2}x_{k3} \dots x_{km}]^T \quad (4)$$

where $k = 1, 2, 3, \dots, m$; and

$m = \text{number of interconnected neurons at the input,}$

The corresponding set of output vectors, \vec{y}_k , with its transpose are expressed in equations (5) and (6).

$$\vec{y}_k = (y_{k1}y_{k2}y_{k3} \dots y_{km}) \quad (5)$$

$$\vec{y}_k^T = [y_{k1}y_{k2}y_{k3} \dots y_{km}]^T \quad (6)$$

and the associated weight vector, \vec{w}_{jk} , to the input vector \vec{x}_k , is expressed in equation (7)

$$\vec{w}_{jm} = [w_{j1}(k) \quad w_{j2}(k) \quad \dots \quad w_{jm}(k)]^T \quad (7)$$

The output \vec{y}_k , can thus be expressed as:

$$y_{kj} = \sum_{i=1}^m w_{ji}(k)x_{ki} \quad (8)$$

In an expanded form, equation (8) can be written as:

$$\vec{y}_k = [w_{j1}(k) \quad w_{j2}(k) \quad \dots \quad w_{jm}(k)] \begin{bmatrix} x_{k1} \\ x_{k1} \\ \vdots \\ x_{km} \end{bmatrix} \quad (9)$$

where $j = 1, 2, 3, \dots, m$;

The matrix form of equation (9) gives the association between the input vector \vec{x}_k and the output vector \vec{y}_k

$$\begin{bmatrix} y_{k1} \\ \vdots \\ y_{km} \end{bmatrix} = \begin{bmatrix} w_{11}(k) & \dots & w_{1m}(k) \\ \vdots & \ddots & \vdots \\ w_{m1}(k) & \dots & w_{mm}(k) \end{bmatrix} \begin{bmatrix} x_{k1} \\ \vdots \\ x_{km} \end{bmatrix} \quad (10)$$

Or, in a more compact form, equation (10) can be expressed as:

$$\vec{y}_k = \vec{w}(k)\vec{x}_k \quad (11)$$

The neural network model is illustrated in **Figure 5**, thus giving graphical representation of equation (11).

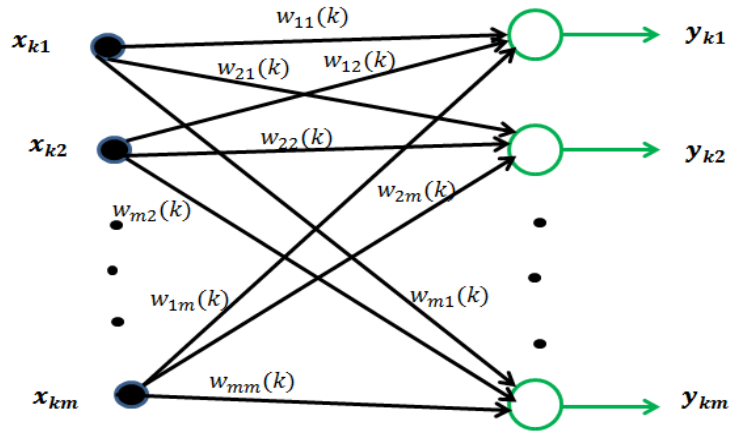


Figure 5. Neural network model of the training phase

3.4. Condition for perfect recall (testing phase)

Upon training of the network, the condition for recall is expressed as in equation (12) ^[11] as follows:

$$\vec{y} = \vec{y}_j + \sum_{k=1}^m (\vec{x}_k^T \cdot \vec{x}_j) \vec{y}_k \quad (12)$$

Equation (13) may be expressed in a more compact form as:

$$\vec{y} = \vec{y}_j + \vec{v}_j \quad (13)$$

where $\vec{v}_j = \sum_{k=1}^m (\vec{x}_k^T \cdot \vec{x}_j) \vec{y}_k$.

For $\vec{v}_j = 0$, (or approximately),

$$\vec{y} = \vec{y}_j \quad (14)$$

Thus, equation (14) presents the condition for perfect recall, \vec{y}_j is the desired or expected output of the system and \vec{v}_j is the noise associated with recalling a pattern from memory.

4. Results and discussions

The result for each image acquired size is 120 by 100 pixels, and this includes image slices resized to 50 by 50 pixel size before extracting features from the images. A convolution neural network machine learning (CNN-ML) model, i.e., a 256 nodes by 256 nodes CNN-ML model, was adopted for training (or learning) and classification, and the result was evaluated for the classifier. The result for the 256 by 256 CNN-ML models is presented in the following sub-section, the training, validation, testing and evaluation.

The preprocessing output for a given input image is shown in **Figures 6 and 7**.



Figure 6. Original input image

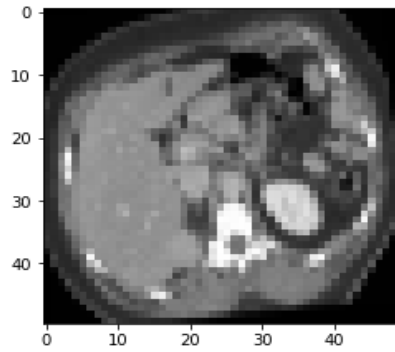


Figure 7. Processed image (50×50 normalization)

4.1. Training of model without validation

The study considered loss per epoch (entire image dataset) for the training. This is presented in **Figures 8** and **9**. It can be seen from the graph pattern that number of epoch increases while loss decreases. This indicates that with more training, the model will be less susceptible to losses.

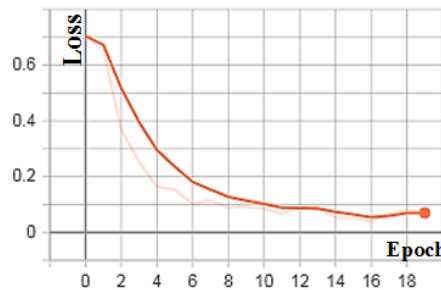


Figure 8. Epoch versus loss without validation

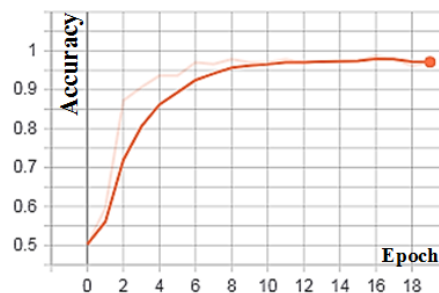


Figure 9. Epoch versus accuracy without validation

The progress of training with increase in number of epoch is shown in **Table 1**. The research further presents elapsed time and increasing batch number. This revealed the relationship between loss and training accuracy.

4.2. Training of model with validation

The result of training model with validation revealed in **Figures 10** and **11** and detailed in **Table 2**. **Figure 11** shows loss per epoch (entire image dataset) training. It can be seen from the graph that when number of epoch increases, loss decreases. This indicates that with more training, the model will be less susceptible to losses. **Figure 10** shows that with increasing number epoch of training, accuracy increases.

Table 1. Training results for 256×256 CNN-ML model training dataset without validation

Epoch	Batch size	Time elapsed (seconds)	Mini-batch loss	Mini-batch accuracy
1/20	8	88	0.7046	0.5021
2/20	8	89	0.6537	0.5957
3/20	8	93	0.3684	0.8723
4/20	8	90	0.2524	0.9064
5/20	8	94	0.1642	0.9362
6/20	8	95	0.1532	0.9362
7/20	8	91	0.1018	0.9702
8/20	8	87	0.1145	0.9660
9/20	8	92	0.0893	0.9787
10/20	8	86	0.0919	0.9702

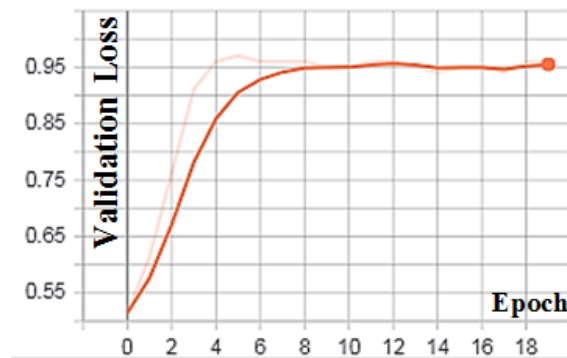


Figure 10. Epoch versus loss without validation

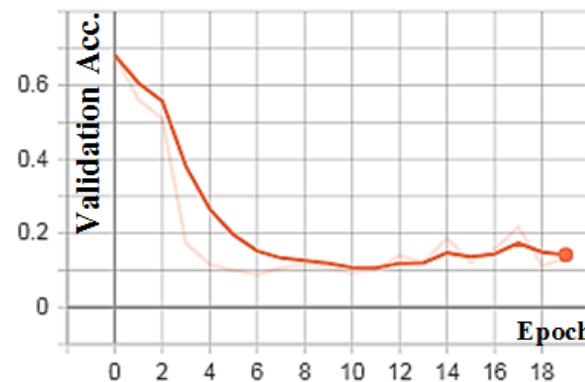


Figure 11. Epoch versus accuracy without validation

Considering **Table 2**, it can be seen that with more elapsed time and increasing batch number, the loss keeps decreasing while on the other hand the accuracy increases.

4.3. Testing stage: Evaluation on the segmentation and classification algorithm

The confusion matrix in **Table 3** was adapted to study performance evaluation on segmentation and classification algorithm on image database. For the database, 10 images were tested on the classification model; 5 of which were actual kidney stone images and the other 5 were non-kidney stone images as revealed in **Table 3**.

Table 2. Training results for the 256×256 CNN-ML model training dataset with validation

Epoch	Batch size	Time elapsed (seconds)	Mini-batch loss	Mini-batch accuracy
1/20	8	88	0.6822	0.5149
2/20	8	89	0.5613	0.6139
3/20	8	93	0.5103	0.7624
4/20	8	90	0.1732	0.9109
5/20	8	94	0.1161	0.9604
6/20	8	95	0.0999	0.9703
7/20	8	91	0.0889	0.9604
8/20	8	87	0.1059	0.9604
9/20	8	92	0.1164	0.9604
10/20	8	86	0.1073	0.9505

Table 3. Confusion matrix table for the image database

	Kidney stone (predicted)	Not kidney stone (predicted)
Kidney stone (actual)	TP = 4	FN = 1
Non-kidney stone (actual)	FP = 0	TN = 5

TP, true positive; FN, false negative; FP, false positive; TN, true negative.

The study adopted **Table 3** in equations (15) to (19) for the evaluation of the model.

$$\text{Accuracy (A)} = \frac{TP+TN}{TP+FN+FP+TN} \quad (15)$$

$$\text{Precision (P)} = \frac{TP}{TP+FP} \quad (16)$$

$$\text{Sensitivity (Se)} = \frac{TP}{TP+FN} \quad (17)$$

$$\text{Specificity (Sp)} = \frac{TN}{TN+FP} \quad (18)$$

$$\text{Effectiveness (E)} = \frac{2 \times \text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad (19)$$

The computation of equations (15) to (19) is summarized in **Table 4**.

Table 4. Segmentation and classification algorithm performance evaluation summary

A	P	Se	Sp	E
0.90	1	0.80	1	0.89

The performance evaluation of model is revealed in **Table 4**. The algorithm concluded that the classification model for the kidney stone database has an accepted (with standard scale of 0-1) accuracy (0.90), sensitivity (0.80) and effectiveness (0.89).

4.4. Error analysis of the 256×256 CNN-ML model

The study considered possibility error in the model with respect performance evaluation in **Table 5** as expressed using equations (21) and (22) as follows. The research considered two types of errors as analyzed.

Table 5.Performance error analysis for the 256×256 CNN-ML model

Sample batch image number	Model without validation accuracy	Model with validation accuracy	Absolute error (AE)	Relative error (RE)
1	0.5021	0.5149	0.0128	± 0.024859
2	0.5957	0.6139	0.0182	± 0.029647
3	0.8723	0.7624	0.1099	± 0.14415
4	0.9064	0.9109	0.0045	± 0.00494
5	0.9362	0.9604	0.0242	± 0.025198
6	0.9362	0.9703	0.0341	± 0.035144
7	0.9702	0.9604	0.0098	± 0.010204
8	0.9660	0.9604	0.0056	± 0.005831
9	0.9787	0.9604	0.0183	± 0.019055
10	0.9702	0.9505	0.0197	± 0.020726
Mean	0.8634	0.85645	0.02571	± 0.031975

This is the absolute difference between the model value and the manual value:

$$AE = |Model\ with\ Validation - Model\ without\ Validation| \quad (20)$$

The mean absolute error (MAE) of the 256×256 CNN-ML Model database of **Table 5** is 0.02571, and relative error (RE) further considered is revealed in equation (21).

$$RE = AE / Model\ Value \quad (21)$$

The RE values indicate reliability of the model values and can be regarded with respect to the model without validation. From **Table 5**, the mean relative eError (MRE) is ± **0.031975**; MRE value is significant from the third decimal point value and this indicates that the neural network deployment is reliable.

5. Conclusion

In this study, deep learning was successfully adopted to classify CT scan images of kidney stone (renal calculi). The technique involved learning model and classification model. The learning and the classification model had satisfactory performances based on the model equations and metric values. It can be concluded that the classification model for the kidney stone image database has an accuracy of 0.90 (on a scale of 0 – 1), a sensitivity of 0.80 (on a scale of 0 – 1) and an effectiveness of 0.89 (on a scale of 0 – 1). The algorithm for the segmentation and classification model could be embedded in future CT diagnosis procedure for medical decision support system and effective patients’ management. Further study should be designed to introduce smart system for kidney stone (renal calculi) diagnosis system.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Zu Berge C, Ulrich C, 2016, Real-time Processing for Advanced Ultrasound Visualization - Computer Aided Medical Procedures, Technische Universität München, Germany.
- [2] Ramya S, Radha N, 2016, Diagnosis of Chronic Kidney Disease Using Machine Learning Algorithms. *International Journal of Innovative Research in Computer and Communication Engineering*, 4(1): 812-820.
- [3] Sadaf F, Shafiya S, Chandini AH, et al., 2018, Chronic Kidney Disease and Stage Detection Using Machine Learning Classifiers, Department of CSE, Vidya Vikas Institute of Engineering and Technology, Mysuru, Karnataka. <https://doi.org/10.21467/proceedings.1.60>
- [4] Huang Q, Zhang F, Li X, 2018, Machine Learning in Ultrasound Computer-Aided Diagnostic Systems: A Survey. *Hindawi BioMed Research International*, 2018: 5137904. <https://doi.org/10.1155/2018/5137904>
- [5] Thein N, Adji TB, Hamamoto K, et al., 2019, Automated False Positive Reduction and Feature Extraction of Kidney Stone Object in 3D CT Images. *International Journal of Intelligent Engineering and System*, 12(2): 62-73. <https://doi.org/10.22266/ijies2019.0430.07>
- [6] Ebrahimi S, Mariano VY, 2015, Image Quality Improvement in Kidney Stone Detection on Computed Tomography Images. *Journal of Image and Graphics*, 3(1): 40-46.
- [7] Freedman D, Radke RJ, Zhang T, et al., 2005, Model-Based Segmentation of Medical Imagery by Matching Distributions. *IEEE Transactions on Medical Imaging*, 24(3): 281-292.
- [8] Udekwe C, Ponnle A, 2019, Application of Cardinal Points Symmetry Landmarks Distribution Model to B-Mode Ultrasound Images of Transverse Cross-section of Thin-walled Phantom Carotid Arteries. *European Journal of Engineering Research and Science*. 4(12): 96-101. <https://doi.org/10.24018/ejers.2019.4.12.1656>
- [9] OpenCV documentations, 2020, OpenCV-Python Tutorials: Image Processing in OpenCV: Smoothing Images. https://docs.opencv.org/master/d4/d13/tutorial_py_filtering.html (Accessed April 4, 2020).
- [10] Nixon M, Aguado A, 2019, Feature Extraction and Image Processing for Computer Vision, Academic Press.
- [11] Sengupta S, 2003, Neural Network and Application (Lecture Series Video), Department of Electronics and Electrical Communication Engineering I. I. T, Kharagpur.