

A Survey on Detection and Classification of Chronic Kidney Disease with a Machine Learning Algorithm

¹Selvarathi C, ²Devipriya P, ³Indumathi R, ⁴Kavipriya K

^{1,2,3,4}Department of Computer Science and Engineering,

M. Kumarasamy College of Engineering, Thalavapalayam-Karur-639113.

¹selvarathic.cse@mkce.ac.in, ²devipriyankl008@gmail.com, ³induragu14@gmail.com
, ⁴kavipriya151199@gmail.com

ABSTRACT

The primary goal of this study is to detect and diagnose chronic kidney diseases (CKDs), specifically kidney stones, cystic kidney disease, and suspected renal carcinoma. CKDs pave the way for a variety of diseases that aren't related to the urinal system. Coronary heart disease, stroke, cardiomyopathy, pulmonary hypertension, and heart valve disease will all become more likely as a result of it. Early diagnosis of chronic kidney disease will save lives and avoid the onset of more severe illnesses. For abdominal research, ultrasound imaging is a commonly used diagnostic tool. Chronic kidney diseases were identified using a framework that included a Histogram of directed gradient function and the KNN Algorithm in this proposed method. The multi-layered Convolution Neural Network (CNN) architecture has been trained for kidney disease classification, and the Batch prediction approach has been tested for chronic kidney disease forecasting. The detection accuracy for kidney disease is reported to be 96.67 percent. The classification of CKD ultrasound using CNN has an accuracy of 85.2 percent.

Keywords

AdaBoost, Chronic Kidney Diseases, HOG, Convolutional Neural Network, Ultrasound Image

INTRODUCTION

The most widely used of the many modern medical diagnostic modalities is ultrasound. Ultrasound images would become more common as a result of their radio-free, noninvasive, and portable properties. Ultrasound photos use high frequency sound waves to acquire the internal elements of the body, making physicians complacent for surgery due to non-intravenous results of diseases. Due to attenuation caused by sound waves, an artefact

is created, resulting in poor diagnosing efficiency. Chronic kidney disease (CKD) is one of the most common diseases that can affect the entire fluidic structure of the human body. A slight hyperechoic shadow indicates the existence of kidney cysts. The cysts are hidden behind a thick-walled septal pattern with a posterior acoustic shadow. Multiple cysts that are complex can have abnormal septal connections. Renal Cell Carcinoma (RCC), also known as body cancer, is another serious type of CKD. The three most common CKDs, kidney stones, cystic kidney disease, and renal cell carcinoma, are investigated for automatic identification and classification in this paper. Kidney stones are often calcium oxalate, uric acid, struvite, and cystic oxide, and they're usually known as echogenic material with a posterior acoustic shadow floating between the calyces infundibulum and the pole area of the kidney. The anatomical structure, form, and scale of kidney stones are all unknown. The presence of a stone will be indicated by a vertical shadow followed by a stone field. Kidney Cancer is a term used to describe a lesion in the cystic kidney. Kidney cancer usually starts as a small tumour and spreads to the lungs. In order to improve radiation-free diagnosis, more research has been done on noise reduction, segmentation, and classification on ultrasound images of kidney diseases.

The medical profession necessitates extremely meticulous findings when diagnosing diseases. Deep learning is a recent phenomenon framed with a grade of concepts put in an order, each of which has a relationship with the other. A posterior acoustic shadow floating between calyces infundibulum and pole area of the kidney. The anatomical structure, form, and scale of kidney stones are all unknown. The presence of a stone will be indicated by a vertical shadow followed by a stone field. A lesion with fluidic form and a small hyperechoic shadow indicates the existence of kidney cysts in a cystic kidney. The cysts are hidden behind a thick-walled septal pattern with a posterior acoustic shadow. Multiple cysts that are complex can have abnormal septal connections. Renal Cell Carcinoma (RCC), also known as Kidney Cancer, is another severe type of CKD. Kidney cancer usually starts as a small tumour and spreads to the lungs. In order to improve radiation-free diagnosis, more research has been done on noise reduction, segmentation, and classification on ultrasound images of kidney diseases. The medical profession necessitates extremely meticulous findings when diagnosing diseases. Deep learning is a relatively new phenomenon that entails placing a grade of concepts

in an order that is related to one another.

LITERATURE SURVEY

Many researchers are working on the estimation of CKD using a variety of classification algorithms. And those researchers get the results they predicted from their model.

Outcomes of various models have been compared by the researchers and found that the Multiclass Decision

Forest algorithm outperforms other algorithms, with a precision of about 99 percent for a reduced dataset of 14 attributes.[1]. Using various machine learning classification algorithms, we're working on reducing diagnosis time and improving diagnosis accuracy. The proposed study looks at how different phases of CKD are classified based on their severity[2].

CKD is a global health issue that is getting worse every day [9]. Many researchers use K-Nearest Neighbor (KNN) to solve classification problems[13]. There have been a variety of experiments that have used machine learning approaches to forecast and diagnose CKD[11]. Medical data mining is used in the collection of expertise and analyses information gathered from research papers, medical records, flowcharts, and evidence tables, transforming these mounds of data into usable facts for decision-making[3]. Based on tree classification in data mining scheme, researchers proposed a new variable precision rough set decision tree classification algorithm based on weighted limit number explicit field.[4].

Using data mining tools, I'm planning a study on forecasting kidney dialysis patients' survival. A data mining methodology is used in this analysis to obtain information about the relationships between these factors and patients survival[5]. K.R.Lakshmi et al. presented a success review on three methods of data mining in order to forecast survivability of kidney dialysis in 2014 [8]. Selected Classification Algorithms for Liver Disease Diagnosis: A Critical Analysis International Journal of Database Management Systems[6]. On a dataset with 400 records and 24 attributes, 12 separate classification algorithms were used[29].

RELATEDWORKS

High frequency sound waves are quickly responses with tissue region. The echoes are produced from relatively small area of tissue region with different intensity and angle. The bumpy region would cause frequency deviation that makes amplitude waves are dropped at different decibels. This deviation makes speckle noise in ultrasound scanning. Anisotropic diffusion filter is a gradient based image denoising method, Diffusion has to be taken place within the homogenous area of an image. Adaptive boosting uses set of features to form a group of classifiers and classifying them into strong and weak classifier. It eliminates weak classifier by updating weights with training classifier. Algorithm in computes different feature set with Adaboost classifier which results better performance in their work. Adaboost classifier with HOG features have significant outcomes on detection of different kidney diseases. A deep convolution neural network is recently used method that produces better results for image classification. Anisotropic diffusion filters are effectively worked for speckle noise reduction without destruct edge information, Combination of Back Propagation and DWT features have computed for medical image compression produced noteworthy solutions. [CNN] showed the comparative results for hyperspectral image cell classification using CNN with discriminative semantic features.

Preprocessing is highlighted in Figure 1.

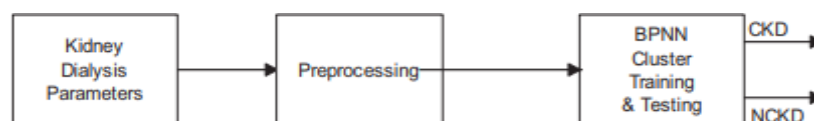


Figure 1. Shows steps in preprocessing

METHODOLOGY

With the help of clinical medical doctors and/or device marking, an enter ultrasonic photograph may contain large markers shown in Figure 2. While prognosis labelling and measuring identifiers annotated by doctors may provide difficult preliminary kidney locations. Those obvious markers can have an effect on the results of boundary detection, even

though they can no longer be removed in advance, original input images can stay standard devicedetails,andthegreenmarkerswerelabelled.

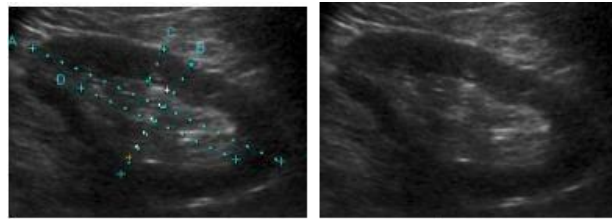


Figure2. Photography with markers

Prior to boundary identification, these markers should be disabled. As a result, the foremost step in the preprocessing module is to identify and remove the identifiers automatically. Because of their consistent grey levels and highly differentiated grey levels among neighbouring pixels, the markers can be successfully identified. Following the identification and elimination of the artificial points, the remaining empty spots should be repaired and restored using preprocessed production green markers were labelled. Prior to boundary identification, these markers should be disabled. Because of their consistent grey levels and highly differentiated grey levels among neighbouring pixels, the markers can be successfully identified.

PROPOSED SYSTEM

The proposed method uses a machine learning algorithm to categorise various categories of kidney cancer. Here the images of affected kidney are given as input dataset. About fifty images of kidney with different conditions are given which will cover nearly all the symptoms of chronic kidney disease. If we are trying to take still more images then the system will start buffering to train so it is enough with fifty images which will cover almost

all symptoms of kidney disease. Here ANN algorithm is used to categorise chronic diseases. It works more efficiently by creating artificial neurons for machine which will later form an artificial neural network and starts to mimic like a human. The better the classification results in terms of accuracy and error rate, the better the study of CKD in its early stages. Details Since there are some missing values in the dataset, pre-processing is

essential. Identification and segmentation of disease affected kidney using a clustering and classification method. In today's world, one of the most important techniques in surgical and treatment preparation for ultrasound images is kidney stone identification and segmentation. Overall processes are highlighted in Figure 3.

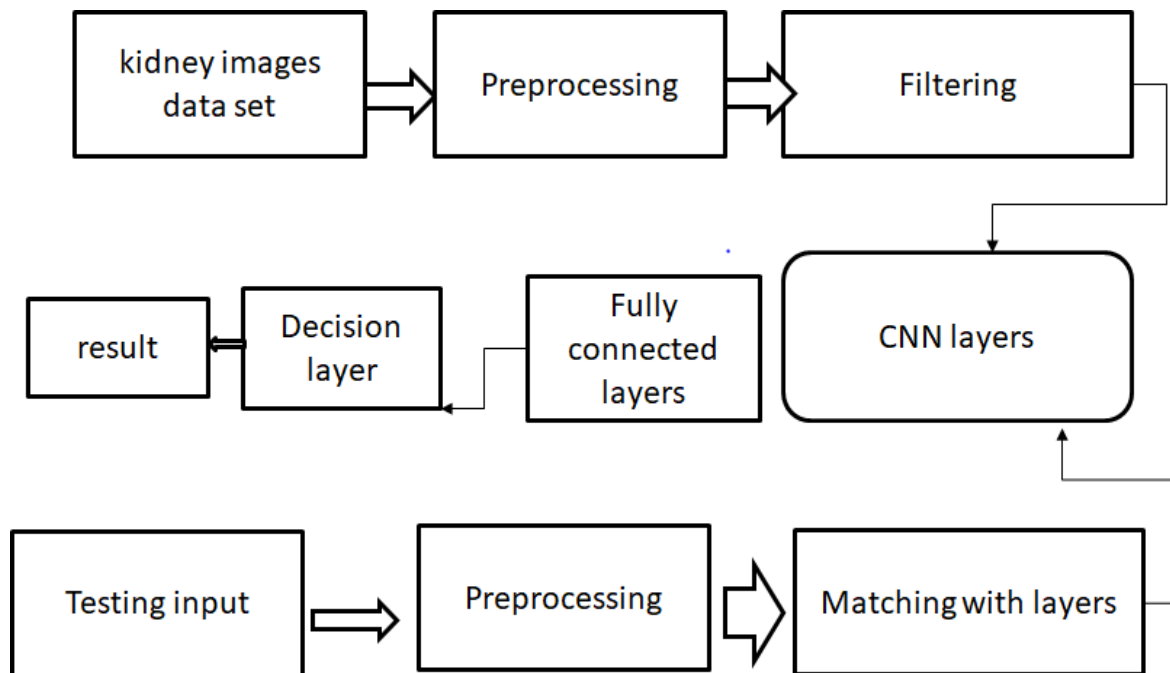


Figure 3. Overall steps in training and testing

IMAGE ACQUISITION

The Image Acquisition Toolbox streamlines the acquisition process by offering a standardised interface for MATLAB and Simulink hardware devices. A database of ultrasonic kidney photographs of various sizes has been assembled. It may be a colour or black-and-white photograph. The picture that was collected is provided as input and used for further processing.

IMAGE PREPROCESSING

In the preprocessing portion, the input image may be of various sizes, contain noise, and be in a variety of colour combinations. These essentials are needed.

SEGMENTATION

Picture segmentation is the method of splitting a digital image into several parts (sets of pixels, also known as super-pixels). The objective of segmentation is to make a picture more meaningful and easier to analyse by simplifying and/or modifying its representation. Boundaries and artefacts are usually located in images using image segmentation.

FEATURE EXTRACTION

The study of data collection, organisation, analysis, and interpretation is known as statistics. It covers every aspect of this, including data collection preparation in terms of survey and experiment design. This is what figures are all about. • Mean • Variance • Skewness • Standard deviation are all statistical features of the picture. The Gray-Level Co-Occurrence Matrix is used to analyse texture (GLCM). Gray-level co-occurrence is a mathematical procedure for estimating analysing texture that takes into account the spatial relationship between pixels. It can be updated according to the process requirements. Image noise is most noticeable in low-signal areas of an image, such as shadows or underexposed images. Filtering algorithms are used to suppress various forms of noise, such as salt and pepper noise, film grains, and so on. The Wiener filter is one of the filters used. The image acquired will be processed in the preprocessing module for proper production. . Some algorithm was used to do the pre-processing. Preprocessing is needed for all images in order to obtain the result in the better way matrix (GLCM), also known as the gray-level spatial dependency matrix. A typical compromise is 16 grey levels and a 30 or 50 pixel picture window.

CLASSIFICATION

The relationship between the data and the classes into which they are categorised must be well known in order to classify a collection of data into various classes or categories. To do this with a machine, the computer must be trained. Training is vital to classification performance. Techniques for classification were reproduced in the beginning. Features are attributes of data elements that are used to categorise them into different groups.

1st The image classifier acts as a discriminant, pitting one class against another. 2) Discriminant value is highest in one class and lowest in others (multiclass). 3) A discriminant value that is positive for one class and negative for another (twoclass).

CONCLUSION

The kidney part of the abdominal CT image is segmented out as the first step in tumour detection. This was achieved by the use of an automated segmentation process. The automatic segmentation method's strength is its generality, which allows it to be applied to a variety of organs without requiring specialisation or individual parameter settings. The results of this study show that using an automated segmentation system decreases the amount of errors that occur when performing manual segmentation. The important aspect of this project is to find a reliable predictor that can help physicians differentiate between different stages of CKD. Observation of medical images has accurately demonstrated that the changing positions of white to black ratio are in proportion to the stages of CKD patients in renal organs, despite the fact that low-resolution ultrasonic images degraded by multiplicative noise are blurred.

REFERENCES

- [1] Gunarathne W.H.S.D, Perera K.D.M, Kahandawaarachchi K.A.D.C.P, "Performance Evaluation on Machine Learning Classification Techniques for Disease Classification and Forecasting through Data Analytics for Chronic Kidney Disease (CKD)", 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering.
- [2] S.Ramya, Dr. N.Radha, "Diagnosis of Chronic Kidney Disease Using Machine Learning Algorithms," Proc. International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 1, January 2016.
- [3] K.Sudhakar & Dr.M.Manimekalai, "Study of Heart Disease Prediction using Data Mining", IJARCSSE, Volume 4, Issue 1, January 2014.
- [4] Xudong Song, Zhanzhi Qiu, Jianwei Mu, "Study on Data Mining Technology and its Application for Renal Failure Hemodialysis Medical Field", International Journal of Advancements

in Computing Technology(IJACT), Volume 4, Number 3, February 2012.

- [5] Andrew Kusiak, Bradley Dixon, Shital Shaha, (2005) Predicting survival time for kidney dialysis patients: a data mining approach, Elsevier Publication, Computers in Biology and Medicine 35, page no 311–327.
- [6] Bendi Venkata Ramana, “A Critical Study of Selected Classification Algorithms for Liver Disease Diagnosis” International Journal of Database Management Systems(IJDMS), Vol.3, No.2, May 2011(pp101–114).
- [7] Murugesan, M., Thilagamani, S. ,” Efficient anomaly detection in surveillance videos based on multi layer perception recurrent neural network”, Journal of Microprocessors and Microsystems, Volume 79, Issue November 2020, <https://doi.org/10.1016/j.micpro.2020.103303>
- [8] K.R.Lakshmi¹, Y.Nagesh² and M.Veera Krishna³, “Performance Comparison of Three Data Mining Techniques for Predicting Kidney Dialysis Survivability”, International Journal of Advances in Engineering & Technology, Mar. 2014, Vol. 7, Issue 1, pp. 242-254.
- [9] Q.-L. Zhang and D. Rothenbacher, “Prevalence of chronic kidney disease in population-based studies: Systematic review,” BMC Public Health, vol. 8, no. 1, p. 117, Dec. 2008. H. Polat, H. D. Mehr and A. Cetin, "Diagnosis of chronic kidney disease based on support vector machine by feature selection methods," J. Med. Syst., vol. 41, (4), pp. 55, 2017.
- [10] Thilagamani, S., Nandhakumar, C. .” Implementing green revolution for organic plant forming using KNN-classification technique”, International Journal of Advanced Science and Technology, Volume 29, Issue 7S, pp. 1707–1712.
- [11] S. Vijayarani, S. Dhayanand, and M. Phil, “Kidney disease prediction using SVM and ANN algorithms,” Int. J. Comput. Bus. Res. (IJCBR), vol. 6, no. 2, pp. 2229–6166, 2015.
- [12] J. Xiao et al, "Comparison and development of machine learning tools in the prediction of chronic kidney disease progression," Journal of Translational Medicine, vol. 17, (1), pp. 119, 2019.
- [13] E. K. Hashi, M. S. U. Zaman, and M. R. Hasan, “An expert clinical decision support system to predict disease using classification techniques,” in Proc. Int.

- Conf. Electr., Comput. Commun. Eng.(ECCE),Feb.2017, pp. 396–400.
- [14]Thilagamani, S., Shanti, N.,” Gaussian and gabor filter approach for object segmentation”, Journal of Computing and Information Science in Engineering, 2014, 14(2), 021006, <https://doi.org/10.1115/1.4026458>
- [15]M. Fatima and M. Pasha, "Survey of machine learning algorithms fordisease diagnostic," JournalofIntelligentLearningSystemsandApplications,vol.9,(01),pp. 1,2017.
- [16]Santhi, P., Lavanya, S., Prediction of diabetes using neural networks, International Journal of Advanced Science and Technology, 2020, 29(7 Special Issue), pp. 1160–1168
- [17]Jinhua Yu, JingluTan,YuanyuanWangb “Ultrasound speckle reductionby a SUSANcontrolledanisotropicdiffusionmethod,”ElsevierTrans.Patternrecognition, vol.43,Sep 2010,pp.3083-3092.
- [18]Rhagini, A., Thilagamani, S. ,”Women defence system for detecting interpersonal crimes”, International Journal of Advanced Science and Technology, 2020, Volume 29,Issue7S, pp. 1669–1675
- [19]J. J. D. Kelleher, B. Mac Namee and A. D'arcy, Fundamentals of Machine Learningfor PredictiveDataAnalytics:Algorithms,WorkedExamples,andCaseStudies.MITPres s, 2015.
- [20]K.Deepa, S.Thilagamani, “Segmentation Techniques for Overlapped Latent Fingerprint Matching”, International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-8 Issue-12, October 2019. DOI: [10.35940/ijitee.L2863.1081219](https://doi.org/10.35940/ijitee.L2863.1081219)
- [21]ZilongHu,JinshanTangCluster”DrivenAnisotropicDiffusionforSpeckleReductioni nUltrasound Images,” In Proc. IEEE International conference for imageprocessing, Aug 2016, pp. 2325-2329.
- [22]Deepa. K , LekhaSree. R , Renuga Devi. B , Sadhana. V , Virgin Jenifer. S ,“Cervical Cancer Classification”, International Journal of Emerging Trends in Engineering Research, 2020, 8(3), pp. 804–807 <https://doi.org/10.30534/ijeter/2020/32832020>.
- [23]Santhi, P., Priyanka, T.,Smart India agricultural information reterival system, International Journal of Advanced Science and Technology, 2020, 29(7 Special

Issue), pp. 1169–1175.

- [24] L. Rubini, “Early stage of chronic kidney disease UCI machine learning repository,” 2015. [Online]. Available: <http://archive.ics.uci.edu/ml/datasets/ChronicKidneyDisease>
- [25] S. A. Shinde and P. R. Rajeswari, “Intelligent health risk prediction systems using machine learning: a review,” IJET, vol. 7, no. 3, pp. 1019–1023, 2018.
- [26] Vijayakumar, P, Pandiaraja, P, Balamurugan, B & Karuppiah, M 2019, ‘A Novel Performance enhancing Task Scheduling Algorithm for Cloud based E-Health Environment’, International Journal of E-Health and Medical Communications , Vol 10, Issue 2, pp 102-117
- [27] Asif Salekin, John Stankovic, “Detection of Chronic Kidney Disease and Selecting Important Predictive Attributes,” Proc. IEEE International Conference on Healthcare Informatics (ICHI), IEEE, Oct. 2016, doi:10.1109/ICHI.2016.36.
- [28] Pinar Yildirim, “Chronic Kidney Disease Prediction on Imbalanced Data by Multilayer Perceptron: Chronic Kidney Disease Prediction,” Proc. 41st IEEE International Conference on Computer Software and Applications (COMPSAC), IEEE, Jul. 2017, doi:10.1109/COMPSAC.2017.84.
- [29] P. Pandiaraja, N Deepa 2019, “A Novel Data Privacy-Preserving Protocol for Multi-data Users by using genetic algorithm”, Journal of Soft Computing , Springer , Volume 23 , Issue 18, Pages 8539-8553.
- [30] A. Perotte¹, R. Ranganath, J. S. Hirsch¹, D. Blei and N. Elhadad¹, “Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis”, Oxford University press.
- [31] N Deepa , P. Pandiaraja, 2020 , “ E health care data privacy preserving efficient file retrieval from the cloud service provider using attribute based file encryption “, Journal of Ambient Intelligence and Humanized Computing , Springer , <https://doi.org/10.1007/s12652-020-01911-5>
- [32] M. A. M. Abushariah, A. A. M. Alqudah, O. Y. Adwan, and Yousef, “Automatic Heart Disease Diagnosis System Based on Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) Approaches”, Journal of Software Engineering and Applications, pp. 1055-1064.
- [33] A. Y. Al-Hyari, M. Al-Taei and M. A. Al-Taei, “Clinical Decision Support

- System for
 Diagnosis and Management of Chronic Renal Failure”, IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), (2013).
- [34] K. Sumathi, P. Pandiaraja 2019, ” Dynamic alternate buffer switching and congestion control in wireless multimedia sensor networks” , Journal of Peer-to-Peer Networking and Applications , Springer , Volume 13, Issue 6, Pages 2001-2010
- [35] Shankar, A., Pandiaraja, P., Sumathi, K., Stephan, T., Sharma, P. ,” Privacy preserving E-voting cloud system based on ID based encryption ” Journal of Peer-to-Peer Networking and Applications , Springer ,
<https://doi.org/10.1007/s12083-020-00977-4>
- [36] Welling Healthcare Private Limited available online at <http://homeo>
- [37] Sumathi, K., Naveena, K., Prashanth, P., Revanthkumar, S., Srikeerthanaa, A.P., E-health based patient surveilling device, International Journal of Emerging Trends in Engineering Research, 2020, 8(3), pp. 792–796
- [38] P. M. Patil, “Review on Prediction of Chronic Kidney disease Using Data Mining Techniques”, International Journal of Computer Science and Mobile Computing, vol.5, no.5, (2016), pp.135.
- [39] K. Strike, K. El Emam, and N. Madhavji, “Software cost estimation with incomplete data,” IEEE Trans. Softw. Eng., vol.27, no.10, pp.890-908, Oct. 2001, doi:10.1109/32.962560.
- [40] A. Gunawardana and G. Shani, “A survey of accuracy evaluation metrics of recommendation tasks,” J. Mach. Learn. Res., vol.10, pp.2935-2962, Dec.2009.
- [41] R. Tishirani and T. Hastie, “Margin trees for high-dimensional classification,” J. Mach. Learn. Res., vol.8, pp.637-652, Mar.2007.
- [42] C. Ding and X. He, “K-means clustering via principal component analysis,” in Proc. 21st Int. Conf. Mach. Learn. (ICML), 2004, pp. 225-232.
- [43] T.M.Carvajal, K.M.Viacrusis, L.F.T.Hernandez, H.T.Ho, D.M.Amalin, and K.Watanabe,
 “Machine learning methods reveal the temporal pattern of dengue incidence using meteorological factors in metropolitan Manila, Philippines,” BMC Infectious

Diseases, vol. 18, no. 1, pp. 115, Dec.2018.

- [44] T.C.-F.Yip,A.J.Ma,V.W.-S.Wong,Y.-K.Tse,H.L.-Y.Chan,P.-C.Yuen,and G. L.-H. Wong, ``Laboratory parameter-based machine learning model for excluding non-alcoholic fatty liver disease (NAFLD) in the general population," *Alimentary Pharmacol. Therapeutics*, vol. 46, no. 4, pp. 447-456, Aug. 2017.
- [45] J. Pahareeya, R. Vohra, J. Makhijani, and S. Patsariya, ``Liver patient classification using intelligent techniques," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 4, no. 2, pp. 295-299, 2014.
- [46] D. Azeez, M. A. Ali, K .B. Gan and I. Saiboon, “Comparison of adaptive neuro-fuzzy inference system and artificial neural networks model to categorize patients in the emergency department”, *Springer Plus* 2013.
- [47] M. Zhamim Hossain and G. Muhammad, “Big Data Voice Pathology Assessment Framework”, *IEEE Access*, (2016). [3] Dr. S. Vijayarani and S. Dhayanand, “Kidney disease prediction using SVM and ANN algorithms”, *International Journal of Computing and Business Research (IJCBR) ISSN (Online)*, vol. 6, no. 2, (2015), pp. 2229-6166.
- [48] M. Kumar, “Prediction of Chronic Kidney Disease Using Random Forest Machine Learning Algorithm”, “*International Journal of Computer Science and Mobile Computing*”, vol. 5, no. 2, (2016), pp. 24-33.
- [49] K. R. Anantha Padmanaban and G. Parthiban, “Applying Machine Learning Technique for Predicting the Risk of Chronic Kidney Disease”, *Indian Journal of Science and Technology*, DOI:10.17485/ijst/2016/v9i29/93880, vol. 9, no. 29.
- [50] D. A. E. H. Omran, A. H. Awad, M. A. E. R. Mabrouk, A. F. Soliman, and A. O. A. Aziz, Application of data mining techniques to explore predictors of HCC in Egyptian patients with HCV-related chronic liver, " *Asian Pac J. Cancer Prevention*, vol. 16, no. 1, pp. 381-385, Feb. 2015.