IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 4, August 2025, pp. 3182~3191
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i4.pp3182-3191

Using the ResNet-50 pre-trained model to improve the classification output of a non-image kidney stone dataset

Kazeem Oyebode[1], Anne Ngozi Odoh[2]
[1]Department of Computer and Information Science, School of Science and Technology, Pan-Atlantic University, Lagos, Nigeria
[2]School of Media and Communication, Pan-Atlantic University, Lagos, Nigeria

Article Info

Article history:

Received Jul 6, 2024
Revised Apr 4, 2025
Accepted Jun 8, 2025

Keywords:

Artificial intelligence
Deep learning
Kidney stone
Pre-trained model
Support vector machine

ABSTRACT

Kidney stone detection based on urine samples seems to be a cost-effective way of detecting the formation of stones. Urine features are usually collected from patients to determine if there is a likelihood of kidney stone formation. There are existing machine learning models that can be used to classify if a stone exists in the kidney, such as the support vector machine (SVM) and deep learning (DL) models. We propose a DL network that works with a pre-trained (ResNet-50) model, making non-image urine features work with an image-based pre-trained model (ResNet-50). Six urine features collected from patients are projected onto 172,800 neurons. This output is then reshaped into a 240 by 240 by 3 tensors. The reshaped output serves as the input to the ResNet-50. The output of this is then sent into a binary classifier to determine if a kidney stone exists or not. The proposed model is benchmarked against the SVM, XGBoost, and two variants of DL networks, and it shows improved performance using the AUC-ROC, Accuracy and F1-score metrics. We demonstrate that combining non-image urine features with an image-based pre-trained model improves classification outcomes, highlighting the potential of integrating heterogeneous data sources for enhanced predictive accuracy.

Corresponding Author:

Kazeem Oyebode
Department of Computer and Information Science, School of Science and Technology
Pan-Atlantic University
Lagos, Nigeria
Email: koyebode@pau.edu.ng

1. INTRODUCTION

Kidney stone is a urological condition that occurs when there is a deposit of acid salts in the kidneys, which reduces their function [1]. In some cases, it blocks the flow of urine and causes agonizing pain to the sufferer [1]. It is believed to affect 1 in 10 individuals [2], with its prevalence increasing globally [2], [3]. Cases of kidney stone disease have been on the rise as reported in [4]. In addition, the lack of access to care in rural and remote areas contributes to this increase [4]. In the USA cases of Kidney stones rose from 3.2 % in 1980 to 10 % in 2014 [5], affecting more men than women [6]. Recent studies from Africa [7] and Asia [8] show a similar trend in the number of cases recorded regarding kidney stones.

This development shows that there is a need to combat kidney stone disease. There are several methods available to detect kidney stones. The most common form of diagnosis is a computed tomography (CT) image scan [2], [3]. However, CT image scans require very high radiation and are expensive for low-income earners, especially in developing countries. Also, these images need to be analyzed by a specialist such as a radiologist or a nephrologist, which requires time, and is labour-intensive [9]. Furthermore, a CT scan may reveal a kidney stone, but in some instances, a specialist may not notice it [10]. Consequently, computer vision aided by AI and

deep learning (DL), is now used to automatically analyze and classify the images for early classification [3], [10], [11]. However, some of these advanced AI models for analyzing images might be found in developing countries.

The study therefore focuses on developing an improved kidney stone detection system that leverages a pre-trained model. But why is this important? This research is important because it proposes a cost-effective approach–as it only requires a urine sample, unlike other methods such as the CT scan or X-ray. One of the goals of this research is also to improve the accuracy of kidney stone detection by fusing a non-image feature into an image-based pre-trained model. This demonstrates that pre-trained models are versatile. The next question to ask is the model's development. This requires projecting the six features elicited from the urine sample and then sending them into 172,800 neurons. After that, the output of this is reshaped and fed into the ResNet-50 [12] and then to a fully connected layer, and then to the classification layer.

The paper proposes a solution that leverages DL to analyze urine samples from humans. This approach is vital because medical experts are rarely present in remote places, except in the city centre. Therefore, another objective of this research is to develop a deep-learning model that classifies whether a patient has kidney stones based on the patient's urine analysis. Thereafter, the patient could be referred to the city centre for further analysis, such as a CT scan. Researchers have also developed a DL model for kidney stone detection as found in [13] and also for classification of kidney images [14].

DL models are developed to reflect the ways neurons are fired in the human brain [15]. They are a subset of machine learning, which is the development of artificial intelligence (AI) systems without precisely programming the machine on how to learn [15]. They have recorded successes in different fields such as transportation [16], health [17], and agriculture [18] with outstanding performances. For example, Albarakati *et al*. [18], a classification model of land, from images sent from remote sensors, is implemented using the "self-attention mechanism" [19] and DL network. The objective was to determine what farm produce could do well in a given area. In the health sector, automated diagnosis tools are becoming more useful to medical practitioners. For example, in remote areas in developing countries where there is a dearth of medical doctors and limited medical resources, nurses who are not specialists, can use these assistive tools to provide the first line of diagnosis and sometimes, even an alternative diagnostic procedure for kidney stone detection. Emergency room staff can also deploy it as a cheaper, faster and more accurate diagnostic procedure [3]. This development can potentially save lives, especially when the condition is detected early. Delayed treatment can lead to renal failures [10].

For improved classification pre-trained models have been used in this study. A pre-trained model is a model that has been trained on thousands of images, such that the model could be used in different domains to solve a problem, which could be an image classification or segmentation problem [20]. Previous uses of the pre-trained model have been for land classification on images [20] and remote sensing [21]. In health, pre-trained models have been used for the classification of brain tumours [22]. In addition, pre-trained models have been used for segmentation with improved outcomes, as seen in [23].

In this research, we demonstrated that pre-trained models could be used in non-image tasks. Particularly, we demonstrated its capacity for kidney stone detection using six features from the Kaggle dataset found in [24]. First, the six-feature dataset is connected to 172,800 neurons, forming the basis to reshape the output of these neurons into a 240 by 240 by 3 input shape by using the selected ResNet-50 pre-trained model. After reshaping, we then passed it into the ResNet-50 pre-trained model, and then its output is flattened and subsequently fed into the 2048 layer neurons, and onward to the classification layer for binary classification. This idea is inspired by authors in [25], [26]. Our proposed model contributes to science in two ways – the introduction of a pre-trained model in a non-image task, and the transformation of the six input features into image-like data that can be accommodated by a pre-trained model.

Stones are not the only diseases of the kidney. Others include congenital abnormalities, cancer, and obstruction of the urinary tract. Kidney stones are of various types, which include "calcium oxalate stones, uric acid stones, calcium phosphate stones and struvite stones" [27]. Different imaging processing techniques such as CT scans, X-rays and ultrasounds are increasingly being used for the detection of internal organ and tissue disorders. These images are then analyzed by human specialists, which could sometimes lead to errors in the classification of illnesses. Also, speckle noise is produced in ultrasound pictures, which increases the difficulty in the manual detection of kidney stones. Thus, it became necessary to deploy automated tools such as image processing together with machine learning algorithms, to detect and classify kidney stones [10]. Support vector machine (SVM) can be used to classify kidney stones as demonstrated in [10].

Several deep-learning models are being used for kidney stone detection using CT images. In their study [28], they used the VGG16 model to classify a CT image, and a human specialist to ensure the accuracy of the detection. Another DL method that is widely used to classify kidney stones using CT images is convolutional neural networks (CNNs) [9], [11], [29], [30]. Some of the CNN variant networks include InceptionNet, GoogleNet, AlexNet, and ImageNet [11]. Irudayaraj [1] used four DL algorithms for classification (VGG16, ResNet-50V2, MobileNetV2, and InceptionNetV3) with InceptionNet producing the most accurate results for detecting kidney stones from CT images. While the above studies showed an

improvement over traditional methods for kidney stone detection by using DL algorithms together with CT scans, there is still a challenge regarding the huge cost of CT images for patients in remote areas of Africa and other developing countries. While there might be experts and advanced AI CT image scans, they are usually found in city centres. These gaps are what our proposed model aims to fill.

This research aims to develop a DL model that could potentially be used as the first line of diagnosis in remote areas without requiring a specialist. However, a nurse is required. The nurse takes the urine sample of a patient and then carries out the required tests. The tests should generate six features: specific gravity, urine pH level, urine osmolality, urine conductivity, urea, and urine calcium. These features are then passed into our developed AI model for classification. If the model gives a positive response, the person is referred to the city centre for a comprehensive AI CT scan. Research has been undertaken on the identification of kidney stones via urine samples. An example is with the SVM. For example, Balbin *et al.* [31] proposed an SVM system that detects calcium in urine samples. The combination of k-nearest neighbors (KNN) and SVM was used for the detection of stones in kidney images [32]. Abraham *et al.* [33] compared the performance of XGBoost with logistic regression for the identification of kidney stones using patients' health data and urine samples. They found out that the XGBoost outperformed the logistic regression in this regard. Furthermore, Alghamdi and Amoudi [34] used an ensemble approach that encapsulates the random forest (RF) for kidney stone detection.

2. METHOD

In developing the proposed model, we had to seek a publicly available dataset. We used the kidney stone dataset as seen in [24]. The dataset has six features: specific gravities, urine pH level, urine osmolality, urine conductivity, urea, and urine calcium. These urine features are important for the formation of stones in the kidney. It therefore means that if we are able to develop and train a model around these features, one could automate the detection process. As mentioned earlier, the proposed architecture (developed with Pytorch) uses a DL model with a pre-trained model. In this instance, we used the ResNet-50 pre-trained model. One reason we chose the ResNet-50 is because of its skip connections to layers earlier to mitigate the vanishing gradient problem [35]. The model architecture has been illustrated in Figure 1.

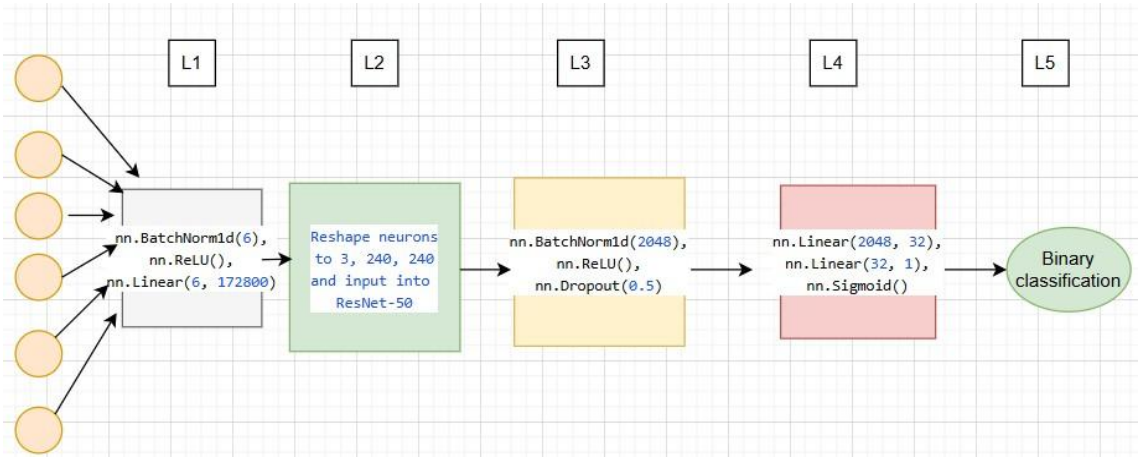

Figure 1. Proposed classification model

The proposed model as shown in Figure 1 has 5 layers L1 to L5. Layer 1 (L1) is where the 6 features of the urine sample are passed. L1 encapsulates the batch normalization later as well as the rectified linear unit (ReLU) layer. The batch normalization normalizes the input data such that they have a mean of zero to achieve an efficient training process [36]. The ReLU helps learn non-linear attributes of the input data [37]. The output of L1 is fed into another layer, which is L2, with 172,800 (3×240×240) neurons. The projected neurons from the output of L1 are reshaped in L2 and then fed into the pre-trained model residing in L2. The output of the pre-trained model is then sent over to L3 where it undergoes normalization as well as the activation layer–ReLU. The output is sent to L4 where the sigmoid gives an output within the 0-1. The output is sent to L5 for binary classification.

We trained the model on 60 epochs, learning rate of 0.0001 with a batch size of 8. The training dataset found in the Kaggle [24] was used for this model (train.csv). The train.csv file has 414 data points.

This train.csv is further split into three categories (subsets). These subsets are the training set, the validation set, and the test set. The training dataset has 291 data points, with ID values from 123 to 413, and the validation dataset has 82 data points, with IDs ranging from 0 to 81. Lastly, the testing dataset has 41 data points with an ID range of 82 to 122. From the dataset, it is clear that the training dataset takes about 70% of the entire data points, the validation dataset takes 20% and then the testing data takes 10%. It is crucial to note that the best model with the least validation error is saved and then used to evaluate the 10 per cent from the test dataset. To validate the performance of the proposed model, we replicated two variants of the proposed model, while excluding the pre-trained model. This is illustrated in Figure 2 and it shows the first variant of Figure 1.

The first variant as shown in Figure 2 omits the pre-trained model. This model aims to establish the relevance of the proposed model. The second variant implements a normal DL model without an elevated number of neurons as seen in the proposed model. This second variant is illustrated in Figure 3 and provides a balanced perspective on the significance of the proposed model. The second variant as shown in Figure 3 also has 6 inputs just like the proposed model. However, it does not have the pre-trained model. The second layer is equipped with 200 neurons, the third (L3) 250, and the fourth (100, 32, and 1). This variant also looks similar to the model proposed in [13], however, in thiers, the DL model had two layers excluding the input layer. The output is then passed onto L4. In addition, we also implemented the SVM and XGBoost models to strengthen our argument.

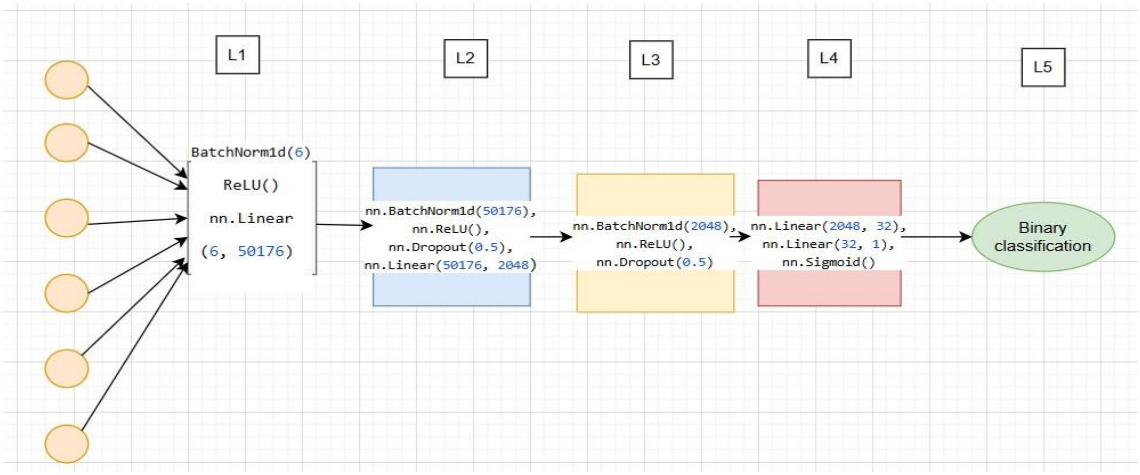

Figure 2. DL's first variant (variant one)

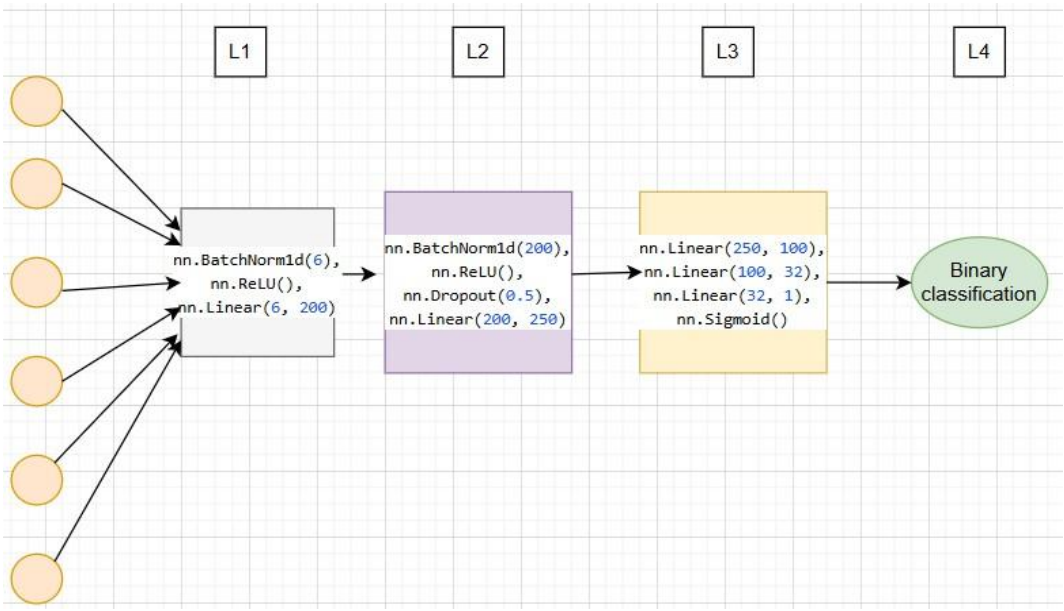

Figure 3. DL second variant

3. RESULTS AND DISCUSSION

We carried out the experiment (test) on the 10% dataset held back (41 data points), as discussed earlier. These data points were passed into the trained models in batches of 8. The predictions of all the considered models were evaluated using the the accuracy in (1), area under the receiver operating characteristic curve (AUC-ROC) [38], and the F1-score metrics in (2) to evaluate the models putforward. The AUC-ROC quantifies the performance of a binary classification model. A score of 1 shows the model performance is 100% while a performance of 0 % means the model performed poorly [38]. This value AUC-ROC indicates the tradeoff between the true positive rate (TPR) in (3) and the false positive rate (FPR) in (4) at various classification thresholds. The F1-score in (2) is a metric that strikes a balance between a model's precision in (5) and recall in (6). True positive (TP), true negative (TN), false negative (FN), and false positive (FP).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FN+FP}} \times 100\% \quad (1)$$

$$\text{F1} - \text{score} = \frac{\text{Precision} \times \text{Recall}}{\text{Precision+ Recall}} \times 100\% \quad (2)$$

$$\text{TPR} = \frac{\text{TP}}{(\text{TP+FN})} \quad (3)$$

$$\text{FPR} = \frac{\text{FP}}{(\text{FP+TN})} \quad (4)$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP+FP})} \quad (5)$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP+FN})} \quad (6)$$

In Figure 4, the ‘training vs validation loss graph’ shows the model overfits as seen from the validation loss–a gap exists between the training loss and the validation loss. One reason for this is that the training dataset is small, and therefore, it might begin to “memorize” [39] the training dataset, leading to overfitting. However, because the model uses a pre-trained model, there is the possibility to generalize, thereby performing well on the test dataset.

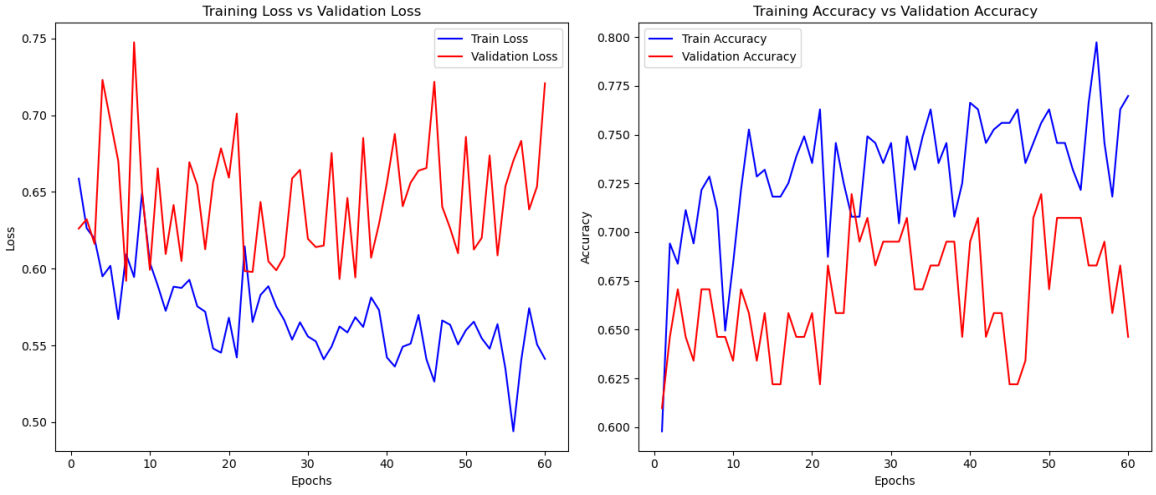

Figure 4. Training and validation loss vs training and validation accuracy on the proposed model

This is evident in Table 1, where it achieves a classification accuracy of 70.242% outperforming all other considered models, such as the DL variants one and two, as well as XGBoost. However, it failed to outperform the SVM- this is because SVM performs well on small dataset [40]. The variant two reflects the DL model developed in [13], however, with only two layers, whereas variant two has 4 layers. Nothwidstanding it did not outperform the proposed model. Another reason for this performance (proposed model) is that

reshaping the six urine features and passing them into the pre-trained model has helped it learn distinct features and textures that ultimately improved the classification outcome. The ResNet-50 played a vital role. Another contributing factor could be that the images used to train the pre-trained model (ResNet-50) aligned with the features of the urine sample. This development has ultimately helped improve its classification outcome. This improvement is also reflected in Table 2, for the F1-score of 67.29 %, as well as an AUC-ROC of 0.70288 in Table 3, as illustrated in Figure 5.

Table 1. Accuracy

Models	Accuracy (%)
Proposed	70.242
Variant one	63.412
Variant two	64.3
SVM	70.73
XGboost	60.98

Table 2. F1-score

Models	F1-score (%)
Proposed	67.29
Variant one	53.91
Variant two	53.5
SVM	64.71
XGboost	57.89

Table 3. AUC-ROC

Models	AUC-ROC
Proposed	0.70288
Variant one	0.66426
Variant two	0.7005
SVM	0.6786
XGboost	0.6238

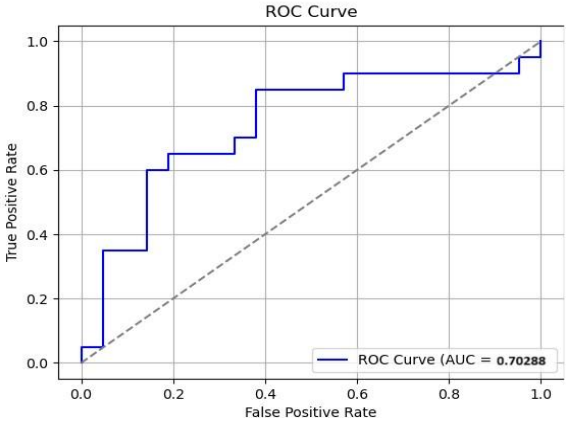

Figure 5. ROC-AUC for the proposed model

Moving on to variant one, it did not over-fit, as illustrated in Figure 6. This is due to the absence of the pre-trained model–far fewer neurons than the proposed, however enough neurons to capture complex patterns to deliver improved performance. The development gave an accuracy score of 63.412 in Table 1, an F1-score of 53.91 in Table 2. The second variant widens the gap between the training loss and the validation loss in Figure 7. One explanation for this could be that there are not enough neurons to learn the complexities of the training data-this means the limited neurons may begin to "memorize" [39] the training data leading to the over-fitting in Figure 7. The first variant has an AUC-ROC in Figure 8 of 0.66426 in Table 3. The ROC AUC figure for the second variant is illustrated in Figure 9. The ROC AUC figures of SVM and XGboost are illustrated in Figures 10 and 11.

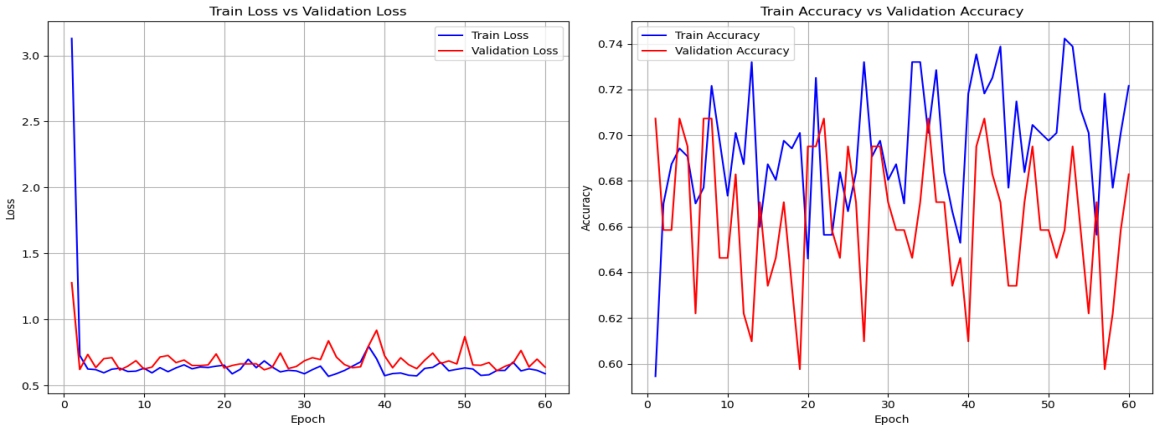

Figure 6. Training and validation loss vs training and validation accuracy on the first variant (variant one)

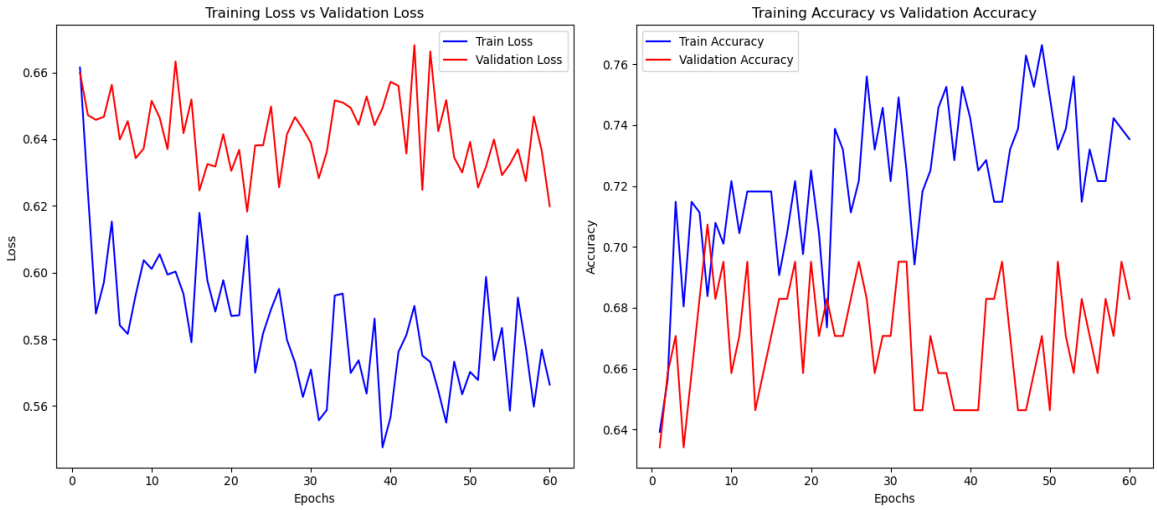

Figure 7. Training and validation loss vs training and validation accuracy on the second variant (variant two)

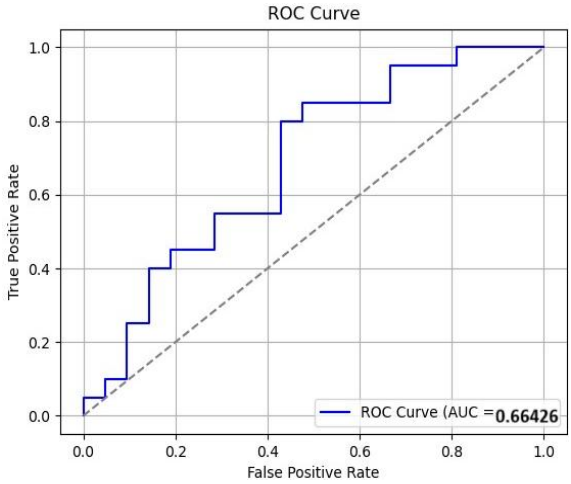

Figure 8. ROC-AUC for variant one

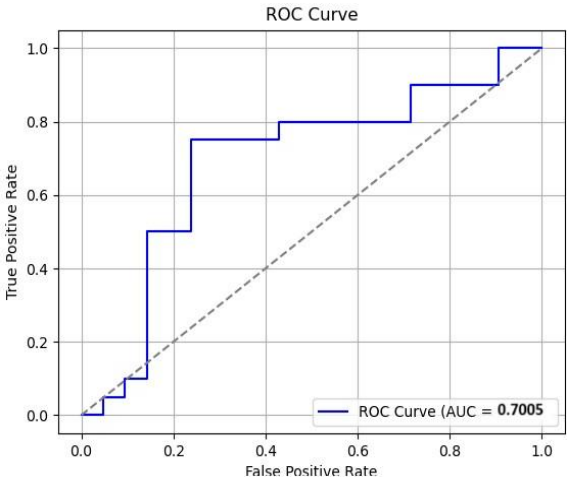

Figure 9. ROC-AUC for variant two

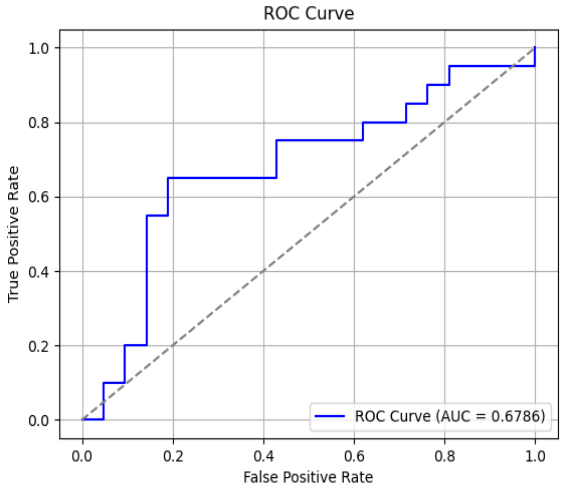

Figure 10. ROC-AUC for SVM

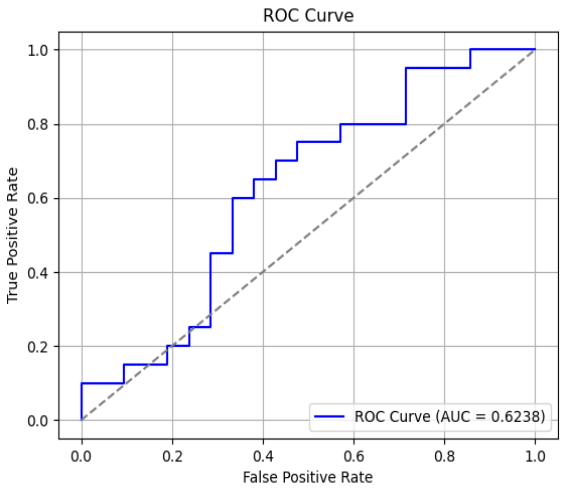

Figure 11. ROC-AUC for XGboost

From the result presented above, it is evident that the ResNet-50 improved the classification accuracy of the model proposed–using the accuracy, F1-score as well as the ROC-AUC metric. The ROC-AUC generates a plot of the TP rate and the FP rate. It depicts a model's scorecard. The TPR gives how often a model predicts a kidney stone disease, while the FPR provides how often a model gives a wrong classification of negative instances as positive [38]. It is worth noting that the proposed model gave the highest ROC-AUC score of 0.70288. This indicates that the model is 0.70288 effective out of a maximum score of 1 in distinguishing if a patient has a kidney stone or not. Other considered models are below the 0.70288 mark. A score of 0.70288 clearly reduces the extent to which a model makes misclassification. Therefore, it is safer to deploy such a model in a real-life scenario than to deploy any of the other considered models. We deployed the proposed model on Huggingface as seen from the URL-https://huggingface.co/spaces/Kazeemkz/Kidneystone_detection using Streamlit and Gradio [41]. For improvement, more dataset needs to collected to assist the model to learn the complex partterns of urine features. This would potentially have improved the classification accuracy.

4. CONCLUSION

This paper proposes a pre-trained model for kidney stone classification based on six features: specific gravity, urine pH level, urine osmolality, urine conductivity, urea, and calcium. These features are projected into a higher-dimensional space of 172,800 and then reshaped into a 240 by 240 by 3 array to fit the input shape of the selected pre-trained model-ResNet-50. The output of the pre-trained model is passed through another layer with 2048 neurons, then to a layer with 32 neurons, and finally to a single neuron for classification. We evaluated the proposed model against two other variants of DL models the SVM and the XGboost models. Using the ROC-AUC, and the F1-score metrics it is established that the model proposed outperformed other considered models. The evaluation demonstrates that an image-based pre-trained model (proposed) could enhance the classification accuracy of non-image-based datasets highlighting the potential of integrating heterogeneous data sources for enhanced predictive accuracy. For improvement, more data needs to be collected to assist the model to learn the complex partterns of urine features. This could potentially have improved the classification accuracy.

FUNDING INFORMATION

No funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Kazeem Oyebode	✓	✓	✓	✓	✓	✓		✓	✓		✓		✓	
Anne Ngozi Odoh	✓	✓				✓		✓	✓	✓		✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that supports the findings of this study is openly available in Kaggle at https://kaggle.com/competitions/playground-series-s3e12.

REFERENCES

[1] A. A. Irudayaraj, "Kidney stone detection using deep learning methodologies," *M.Sc. Project*, School of Computing, National

College of Ireland, Dublin, Ireland, 2022.
[2] O. Sabuncu and B. Bilgehan, "Performance evaluation for various deep learning (DL) methods applied to kidney stone diseases," in *2021 International Conference on Forthcoming Networks and Sustainability in AIoT Era, FoNeS-AIoT 2021*, 2021, pp. 1–3, doi: 10.1109/FoNeS-AIoT54873.2021.00010.
[3] Y. Y. Liu, Z. H. Huang, and K. W. Huang, "Deep learning model for computer-aided diagnosis of urolithiasis detection from kidney–ureter–bladder images," *Bioengineering*, vol. 9, no. 12, 2022, doi: 10.3390/bioengineering9120811.
[4] M. Almusafer *et al.*, "Unveiling the burden of nephrolithiasis in low- and lower-middle-income countries: a review on its presentation, risk factors, treatment practices, and future directions," *Société Internationale d'Urologie Journal*, vol. 5, no. 5, pp. 361–370, Oct. 2024, doi: 10.3390/siuj5050055.
[5] T. Alelign and B. Petros, "Kidney stone disease: an update on current concepts," *Advances in Urology*, vol. 6, no. 2, pp. 28–35, 2018, doi: 10.1155/2018/3068365.
[6] C. D. Scales, A. C. Smith, J. M. Hanley, and C. S. Saigal, "Prevalence of kidney stones in the United States," *European Urology*, vol. 62, no. 1, pp. 160–165, Jul. 2012, doi: 10.1016/j.eururo.2012.03.052.
[7] N. Borumandnia *et al.*, "Longitudinal trend of urolithiasis incidence rates among world countries during past decades," *BMC Urology*, vol. 23, no. 1, Oct. 2023, doi: 10.1186/s12894-023-01336-0.
[8] J. Li, Y. Zhao, Z. Xiong, and G. Yang, "Global, regional, and national incidence and disability-adjusted life-global, regional, and, national incidence and disability-adjusted life-years for urolithiasis in 195 countries and territories, 1990–2019: results from the global burden of disease stud," *Journal of Clinical Medicine*, vol. 12, no. 3, Jan. 2023, doi: 10.3390/jcm12031048.
[9] A. Caglayan, M. O. Horsanali, K. Kocadurdu, E. Ismailoglu, and S. Guneyli, "Deep learning model-assisted detection of kidney stones on computed tomography," *International Brazilian Journal of Urology*, vol. 48, no. 5, pp. 830–839, 2022, doi: 10.1590/S1677-5538.IBJU.2022.0132.
[10] P. K. Min, D. Bhattacharyya, B. C. Min, T. H. Kim, and K. Mito, "Enhanced kidney stone identification using ultrasonographic images in image processing," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 4, pp. 477–484, 2024.
[11] B. Reuben and C. Narmadha, "Effective kidney stone prediction based on optimized YOLOv7 segmentation and deep learning classification," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 1, pp. 183–192, 2024.
[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
[13] A. Jarghon and S. S. Abu-Naser, "Predicting kidney stone presence from urine analysis: a neural network approach using JNN," *International Journal of Academic Information Systems Research*, vol. 7, no. 9, pp. 32–39, 2023.
[14] J. Rubia, S. Shibi, B. Lincy, J. P. Catherin, Vigneshwaran, and E. Nithila, "Automatic kidney disease prediction using deep learning techniques," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 36, no. 3, pp. 1798–1806, 2024, doi: 10.11591/ijeecs.v36.i3.pp1798-1806.
[15] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.
[16] X. Yin, G. Wu, J. Wei, Y. Shen, H. Qi, and B. Yin, "Deep learning on traffic prediction: methods, analysis, and future directions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 4927–4943, Jun. 2022, doi: 10.1109/TITS.2021.3054840.
[17] Q. Ran and D. Han, "Design of mental health consultation system based on deep learning algorithm," in *2023 International Conference on Computer Science and Automation Technology (CSAT)*, Oct. 2023, pp. 257–262, doi: 10.1109/CSAT61646.2023.00074.
[18] H. M. Albarakati *et al.*, "A novel deep learning architecture for agriculture land cover and land use classification from remote sensing images based on network-level fusion of self-attention architecture," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 6338–6353, 2024, doi: 10.1109/JSTARS.2024.3369950.
[19] A. Vaswani *et al.*, "Attention is all you need," in *31st Conference on Neural Information Processing Systems (NIPS 2017)*, 2017, pp. 1–11.
[20] B. Zhao, B. Huang, and Y. Zhong, "Transfer learning with fully pretrained deep convolution networks for land-use classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 9, pp. 1436–1440, Sep. 2017, doi: 10.1109/LGRS.2017.2691013.
[21] Z. Huang, M. Zhang, Y. Gong, Q. Liu, and Y. Wang, "Generic knowledge boosted pretraining for remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–13, 2024, doi: 10.1109/TGRS.2024.3354031.
[22] K. Muhammad, S. Khan, J. Del Ser, and V. H. C. de Albuquerque, "Deep learning for multigrade brain tumor classification in smart healthcare systems: a prospective survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 2, pp. 507–522, Feb. 2021, doi: 10.1109/TNNLS.2020.2995800.
[23] W. Cai, L. Xie, W. Yang, Y. Li, Y. Gao, and T. Wang, "DFTNet: dual-path feature transfer network for weakly supervised medical image segmentation," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, no. 4, pp. 2530–2540, Jul. 2023, doi: 10.1109/TCBB.2022.3198284.
[24] W. Reade and A. Chow, "Binary classification with a tabular kidney stone prediction dataset," *Kaggle*, 2023. [Online]. Available: https://kaggle.com/competitions/playground-series-s3e12
[25] A. Sharma, E. Vans, D. Shigemizu, K. A. Boroevich, and T. Tsunoda, "DeepInsight: a methodology to transform a non-image data to an image for convolution neural network architecture," *Scientific Reports*, vol. 9, no. 1, Aug. 2019, doi: 10.1038/s41598-019-47765-6.
[26] Y. Zhu *et al.*, "Converting tabular data into images for deep learning with convolutional neural networks," *Scientific Reports*, vol. 11, no. 1, May 2021, doi: 10.1038/s41598-021-90923-y.
[27] W. Zhu, R. Zhou, Y. Yao, T. Campbell, R. Jain, and J. Luo, "SegPrompt: using segmentation map as a better prompt to finetune deep models for kidney stone classification," in *Proceedings of Machine Learning Research*, 2023, vol. 227, pp. 1680–1690.
[28] B. Manoj, N. Mohan, S. K. S, and S. K. P, "Automated detection of kidney stone using deep learning models," in *2022 2nd International Conference on Intelligent Technologies (CONIT)*, Hubli, India, Jun. 2022, pp. 1–5, doi: 10.1109/CONIT55038.2022.9847894.
[29] K. M. Black, H. Law, A. Aldoukhi, J. Deng, and K. R. Ghani, "Deep learning computer vision algorithm for detecting kidney stone composition," *BJU International*, vol. 125, no. 6, pp. 920–924, Jun. 2020, doi: 10.1111/bju.15035.
[30] P. S. Ramesh, S. Patel, K. D. Bamane, Y. UshaRani, M. Tiwari, and T. Karthikeyan, "Automatic kidney stone detection using deep learning method," *Journal of Advanced Zoology*, vol. 44, no. 4, pp. 100–109, Nov. 2023, doi: 10.17762/jaz.v44iS4.2176.
[31] J. J. R. Balbin, G. V. Magwili, L. D. Valiente, D. L. B. Gawaran, N. E. R. Lumapas, and A. M. Umali, "Detection and identification of triple phosphate crystals and calcium oxalate crystals in human urine sediment using haar feature, adaptive boosting, and support vector machine via openCV," in *Proceedings of the 2020 10th International Conference on Biomedical Engineering and*

Technology, Sep. 2020, pp. 34–39, doi: 10.1145/3397391.3397415.
[32] J. Verma, M. Nath, P. Tripathi, and K. K. Saini, "Analysis and identification of kidney stone using Kth nearest neighbour (KNN) and support vector machine (SVM) classification techniques," *Pattern Recognition and Image Analysis*, vol. 27, no. 3, pp. 574–580, Jul. 2017, doi: 10.1134/S1054661817030294.
[33] A. Abraham, N. L. Kavoussi, W. Sui, C. Bejan, J. A. Capra, and R. Hsi, "Machine learning prediction of kidney stone composition using electronic health record-derived features," *Journal of Endourology*, vol. 36, no. 2, pp. 243–250, Feb. 2022, doi: 10.1089/end.2021.0211.
[34] H. Alghamdi and G. Amoudi, "Using machine learning for non-invasive detection of kidney stones based on laboratory test results: a case study from a Saudi Arabian hospital," *Diagnostics*, vol. 14, no. 13, Jun. 2024, doi: 10.3390/diagnostics14131343.
[35] Z. Duan, F. Wang, B. Wang, G. Luo, and Z. Jiang, "An adapted ResNet-50 architecture for predicting flow fields of an underwater vehicle," *IEEE Access*, vol. 12, pp. 66398–66407, 2024, doi: 10.1109/ACCESS.2024.3399077.
[36] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on ImageNet classification," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015, pp. 1026–1034, doi: 10.1109/ICCV.2015.123.
[37] V. M. Vargas, P. A. Gutiérrez, J. Barbero-Gómez, and C. Hervás-Martínez, "Activation functions for convolutional neural networks: proposals and experimental study," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 3, pp. 1478–1488, Mar. 2023, doi: 10.1109/TNNLS.2021.3105444.
[38] J. A. Hanley and B. J. McNeil, "A method of comparing the areas under receiver operating characteristic curves derived from the same cases," *Radiology*, vol. 148, no. 3, pp. 839–843, 1983, doi: 10.1148/radiology.148.3.6878708.
[39] C. Zhang, B. Recht, S. Bengio, M. Hardt, and O. Vinyals, "Understanding deep learning requires rethinking generalization," in *5th International Conference on Learning Representations, ICLR 2017-Conference Track Proceedings*, 2017.
[40] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," *BJU international*, vol. 101, no. 1, pp. 1396–1400, 2008.
[41] K. Oyebode, "Kidney stone detection note- for research purpose only-," *Hugging Face-Spaces*. [Online]. Available: https://huggingface.co/spaces/Kazeemkz/Kidneystone_detection

BIOGRAPHIES OF AUTHORS

Kazeem Oyebode holds a Ph.D. in computer engineering with a specialization in computer vision. He also received a Master of Science (M.Sc.) in software engineering from the University of the West of England. He completed his postdoctoral research at Tshwane University of Technology, South Africa, in 2018. He has developed solutions in the field of computer vision for seismic volume analysis and security applications. He teaches at the School of Science and Technology at Pan-Atlantic University, Lagos, Nigeria, where he also conducts research. His research areas of interest include computer vision and image processing. He can be contacted at email: koyebode@pau.edu.ng.

Anne Ngozi Odoh holds a Bachelor of Science (B.Sc.) in computer science from the University of Benin, Nigeria. She also received a Master of Business Administration (MBA) and a Doctorate in Business Administration from the Lagos Business School and the University of Bradford, United Kingdom respectively. She is currently is a senior faculty member at the School of Media and Communication, Pan-Atlantic University in Lagos, Nigeria, where she facilitates courses in digital business and technology management, new media and marketing communication at the graduate and undergraduate levels. Her research interests include digital transformation, artificial intelligence, and information systems. She can be contacted at email: aodoh@pau.edu.ng.