Developed Segmentation and Classification Algorithms for Computed Tomographic Images of Human Kidney Stone (Renal calculi)

Abstract

Significant height of accuracy is required in healthcare decision support, as 20% of Computed tomography (CT) scan is associated with error. This study adopts machine learning algorithm to analyze a stream of anonymous CT scan kidney images. We train the deep learning network with a control volume of data, totaling three hundred and thirty-six (336) kidney stone images ten (10) testing images. We divide the training images into two sets (folders) as follows: we label one as STONE (containing 167 images) and the other as NO-STONE (containing 169 images). Ten (10) iterations were performed for the training of the model. The network layers structure are as follows: input layer with 2-D convolutional neural network machine learning (CNN-ML), ReLU activation, Maxpooling, fully connected (dense) layer and (including) the Sigmoid activation layer. The training adopts a batch size of eight (8) with 10% validation. The output result, upon testing the model, records an accuracy of 90%, sensitivity value of 80% and effectiveness of 89%. The segmentation and classification algorithm model could be embedded in future computed tomography (CT) diagnostic procedure to enhance medical decision support and accuracy.

Keywords: CT scan, Diagnostics, CNN, Accuracy and kidney Stone.

Introduction

With 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 to further improve 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) facilitate auxiliary 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 computed tomography (CT) [2, 3], this involves use of quality patients' image features information 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 makes significant impact in clinical decision output and patients' management; hence the 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 reveal that the need for an efficient analytical methodology to detect unknown and valuable information beyond traditional stochastic empirical view of healthcare data is long overdue; novel methods that can produce better information to support clinical decision, diagnosis and trials. Applying data mining techniques in health care domain can enhance patients' management and medical diagnosis. In the health care industry, data mining operations are incorporated into intelligent devices which are adopted for use in prediction models in public health [4, 5].

The process of extracting useful information from a stream of data 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.[6]

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

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 [7, 8]. However, reading CT scan imaging could be difficult during diagnosis with respect to patients' time management [9, 10]; hence computer-aided diagnosis (CAD) expertise systems (artificial intelligence application) are very useful towards effective patients' management [11, 12]. The adoption of deep learning in image segmentation and classification led to knowledge management exposure.[13] Figures 1a and 1b show samples of CT scan images with various stone sizes of the kidney.

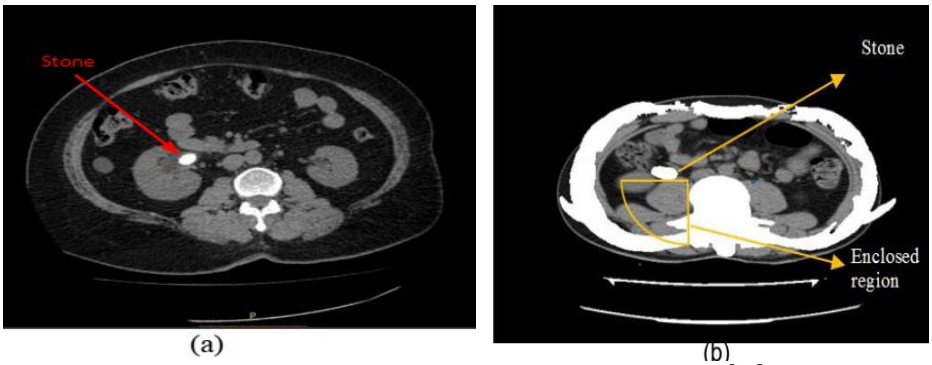

Fig. 1: CT scan images of kidney with stones: (a) small stone [14] and (b) large stone [15]

The heights of CT scan in medical decision have proven worthy in patient's management and decision support towards diagnostic accuracy. CT scan and conventional x-rays can be adopted in anatomy decision support and clinical procedures. In conventional x-rays, the structures (organs and body framework –bones) could overlap; this could distort clinical decision accuracy and diagnosis time [16, 17].

In the context of adapting expert system in clinical decision and patients' management imagery, [18]identified four key segmentation-related challenges as follow; (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, typically, 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[18, 19]

Method

Image Dataset Acquisition

The dataset images available are passed through [20] preprocessing steps to enhance its contrast as pure gray scale. The operation of adopting of the image dataset into the design method is illustrated in figure 2.

Fig. 2: Block diagram for image operation

a. Image Pixel Extraction

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.

$$N_{x,y} = \begin{matrix} w_0\,\mathrm{X}\mathbf{O}_{x-1,y-1} + w_1\,\mathrm{X}\mathbf{O}_{x,y-1} + w_2\,\mathrm{X}\mathbf{O}_{x+1,y-1} + \\ w_3\,\mathrm{X}\mathbf{O}_{x-1,y} + w_4\,\mathrm{X}\mathbf{O}_{x,y} + w_5\,\mathrm{X}\mathbf{O}_{x+1,y} + \\ w_6\,\mathrm{X}\mathbf{O}_{x-1,y+1} + w_7\,\mathrm{X}\mathbf{O}_{x,y+1} + w_8\,\mathrm{X}\mathbf{O}_{x+1,y} + \end{matrix} \forall x, y \in 2, N-1. \quad (1)$$

Pixel values are multiplied by the corresponding weighting coefficient and are added to an overall sum. The sum (usually) evaluates a new value for the centre pixel (where the template is centred) and this becomes the pixel in a new output image. If the template's position is not yet at the end of a line, the template is then moved horizontally by one pixel and the process repeats.[21] 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.[21]

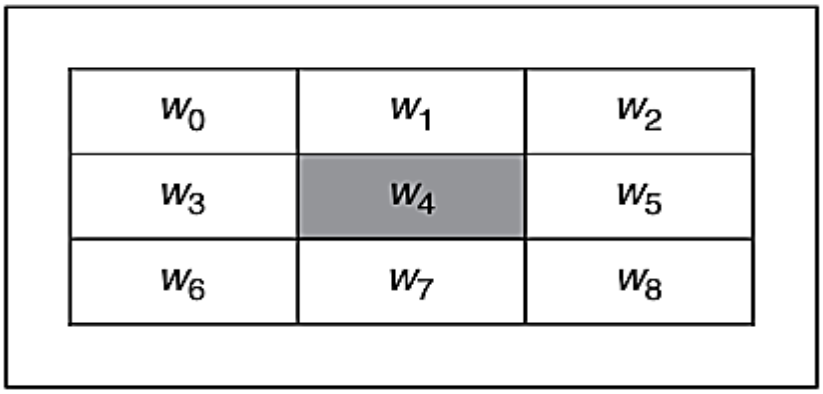

Fig. 3: 3×3 Template and weighting coefficients

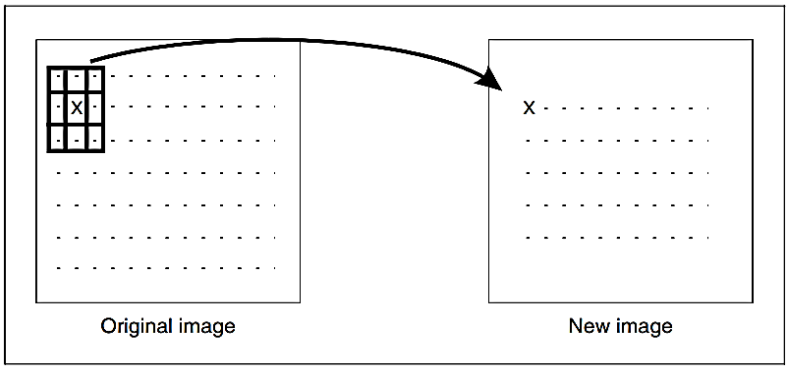

Fig. 4: Template convolution process (Nixon and Aguado, 2019)

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

$N = W * O.$ (2)

Where, N is the new image which results from convolving the template W (of weighting coefficients) with the image O.

b. 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 $\overrightarrow{x_k}$, (which represents the image matrix) with its corresponding transpose are expressed in equations (3) and (4).

$$\overrightarrow{x_k} = (x_{k1}x_{k2}x_{k3} \dots x_{km}). \quad (3)$$

$$\overrightarrow{x_k} = [x_{k1}x_{k2}x_{k3} \dots x_{km}]^T. \quad (4)$$

Where *k = 1, 2, 3, …, m; and m = number of interconnected neurons at the input,*

The corresponding set of output vectors, $\overrightarrow{y_k}$, with its transpose are expressed in equations (5) and (6).

$$\overrightarrow{y_k} = (y_{k1}y_{k2}y_{k3} \dots y_{km}). \quad (5)$$

$$\overrightarrow{y_k} = [y_{k1}y_{k2}y_{k3} \dots y_{km}]^T. \quad (6)$$

And the associated weight vector, $\overrightarrow{w_{jk}}$, to the input vector $\overrightarrow{x_k}$, is expressed in equation (7)

$$\overrightarrow{w_{jm}} = [w_{j1}(k) \quad w_{j2}(k) \quad \dots \quad w_{jm}(k)]^T. \quad (7)$$

The output $\overrightarrow{y_k}$, can thus be expressed as:

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

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

$$\overrightarrow{y_{kJ}} = [w_{j1}(k) \quad w_{j2}(k) \quad \dots \quad w_{jm}(k)] \begin{bmatrix} x_{k1} \\ x_{k1} \\ \vdots \\ x_{km} \end{bmatrix}. \quad (9)$$

Where *j = 1, 2, 3, …, m;*

The matrix form of equation (9) gives the association between the input vector $\overrightarrow{x_k}$ and the output vector $\overrightarrow{y_k}$

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

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

$$\overrightarrow{y_k} = \overrightarrow{w(k)}\overrightarrow{x_k}. \quad (11)$$

The neural network model is illustrated in figure 5, which thus shows a graphical representation of equation (11).

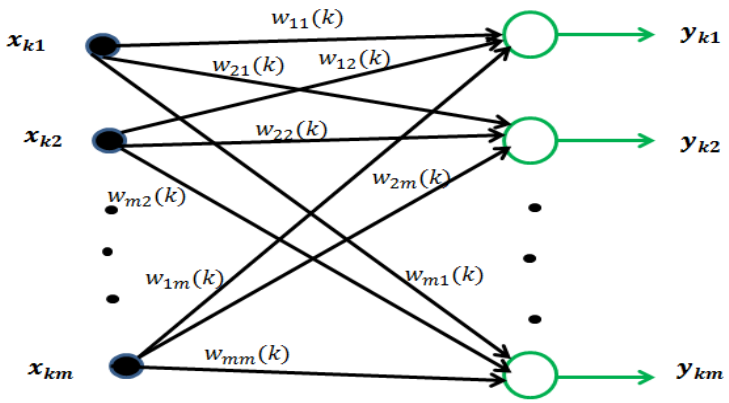

Fig. 5: Neural Network model of the training phase.

c. Condition for Perfect Recall (Testing Phase)

Upon training of the network, the condition for recall is expressed as in equation (12)[22] as follows:

$$\vec{y} = \overrightarrow{y_j} + \sum_{\substack{k=1 \\ k \neq j}}^{m} (\overrightarrow{x_k^T}.\ \overrightarrow{x_j})\ \overrightarrow{y_k}. \quad (12)$$

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

$$\vec{y} = \overrightarrow{y_j} + \overrightarrow{v_j}. \quad (13)$$

Where $\overrightarrow{v_j} = \sum_{\substack{k=1 \\ k \neq j}}^{m} (\overrightarrow{x_k^T}.\ \overrightarrow{x_j})\ \overrightarrow{y_k}$

For $\overrightarrow{v_j}$= 0, (or approximately zero),

$$\vec{y} = \overrightarrow{y_j}. \quad (14)$$

Equation (14) thus represents the condition for perfect recall. $\overrightarrow{y_j}$ is the desired or expected output of the system and $\overrightarrow{v_j}$ is the noise associated with recalling a pattern from memory.

Results and discussions

The images we acquire for the training are of size 120 by 100 pixels. We resize the image slices to 50 by 50 pixel size before extracting features from the images. We adopt a Convolution Neural Network Machine Learning (CNN-ML) model of 256 nodes by 256 nodes CNN-ML model for training (or learning) and classification; and evaluate the result for the classifier. We present the results for the 256 by 256 CNN-ML model in the following sub-sections for the training, validation, testing and evaluation of the model.

Training result for the 256x256 CNN-ML MODEL

The preprocessing output for a given input image is shown in figures 6 and 7

Fig. 6 original input image

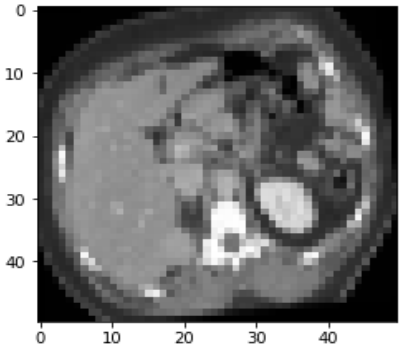

Fig. 7 processed image (50x50 normalization)

a. Training of Model Without Validation

The study considers loss per epoch (entire image dataset) for the training; we present this information in figures 8 and 9. We see from the graph pattern that with increasing number of epochs, loss decreases. This indicates that with more training, the model will be less susceptible to losses.

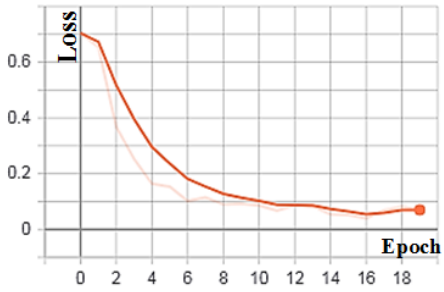

Fig. 8: Epoch versus Loss without Validation

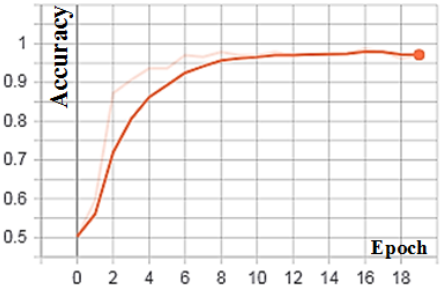

Fig. 9: Epoch versus Accuracy without Validation

The progress of training with increase in number of epochs is shown in Table 1. We further present elapse time and increasing batch number which help to reveal the relationship between loss and training accuracy.

Table 1: Training Results for *256x256 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

b. Training of Model with Validation

The result of the training model with validation is in figures 10 –11 and the details are in Table 2. Figure 10 shows that, with increasing number of epoch training, accuracy increases. Figure 11 shows loss per epoch (entire image dataset) training. The graph shows that the loss decreases with increasing number of epochs. This indicates that with more training, the model will be less susceptible to losses.

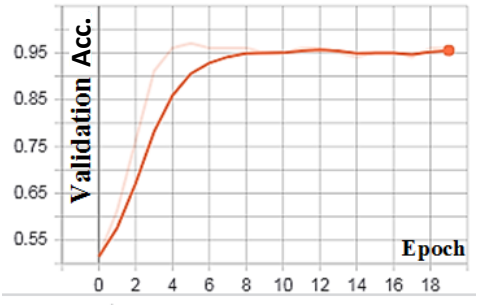

Fig 10: Epoch versus Accuracy Validation

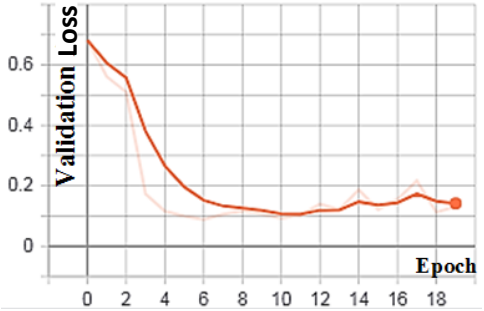

Fig 11: Epoch versus Loss Validation

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.

Table 2: Training Results for the *256x256 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

c. Testing Stage: Evaluation on the Segmentation and Classification Algorithm

We adapt the confusion matrix of Table 3 to carry out the performance evaluation on the segmentation and classification algorithm on image database. From the database, 10 images are tested on the classification model; 5 of which are actual kidney stone images and the other 5 are non-kidney stone images as in in Table 3.

Table 3: The confusion matrix table for the image database

	Kidney Stone (predicted)	Not Kidney Stone (predicted)
Kidney Stone (actual)	TP = 4	FN = 1
No Kidney Stone (actual)	FP = 0	TN = 5

Where, TP represent True Positive; FN is False Negative; FP is False Positive and TN is True Negative. The study applies Table 3 to equations (15) to (19) for the evaluation of the model.

$$\text{Accuracy}\,(\mathbf{A}) = \frac{TP + TN}{TP + FN + FP + FN}. \tag{15}$$

$$\text{Precision}(\mathbf{P}) = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{16}$$

$$\text{Sensitivity}\,(\mathbf{Se}) = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{17}$$

$$Specificity\ (\mathbf{Sp}) = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}. \tag{18}$$

$$Effectiveness\ (\mathbf{E}) = \frac{2 \times \mathrm{Precision} \times \mathrm{Sensitivity}}{\mathrm{Precision} + \mathrm{Sensitivity}}. \tag{19}$$

$$\mathbf{E} = \frac{2 \times \mathrm{P} \times \mathrm{Se}}{\mathrm{P} + \mathrm{Se}}$$

The computation of equations (15) to (19) are 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 the model reveal, as shown in Table 4, that the classification model for the Kidney Stone database has acceptable values, using a standard scale of 0-1. The values are: accuracy (0.90), sensitivity (0.80) and effectiveness (0.89).

d. Error Analysis of the 256x256 CNN-ML Model

The study considers error possibility in the model with respect to the performance evaluation in Table 5. Equations (20) and (21) expresses the error.

Table 5: Performance Error Analysis for the 256x256 CNN-ML Model

Sample Batch Image Number	**Model without Validation Acc.**	**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**

Absolute Error (AE) and Relative Error (RE)

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 256x256 CNN-ML Model result of Table 5 is 0.02571. The relative error (RE) is from equation (21).

$$RE = \frac{AE}{Model Value} \quad (21)$$

The RE value indicates the reliability of the model values. From Table 5, the Mean Relative Error (MRE) is ± 0.031975; MRE value is significant from the third decimal point value and this indicates that the Neural Network model is reliable.

Conclusion

In this study, we successfully adopt deep learning to classify CT scan kidney stone *(Renal Calculi)* images. The technique involves a learning model and a classification model. The learning and the classification model have satisfactory performances based on the model evaluation equations values. The study shows 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 can be embedded in future computed tomography (CT) diagnostic devices for medical decision support system and effective patients' management. Further study aims to introduce smart system for kidney stone *(Renal Calculi)* diagnosis system.

References

[1] zu Berge, C. and Ulrich, C, (2016): Real-time Processing for Advanced Ultrasound Visualization. *Computer Aided Medical Procedures*. Published Doctoral Dissertation, Technische Universität München, Germany.

[2] Summers R. M. (2016). Progress in Fully Automated Abdominal CT Interpretation. *AJR. American journal of roentgenology*, *207*(1), 67–79. https://doi.org/10.2214/AJR.15.15996

[3] DeFreitas, M.J., Katsoufis, C.P., Infante, J.C. *et al.* The old becomes new: advances in imaging techniques to assess nephron mass in children. *Pediatr Nephrol* **36,** 517–525 (2021). https://doi.org/10.1007/s00467-020-04477-8

[4] Ramya, S. and 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.

[5] Radha, N. and Ramya, S., (2015): Performance analysis of machine learning algorithms for predicting chronic kidney disease. *International Journal of Computer Science & Engineering Open Access* (*3*),72-76.

[6] Sadaf, F., Shafiya, S., Chandini, A. H., Nagana Devi, G. J. and Akshatha, M. (2018): Chronic Kidney Disease and Stage Detection Using Machine Learning Classifiers. *Department of CSE, Vidya Vikas Institute of Engineering and Technology, Mysuru, Karnataka.*DOI: https://doi.org/10.21467/proceedings.1.60 (01/04/2020).

[7] Mohammadshahi, M., Alipouri Sakha, M., Esfandiari, A., Shirvani, M., & Akbari Sari, A. (2019). Cost Effectiveness of Mobile versus Fixed Computed Tomography and Magnetic Resonance Imaging: A Systematic Review. *Iranian journal of public health*, *48*(8), 1418–1427.

[8] Kasban, H., El-bendary, M., Salama, D. (2015). A Comparative Study of Medical Imaging Techniques. International Journal of Information Science and Intelligent System. 4. 37-58.

[9] Committee on Diagnostic Error in Health Care; Board on Health Care Services; Institute of Medicine; The National Academies of Sciences, Engineering, and Medicine; Balogh EP, Miller BT, Ball JR, editors. Improving Diagnosis in Health Care. Washington (DC): National Academies Press (US); 2015 Dec 29. 2, The Diagnostic Process. Available from: https://www.ncbi.nlm.nih.gov/books/NBK338593/

[10] National Academies of Sciences, Engineering, and Medicine. 2015. *Improving Diagnosis in Health Care*. Washington, DC: The National Academies Press. https://doi.org/10.17226/21794.

[11] Sudipto Datta, Ranjit Barua and Jonali Das (December 12th 2019). Application of Artificial Intelligence in Modern Healthcare System, Alginates - Recent Uses of This Natural Polymer, Leonel Pereira, IntechOpen, DOI: 10.5772/intechopen.90454.

[12] Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future healthcare journal*, *6*(2), 94–98. https://doi.org/10.7861/futurehosp.6-2-94

[13] Huang, Q., Zhang, F. and Li, X., (2018): Machine learning in ultrasound computer-aided diagnostic systems: a survey. *Hindawi BioMed research international, 2018.Available online* at https://doi.org/10.1155/2018/5137904.

[14] Thein, N., Adji, T.B., Hamamoto, K. and Nugroho, H.A., (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).* DOI: 10.22266/ijies2019.0430.07

[15] Ebrahimi, S. and Mariano, V.Y., (2015): Image quality improvement in kidney stone detection on computed tomography images. *Journal of Image and Graphics*, *3*(1), 40-46.

[16] Fatma El-Zahraa Ahmed El-Gamal, Mohammed Elmogy, Ahmed Atwan, Current trends in medical image registration and fusion, Egyptian Informatics Journal, Volume 17, Issue 1, 2016, Pages 99-124, ISSN 1110-8665, https://doi.org/10.1016/j.eij.2015.09.002.

[17] INTERNATIONAL ATOMIC ENERGY AGENCY, Diagnostic Radiology Physics, , IAEA, Vienna (2014).

[18] Freedman, D., Radke, R.J., Zhang, T., Jeong, Y., Lovelock, D.M. and Chen, G.T., (2005): Model-based segmentation of medical imagery by matching distributions. *IEEE transactions on medical imaging*, *24*(3), 281-292.

[19] Udekwe, C. and 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 (Dec. 2019), 96-101. DOI: https://doi.org/10.24018/ejers.2019.4.12.1656.

[20] OpenCV documentations (2020): *OpenCV-Python Tutorials: Image Processing in OpenCV: Smoothing Images.*https://docs.opencv.org/master/d4/d13/tutorial_py_filtering.html, accessed 15/04/2020.

[21] Nixon, M. and Aguado, A., (2019): *Feature extraction and image processing for computer vision*. Academic Press.

[22] Sengupta, S (2003): Neural Network and Application (Lecture Series Video). *Department of Electronics and Electrical Communication Engineering I. I. T, Kharagpur.*

[23] OpenCV documentations (2020): *OpenCV-Python Tutorials: Image Processing in OpenCV: Smoothing Images.*https://docs.opencv.org/master/d4/d13/tutorial_py_filtering.html, accessed 15/04/2020.