**RESEARCH**                                                                                                    **Open Access**

# DeepVAQ : an adaptive deep learning for prediction of vascular access quality in hemodialysis patients

Sarayut Julkaew[1], Thakerng Wongsirichot[2*], Kasikrit Damkliang[2] and Pornpen Sangthawan[3]

## Abstract

**Background**  Chronic kidney disease is a prevalent global health issue, particularly in advanced stages requiring dialysis. Vascular access (VA) quality is crucial for the well-being of hemodialysis (HD) patients, ensuring optimal blood transfer through a dialyzer machine. The ultrasound dilution technique (UDT) is used as the gold standard for assessing VA quality; however, its limited availability due to high costs impedes its widespread adoption. We aimed to develop a novel deep learning model specifically designed to predict VA quality from Photoplethysmography (PPG) sensors.

**Methods**  Clinical data were retrospectively gathered from 398 HD patients, spanning from February 2021 to February 2022. The DeepVAQ model leverages a convolutional neural network (CNN) to process PPG sensor data, pinpointing specific frequencies and patterns that are indicative of VA quality. Meticulous training and fine-tuning were applied to ensure the model's accuracy and reliability. Validation of the DeepVAQ model was carried out against established diagnostic standards using key performance metrics, including accuracy, specificity, precision, F-score, and area under the receiver operating characteristic curve (AUC).

**Result**  DeepVAQ demonstrated superior performance, achieving an accuracy of 0.9213 and a specificity of 0.9614. Its precision and F-score stood at 0.8762 and 0.8364, respectively, with an AUC of 0.8605. In contrast, traditional models like Decision Tree, Naive Bayes, and kNN demonstrated significantly lower performance across these metrics. This comparison underscores DeepVAQ's enhanced capability in accurately predicting VA quality compared to existing methodologies.

**Conclusion**  Exemplifying the potential of artificial intelligence in healthcare, particularly in the realm of deep learning, DeepVAQ represents a significant advancement in non-invasive diagnostics. Its precise multi-class classification ability for VA quality in hemodialysis patients holds substantial promise for improving patient outcomes, potentially leading to a reduction in mortality rates.

**Keywords**  Deep learning, Hemodialysis, Vascular access

*Correspondence:
Thakerng Wongsirichot
thakerng.w@psu.ac.th
[1] College of Digital Science, Prince of Songkla University, Hat Yai, Songkhla, Thailand
[2] Division of Computational Science, Faculty of Science, Prince of Songkla University, Hat Yai, Songkhla, Thailand
[3] Department of Medicine, Division of Nephrology, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla, Thailand

Julkaew *et al. BMC Medical Informatics and Decision Making*        (2024) 24:45

Page 2 of 11

## Background

The number of end-stage kidney disease (ESKD) patients on hemodialysis (HD) has been rapidly increasing worldwide and is accompanied by the high burden of HD vascular access (VA)-related complications. Globally, the HD patients comprise approximately 69% of all individuals receiving renal replacement therapy and 89% of those undergoing dialysis procedures [1]. In Southeast Asia, the prevalence rate (incidence per million persons) of HD patients is approximately 473.3 compared with higher and lower income countries at 305.8 and 167.5, respectively. The prevalence rates are expected to increase by 10% over five years [2, 3]. VA serves as a critical lifeline for HD patients, providing the necessary means for their dialysis treatment [4]. There are two commonly used types of VAs: AVF (Arteriovenous Fistula) and AVG (Arteriovenous Graft). AVF involves directly connecting an artery and a vein, while AVG utilizes a synthetic graft to establish the connection. Maintaining a properly functioning VA is of utmost importance for effective HD management. Any malfunction in the VA can lead to dialysis insufficiency and significantly increase the morbidity and mortality risks for HD patients. Stenosis and thrombosis are the primary causes of VA dysfunction, necessitating timely detection and intervention to ensure the well-being and survival of HD patients [5].

The prevalent method for detecting VA stenosis and thrombosis in HD patients involves measuring VA blood flow via ultrasound dilution (VABF-UD), a technique considered the gold standard in HD treatment. This method has notably improved patient care by facilitating timely interventions, maintaining VA functionality, and mitigating risks associated with inadequate dialysis [6, 7]. However, the broader application of VABF-UD is constrained by factors such as cost and limited accessibility, particularly in developing or under-developed countries [8]. These barriers restrict its integration into routine clinical practice, limiting its benefits to a broader patient population.

In this study, we propose the utilization of low-cost PPG sensors, known for their effectiveness in capturing biosignals, combined with a deep learning model based on fine-tuned parameters in a convolutional neural network (CNN) architecture called "DeepVAQ". The resulting DeepVAQ model demonstrates remarkable accuracy in early detection and prediction of VA quality in HD patients, offering a promising and effective approach for forecasting VA quality. This study contributes to the development of screening tests specifically tailored to evaluate VA quality in HD patients, enhancing early detection, patient care, and treatment outcomes.

## Literature review

Recent advancements in monitoring health conditions and diagnosing diseases have increasingly capitalized on the integration of sensor technology with machine learning. Prominent among these technologies are Photoplethysmography (PPG) sensors, non-invasive devices that monitor blood volume variations in the vascular system [9]. These sensors function by detecting changes in light absorption or reflection, which facilitates the estimation of essential physiological parameters such as pulse rate, blood flow, and oxygen saturation. These parameters are crucial for non-invasive health monitoring [10, 11]. The application of wearable PPG sensors in healthcare has been extensively documented [12], particularly their utility in monitoring blood circulation changes, exemplified by their use in continuous heart rate monitoring. These sensors are also promising for the early detection of cardiovascular diseases and for real-time monitoring in clinical environments. The versatility of PPG sensors across various healthcare applications, including their use in vascular occlusion training, is well-established [13, 14]. Advances in PPG signal analysis for biomedical purposes have seen considerable progress, especially with the incorporation of sophisticated analytical methods and various sensor types. Deep learning algorithms, particularly the CNN-LSTM model applied to PPG signals, have shown to outperform other algorithms [15–17]. This highlights the potential for further research into model architectures, hyperparameters, and time–frequency representations to refine PPG signal analysis. It is imperative to note the unique nature of vascular access in HD patients, which demands specialized consideration compared to other patient groups. While PPG sensors and deep learning techniques hold significant promise in healthcare, targeted research on their efficacy in evaluating VA quality for HD patients is scant. Advancing this line of inquiry is vital to harness the full capabilities of these technologies in enhancing HD patient care.

## Methods

The experiment comprised two main sections: data collection (a) and classification (b) as shown in Fig. 1.

Data Collection (a): Data were gathered from 398 HD patients with both types of VA, Arteriovenous Fistula (AVF) and Arteriovenous Graft (AVG). The patient cohort included 246 males (61.81%) and 152 females (38.19%), with ages ranging from 29 to 90 years and an average age of 63 years. Among these patients, 155 (38.94%) had an AVG and 243 (61.06%) had an AVF. The VA locations were on the left arm for 276 patients (75.20%) and on the right arm for 91 patients (24.80%), with 141 (36.34%) on the upper arm and 247 (63.66%) on the forearm. This demographic
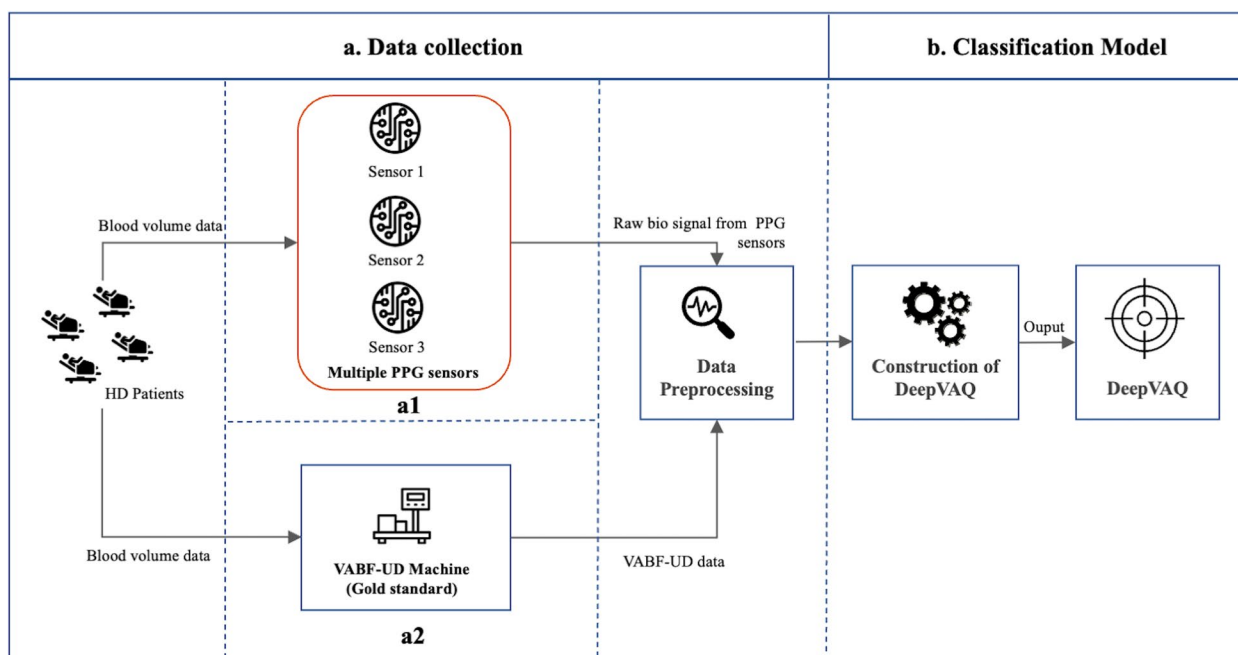
**Fig. 1** Research experimental design

**Table 1** Demographic and clinical characteristics of HD patients

| Variable | Category | HD patients (*n*) | HD patients (%) |
|---|---|---|---|
| Gender | Male | 246 | 61.81% |
|  | Female | 152 | 38.19% |
| Age | Maximum | 90 |  |
|  | Minimum | 29 |  |
|  | Average | 63 |  |
| Type of Vascular Access | AVG | 155 | 38.94% |
|  | AVF | 243 | 61.06% |
| Arm | Left | 276 | 75.20% |
|  | Right | 91 | 24.80% |
| Location on arm | Upper arm | 141 | 36.34% |
|  | Forearm | 247 | 63.66% |

and clinical information is detailed in Table 1. The data collection involved two sources: multiple PPG sensors (a1) and the VABF-UD machine (a2). For the PPG data collection, we employed the SEN0203 model PPG sensors [18], equipped with the SON1303 IC/Module. Three PPG sensors were strategically placed over each patient's VA site on the skin by expert staff specializing in VA, to ensure accurate positioning and data collection. This placement was standardized to ensure consistent data quality, with sensors attached using pads, and the skin area was cleaned and prepared. The PPG

sensors recorded signals for a continuous period of 5 min before the patients' dialysis sessions. Concurrently, VABF data (ml/min) were collected from the patients using a device manufactured by Transonic Systems Inc [19] during their dialysis procedures.

Classification Model (b): In this section, DeepVAQ was constructed by conducting a series of experiments with different parameter settings. Each experiment underwent comprehensive evaluation using multiple performance measurements to identify the optimal parameters for the proposed model.

In previous studies [20, 21], researchers have established optimal criteria for detecting stenosis and thrombosis. Among the various measurement techniques, the VABF (vascular access blood flow) standard has been identified as the most effective predictor of impending stenosis and thrombosis. We utilized a combination of the aforementioned VABF criteria to classify the dataset and predict the quality of VA. Through the use of statistical methods [22], we determined the range of VABF rates associated with different VA quality classes, which are summarized in Table 2.

The classification allows for the identification of VA conditions associated with a high chance of stenosis and thrombosis (Class 1), as well as those prone to such complications (Class 2). Additionally, it distinguishes between VA with good (Class 3), high (Class 4), and excellent (Class 5) quality.

**Table 2** Classification of VA Quality based on VABF Rate

| Class | VA quality | VABF rate Minimum (ml/min) | VABF rate Maximum (ml/min) | Description |
|---|---|---|---|---|
| 1 | Poor | 0 | 450 | High chance of VA stenosis and thrombosis |
| 2 | Below Medium | 451 | 900 | Prone to VA stenosis and thrombosis |
| 3 | Medium | 901 | 1,350 | Good VA quality |
| 4 | Good | 1,351 | 1,800 | High VA Quality |
| 5 | Excellent | $\geq 1,801$ | | Highest VA Quality |

### Datasets and preprocessing

According to Fig. 2, PPG sensors were utilized to gather data from HD patients, resulting in a total of 27,000 measurements (9,000 values per sensor) in a 1D-formatted data representation. This corresponds to an approximate sampling rate of 30 values per second. To ensure consistent scaling of the features, the raw dataset was normalized using the min–max method, which rescales the values between 0 and 1 [23]. The normalization process is governed by Eq. (1):

$$x' = \frac{x_i - x_{mn}}{x_m - x_{mn}} \tag{1}$$

where $X_{mn}$ denotes minimum value, $X_m$ denotes maximum value, $X_i$ denotes input value, and $X'$ denotes normalized data.

To mitigate noise from the PPG sensors, we employed a Kalman filter, which utilizes a linear estimator to estimate the value of an unknown variable over time by considering previous data and weighing it against the actual data [24]. Applying the Kalman filter significantly enhances the accuracy and reliability of the PPG sensor data.

To facilitate the classification process, the PPG sensor data was assigned labels based on five distinct classes corresponding to the range of VA quality, as presented in Table 2. These classes serve as reference points for categorizing the VA quality based on the specific range of values associated with the PPG sensor measurements.

The collected dataset initially had an imbalance, which could introduce bias and impact prediction accuracy. To address this issue and avoid overfitting, we employed the Synthetic Minority Oversampling Technique (SMOTE) [25]. SMOTE, a widely used method, creates new synthetic samples for the minority class by utilizing a k-nearest neighbor algorithm. By augmenting the representation of the minority class with these synthetic samples, we aimed to rebalance the dataset and improve the model's predictive capabilities for both majority and minority classes. This adjustment through SMOTE ensured a more balanced representation of each class in the dataset, reducing bias and enhancing overall performance. As a result, all classes consisted of 125 samples, resulting in a total of 625 samples in the
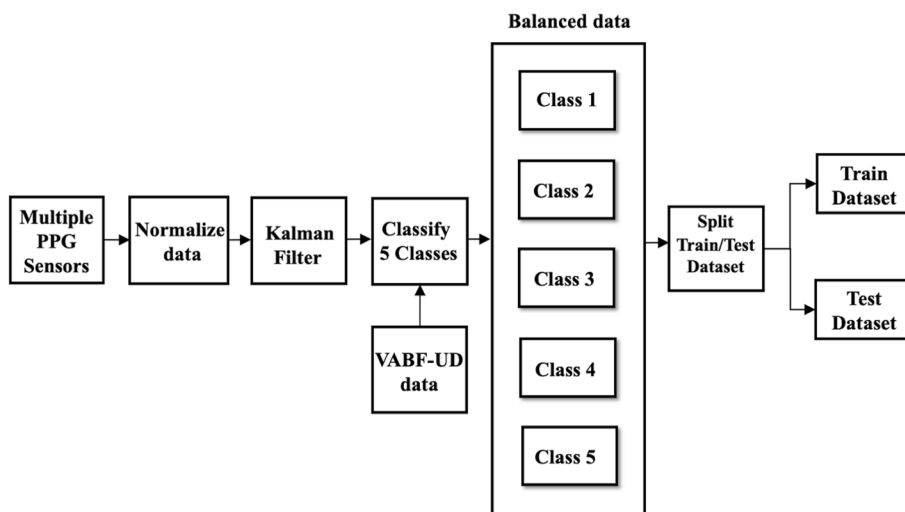


**Fig. 2** Dataset and preprocessing

**Table 3** Classes distribution before and after SMOTE sampling

| Class | VA quality | Number of samples | |
|---|---|---|---|
| | | **Before SMOTE** | **After SMOTE** |
| 1 | Poor | 22 | 125 |
| 2 | Below Medium | 125 | 125 |
| 3 | Medium | 108 | 125 |
| 4 | Good | 55 | 125 |
| 5 | Excellent | 28 | 125 |

dataset. Table 3 presents the number of samples for each class before and after the application of SMOTE.

We further partitioned the dataset into training and testing sets using a 70:30 ratio. The training set, comprising 70% of the data, was used for model training and parameter optimization. The remaining 30% of the data was reserved for evaluating the trained model's performance on unseen samples.

Table 4 illustrates the distinct waveforms of VA quality data captured by the PPG sensor. The waveforms exhibit intricate and diverse patterns within a frequency range of 0–300. These patterns vary significantly across different classes, highlighting the complexity and uniqueness of each VA quality category.

**DeepVAQ model construction**

The construction of the DeepVAQ model employed in this study can be described as a highly sophisticated and intricate model as show in Fig. 3, surpassing the simplicity of CNN architectures. It stands out due to its extensive complexity and advanced design elements. One notable aspect is the utilization of two layers of 1D-CNN, which significantly enhances its capacity to extract meaningful features which are then converted into vectors based on Eq. (2). By incorporating multiple filters such as {8, 16, 32} and varying kernel sizes {3, 5, 7} [26, 27], the model becomes capable of capturing a wide range of intricate patterns and nuanced details present in the data.

**Table 4** Waveform of VA quality data from PPG sensor

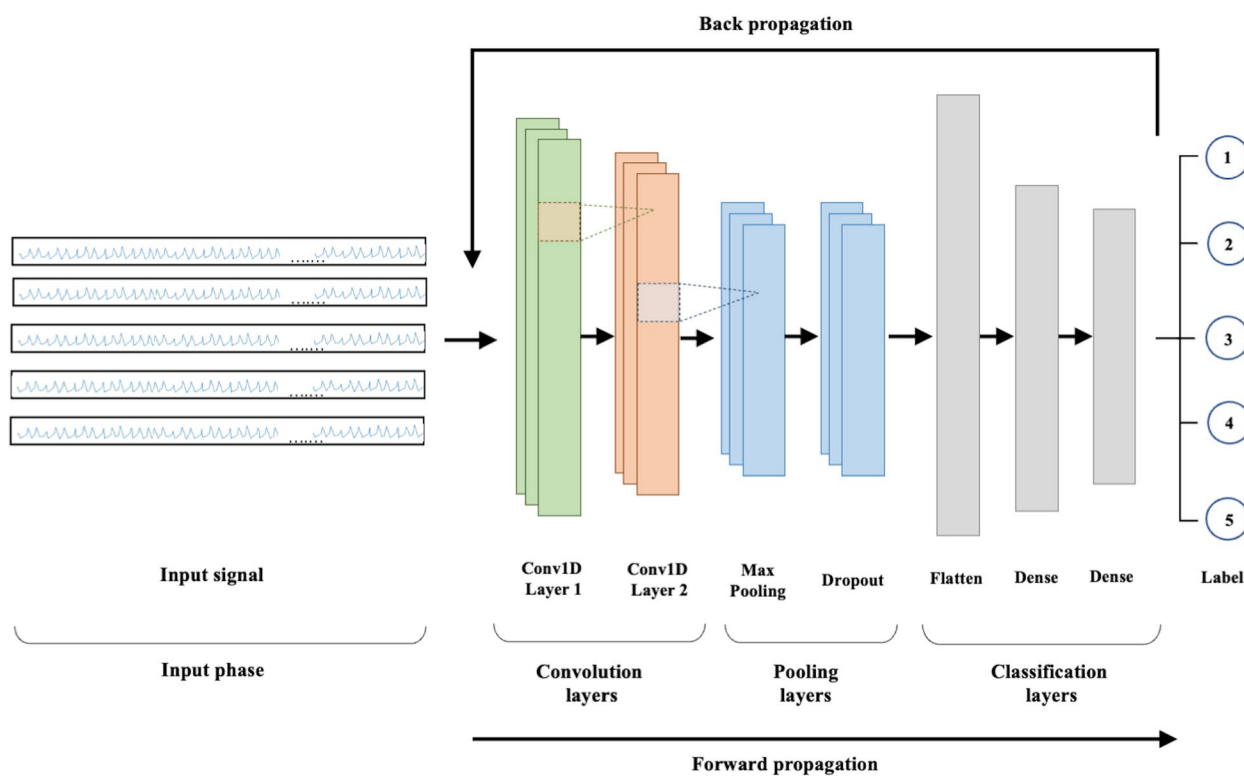| Class | VA quality | Waveform |
|---|---|---|
| 1 | Poor |  |
| 2 | Below Medium |  |
| 3 | Medium |  |
| 4 | Good |  |
| 5 | Excellent |  |

**Fig. 3** Convolution neural network architecture

$$y(n) = \begin{cases} \sum_{i=0}^{k} x(n+i)h(i) & if n = 0. \\ \sum_{i=0}^{k} x(n+i+(s-1))h(i), & otherwise. \end{cases}$$

(2)

Where x is the input to the convolution layer of length n, h the kernel of length k, and s represents the kernel window shift positions (number of strides) after each convolution.

The DeepVAQ model's sophistication is enhanced through customized parameter tuning. The selection of appropriate filters and kernel sizes optimizes feature extraction, which is crucial for discerning and classifying diverse data patterns. The incorporation of the Leaky ReLU activation function [28] introduces a small negative slope for negative inputs, thereby improving the model's ability to handle complex data patterns. The model is further optimized using the Adam optimizer [29]. This algorithm adapts the learning rate based on the gradient of the loss function, promoting faster convergence and improved performance. The dynamic adjustment of the learning rate by the Adam optimizer aids in efficiently navigating the loss landscape, optimizing parameter values for better training outcomes.

Within its convolutional architecture, DeepVAQ incorporates max pooling and dropout layers. Max pooling, with a $2 \times 2$ size, reduces the dimensionality of the feature maps, focusing on important features. The inclusion of a dropout layer, set at a rate of 0.5, acts as a regularization technique. It prevents overfitting by randomly dropping neuron outputs during training, thereby enhancing the model's generalization and adaptability to varied datasets.

The classification layers of the model are instrumental for multi-class classification. These layers consolidate feature vectors from preceding layers into a synthesized form for processing in a dense layer, which is pivotal for final classification. The use of the one-hot encoding technique [30] ensures accurate differentiation among the five VA quality classes.

In the training process, the DeepVAQ model utilizes the cross-entropy loss function [31], a widely-accepted measure for multi-class classification tasks. This loss function quantifies the dissimilarity between the predicted class probabilities and the actual ground truth labels, serving as a crucial metric for evaluating the model's performance and guiding the optimization process. The number of training epochs is strategically set to balance the model's complexity and efficiency.

**DeepVAQ model evaluation**

The DeepVAQ model utilized a ten-fold cross-validation approach on the test set to validate its performance across multiple iterations, enhancing the reliability of the

results. The evaluation metrics included accuracy, sensitivity, specificity, precision, and F-score, which are commonly used classification performance measurements [32]. These metrics provide a comprehensive assessment of the model's predictive capabilities and its ability to correctly classify different VA quality classes. The evaluation was conducted on the test set, and the mean and standard deviation of the performance measurements are presented in Table 5.

The development of the DeepVAQ model was built using the Keras high-level API [33] in Python, providing a powerful and user-friendly framework for deep learning. The experiments were conducted on a robust computational infrastructure, featuring an Intel Core i7-7700 processor with a clock speed of 3.60 GHz, 16-GB DDR4 RAM, and a 512-GB solid-state drive (SSD). Complementing the hardware, an NVIDIA Quadro-620 GPU with 2-GB GDDR5 memory was employed for efficient processing of the complex computations involved in model training and evaluation.

## Results

The DeepVAQ model was evaluated through a series of experiments, where different parameter settings, filter sizes, and kernel sizes were tested. In each experiment, the filter sizes (i) were set to 8, 16, and 32, while

the kernel sizes (j) were set to 3, 5, and 7. The results of these experiments are presented in Table 6. Among all the experiments, exp((16,5)) achieved the highest average performance across all measurements, with an accuracy of $0.9106 \pm 0.08$, sensitivity of $0.7768 \pm 0.17$, specificity of $0.9441 \pm 0.07$, precision of $0.7984 \pm 0.22$, F-Score of $0.7829 \pm 0.18$, and AUC of $0.8605 \pm 0.11$. A comprehensive overview of the performance stands out as the best performing experiment, demonstrating the effectiveness of the DeepVAQ model in accurately classifying VA quality.

Figure 4 illustrates the relationship between the filter sizes and the validation accuracy and loss. It can be observed that as the filter size increases, the validation accuracy improves, reaching a value of 0.8, while the validation loss decreases to 0.7.

The performance of the DeepVAQ model at the configured parameter setting of $\exp_{(16,5)}$ was evaluated for each VA quality class. The results, summarized in Table 7, demonstrate the accuracy, sensitivity, specificity, precision, and F-Score achieved by the model for each class. For class 1, the model achieved an accuracy of 0.9894, high sensitivity (0.9474) and specificity (1.0000), precise predictions (precision of 1.0000), and a balanced F-Score of 0.9730. Class 4 also had notable results, with an accuracy of 0.9574, sensitivity of 0.8421, specificity of 0.9867, precision of 0.9412, and an F-Score of 0.8889. Additionally, class 5 exhibited exceptional accuracy of 0.9947, sensitivity of 0.973, specificity of 1.0000, precision of 1.0000, and an F-Score of 0.9863. Although classes 2 and 3 had relatively lower accuracies (0.8404 and 0.8245, respectively), the model still demonstrated reasonable performance in differentiating these classes.

To assess the performance of DeepVAQ, we utilized ROC curves and calculated the AUC (Area Under the Curve). Figure 5 illustrates the ROC curve for the DeepVAQ model across multiple classifications. The AUC

**Table 5** Performance measurements

| Performance measurements | Formula |
|---|---|
| Accuracy | $(TP+TN)/(TP+FP+TN+FN)$ |
| Sensitivity | $TP/(TP+FN)$ |
| Specificity | $TN/(TN+FP)$ |
| Precision | $TP/(TP+FP)$ |
| F-Measure | $(2 \times Precision \times Recall)/$ $(Precision + Recall)$ |

*TP* True Positive, *TN* True Negative, *FP* False Positive, *FN* False Negative

**Table 6** Performance measurements for different experiments

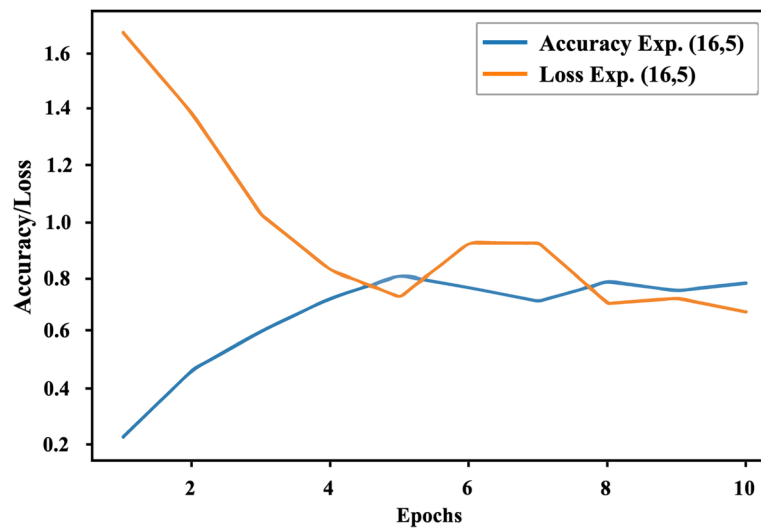| Experiment $\exp_{(i,j)}$ | Accuracy | Sensitivity | Specificity | Precision | F-Score | AUC | Loss validation |
|---|---|---|---|---|---|---|---|
| $\exp_{(8,3)}$ | 0.9021 +—0.09 | 0.7558 +—0.20 | 0.9388 +—0.07 | 0.7698 +—0.24 | 0.7610 +—0.22 | 0.8473 +—0.13 | 1.2521 +—1.02 |
| $\exp_{(8,5)}$ | 0.8915 +—0.11 | 0.7267 +—0.37 | 0.9320 +—0.13 | 0.7391 +—0.28 | 0.6996 +—0.35 | 0.8294 +—0.18 | 1.4008 +—1.90 |
| $\exp_{(8,7)}$ | 0.8851 +—0.12 | 0.7111 +—0.36 | 0.9280 +—0.15 | 0.7957 +—0.29 | 0.6939 +—0.33 | 0.8196 +—0.18 | 1.4810 +—1.85 |
| $\exp_{(16,3)}$ | 0.8745 +—0.07 | 0.6863 +—0.27 | 0.9215 +—0.03 | 0.6681 +—0.17 | 0.6704 +—0.22 | 0.8039 +—0.14 | 1.6079 +—1.40 |
| $\exp_{(16,5)}$[a] | 0.9106 +—0.08[a] | 0.7768 +—0.17[a] | 0.9441 +—0.07[a] | 0.7984 +—0.22[a] | 0.7829 +—0.18[a] | 0.8605 +—0.11[a] | 1.1441 +—0.86[a] |
| $\exp_{(16,7)}$ | 0.9000 +—0.10 | 0.7512 +—0.25 | 0.9376 +—0.08 | 0.7655 +—0.25 | 0.7490 +—0.25 | 0.8444 +—0.15 | 1.2754 +—1.30 |
| $\exp_{(32,3)}$ | 0.8830 +—0.11 | 0.7077 +—0.23 | 0.9269 +—0.07 | 0.7287 +—0.28 | 0.7169 +—0.26 | 0.8173 +—0.16 | 1.4985 +—1.19 |
| $\exp_{(32,5)}$ | 0.8915 +—0.11 | 0.7300 +—0.25 | 0.9323 +—0.09 | 0.7582 +—0.29 | 0.7335 +—0.26 | 0.8311 +—0.15 | 1.3840 +—1.30 |
| $\exp_{(32,7)}$ | 0.9021 +—0.09 | 0.7560 +—0.19 | 0.9388 +—0.07 | 0.7794 +—0.25 | 0.7635 +—0.21 | 0.8474 +—0.13 | 1.2506 +—0.98 |

[a] Best classification result

**Fig. 4** Validation accuracy and loss in $exp_{(16,5)}$

**Table 7** Performance measurements for DeepVAQ model at $exp_{(16,5)}$

| Approach | Class | Accuracy | Sensitivity | Specificity | Precision | F-Score |
|---|---|---|---|---|---|---|
| **DeepVAQ model** | 1 | 0.9894 | 0.9474 | 1.0000 | 1.0000 | 0.9730 |
| | 2 | 0.8404 | 0.5405 | 0.9139 | 0.6061 | 0.5714 |
| | 3 | 0.8245 | 0.7105 | 0.8533 | 0.5510 | 0.6207 |
| | 4 | 0.9574 | 0.8421 | 0.9867 | 0.9412 | 0.8889 |
| | 5 | 0.9947 | 0.973 | 1.0000 | 1.0000 | 0.9863 |



**Fig. 5** ROC curve for DeepVAQ

value provides an indication of the model's predictive capability, with a higher value suggesting better prediction accuracy. In our evaluation, each class achieved a true positive rate exceeding 0.5. The position of the ROC curve closest to the upper-left corner represents a balance between high sensitivity and specificity, indicating an optimal classification performance. This effectively distinguished between different VA quality classes, demonstrating its ability to accurately identify true positives and true negatives.

The confusion matrix, displayed in Fig. 6, provides a comprehensive assessment of the proposed DeepVAQ model. It offers valuable insights into the model's performance on the testing data, consisting of 125 records in the test set. The predictions obtained on the test set closely align with the training phase results, highlighting the model's consistency.

Figure 6 illustrates the confusion matrix, where each column represents the assigned VA quality labels, and each row corresponds to the true values. The highlighted colors within the matrix signify correct predictions, demonstrating the model's accuracy in classifying the VA quality classes. Notably, class 1 and class 5 achieved flawless predictions, with all samples correctly identified.

The demonstrates high accuracy in predicting VA quality, particularly in classifying VA quality class 1 as "Poor" and class 5 as "Excellence". Class 1, representing the "Poor" state, indicates a high chance of VA stenosis and thrombosis, making it clinically significant. The model's accuracy in class 5 reflects its ability to identify the highest VA quality. These findings highlight the DeepVAQ

model's effectiveness in distinguishing between different levels of VA quality, enabling clinicians to make informed decisions based on the predicted VA quality. The model shows promise in identifying patients at higher risk and assessing overall VA quality.

## Discussion

The empirical evidence presented in this study confirms the DeepVAQ model's exceptional capability in classifying VA quality in HD patients. Our comprehensive comparative analysis, detailed in Table 8, demonstrates DeepVAQ's superior performance metrics against established machine learning models, including Decision Tree [34], Naive Bayes [35], Support Vector Machine (SVM) [36], and k-Nearest Neighbors (kNN) [37]. These findings are not merely academic but carry profound implications for clinical practice. DeepVAQ's high accuracy and precision in non-invasive VA quality monitoring can potentially transform patient management, reducing the reliance on invasive procedures and facilitating proactive healthcare strategies.

Acknowledging the limitations of our research, we recognize that the controlled test conditions of our study may not fully encapsulate the complexities of real-world clinical environments. Future studies are required to validate DeepVAQ in diverse clinical settings, ensuring its efficacy across a broader patient population. Additionally, while our model exhibits robustness in classification, the overlap observed between classes 2 and 3 necessitates further refinement of the model's discriminative power. Advanced noise reduction and feature selection
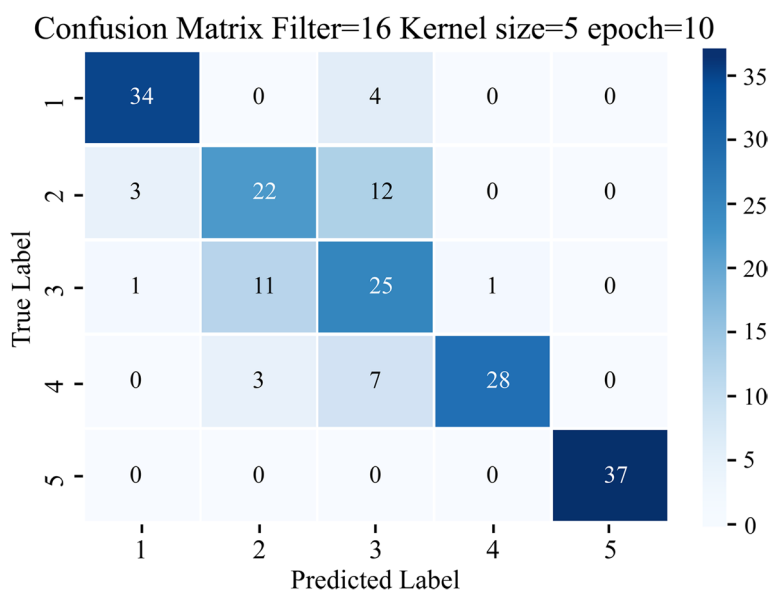


**Fig. 6** Confusion matrix of DeepVAQ for each class of HD patients

**Table 8** A comparison of prediction results of models

| Models | Accuracy | Sensitivity | Specificity | Precision | F-Score |
|---|---|---|---|---|---|
| Decision Tree [34] | 0.2531 | 0.1563 | 0.6550 | 0.0679 | 0.3051 |
| Naive Bayes [35] | 0.5816 | 0.3041 | 0.7159 | 0.4348 | 0.2489 |
| k-Nearest-Neighbours (*k*NN) [37] | 0.4735 | 0.2492 | 0.6912 | 0.3237 | 0.1230 |
| DeepVAQ model (A proposed model) | 0.9213[a] | 0.8431[a] | 0.9614[a] | 0.8762[a] | 0.8364[a] |

[a] Best classification result

techniques, such as Recursive Least Square (RLS) and Least Mean Square Error (LMS), will be pivotal in addressing this challenge.

The potential for AI and machine learning to revolutionize the field of HD patient care is unmistakable. DeepVAQ stands as a testament to this potential, paving the way for integrating such models into routine clinical practice, enhancing patient-centered care, and setting a new benchmark in the management of vascular access.

## Conclusions

The proposed DeepVAQ model demonstrates successful classification of VA quality in HD patients by utilizing multiple PPG sensors. It was developed using a customized hyperparameter setting within the CNN architecture, DeepVAQ achieved superior classification performance when compared to other machine learning models, as evidenced by its highest scores across various performance measurements. This approach offers a non-invasive and cost-effective means of predicting VA quality, which could potentially be accessible to the wider public. The findings highlight the potential of DeepVAQ as a promising tool for early detection of VA quality deterioration in HD patients. By accurately classifying VA quality, this model has the potential to greatly benefit the quality of life for HD patients worldwide. While DeepVAQ performed exceptionally well overall, further enhancements can be explored to improve its ability to differentiate between closely related classes, specifically classes where some overlap was observed.

## Abbreviations

| | |
|---|---|
| VA | Vascular access |
| ESKD | End-stage kidney disease |
| HD | Hemodialysis |
| VABF-UD | VA blood flow using ultrasound dilution |
| PPG | Photopletysmography |
| CNN | Convolutional neural network |
| DeepVAQ | Deep learning for prediction of vascular access quality in hemodialysis patients |

## Authors' contributions
S. J., T. W., K. D., and P. S. contributed to conceptualization. S. J. and T. W. contributed to data validation. S. J., T. W., and P. S. contributed to resources. S. J., T. W., and P. S. contributed to data curation. S. J. and T. W. contributed to manuscript writing. S. J., T. W., K. D., and P. S. contributed to visualization. T. W. was the supervisor and project administrator. All authors have read and approved the final manuscript.

## Availability of data and materials
The datasets analyzed during the current study is available from the corresponding author on reasonable request.

## Declarations

### Ethics approval and consent to participate
This study protocol was approved by the Ethics Committee of Faculty of Medicine, Prince of Songkla University, Thailand (Certificate of Approval No. REC.63–411-14–1). Consent to participate: Written informed consent was obtained from all participants in this study. All methods were carried out in accordance with relevant guidelines and regulations.

### Consent for publication
As our manuscript does not include any individual data or sensitive personal information, therefore consent for publication is "Not Applicable" in this case.

### Competing interests
The authors declare no competing interests.

## References
1. Bello AK, et al. Epidemiology of haemodialysis outcomes. Nat Rev Nephrol. 2022;18(6):378–95.
2. Chapter 1: Definition and classification of CKD. Kidney Int suppl. 2013;3(1):19–62. https://doi.org/10.1038/kisup.2012.64.
3. Alexander S, et al. Impact of national economy and policies on end-stage kidney care in South Asia and Southeast Asia. Int J Nephrol. 2021;2021:6665901.

4.  Yeh L-M, Chiu SY-H, Lai P-C. The impact of vascular access types on hemodialysis patient long-term survival. Sci Rep. 2019;9(1):10708.
5.  Hung TW, et al. Association of permanent vascular access dysfunction with subsequent risk of cardiovascular disease: a population-based cohort study. J Pers Med. 2022;12(4):598.
6.  Pedro P, Ana M. In: David CG, Aleksandar K, editors. Early Detection and Endovascular Intervention to Correct Dialysis Vascular Access Malfunction, in Cardiac Diseases. Rijeka: IntechOpen; 2020. p. Ch. 11.
7.  Lok CE, et al. KDOQI clinical practice guideline for vascular access: 2019 update. Am J Kidney Dis. 2020;75(4):S1–164.
8.  Prasad N, Jha V. Hemodialysis in Asia. Kidney Dis (Basel). 2015;1(3):165–77.
9.  Wang D, et al. Photoplethysmography-based blood pressure estimation combining filter-wrapper collaborated feature selection with LASSO-LSTM model. IEEE Trans Instrum Meas. 2021;70:1–14.
10. Schrumpf F, et al. Assessment of Non-Invasive Blood Pressure Prediction from PPG and rPPG Signals Using Deep Learning. Sensors. 2021;21(18):6022.
11. Schrumpf F, Frenzel P, Aust C, Osterhoff G, Fuchs M. Assessment of deep learning based blood pressure prediction from PPG and rPPG signals. Nashville: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2021. pp. 3815–3825. https://doi.org/10.1109/CVPRW53098.2021.00423.
12. Castaneda D, et al. A review on wearable photoplethysmography sensors and their potential future applications in health care. Int J Biosens Bioelectron. 2018;4(4):195–202.
13. Moraes JL, et al. Advances in Photopletysmography Signal Analysis for Biomedical Applications. Sensors. 2018;18(6):1894.
14. Pantoni C, et al. Study of heart rate autonomic modulation at rest in elderly patients with chronic obstructive pulmonary disease. Revista Brasileira de Fisioterapia, São Carlos, v. 11, n. 1. 2007. p. 33-38. Available from: https://www.researchgate.net/publication/228473777_STUDY_OF_HEART_RATE_AUTONOMIC_MODULATION_AT_REST_IN_ELDERLY_PATIENTS_WITH_CHRONIC_OBSTRUCTIVE_PULMONARY.
15. Esgalhado F, et al. The application of deep learning algorithms for PPG signal processing and classification. Computers. 2021;10(12):158.
16. Tazarv A, Levorato M. A deep learning approach to predict blood pressure from PPG Signals. Annu Int Conf IEEE Eng Med Biol Soc. 2021;2021:5658–62.
17. Bian D, Mehta P, Selvaraj N. Respiratory rate estimation using PPG: a Deep learning approach. Annu Int Conf IEEE Eng Med Biol Soc. 2020;2020:5948–52.
18. DFRobot. Gravity: PPG Heart Rate Monitor Sensor for Arduino (Analog/Digital). 2024; Available from: https://www.dfrobot.com/product-1540.html. [Cited 2024 January 8]
19. Inc., T.S. Best Practices: Transonic® Hemodialysis Surveillance. 2023. [Cited February 2023].
20. Tessitore N, et al. Diagnostic accuracy of ultrasound dilution access blood flow measurement in detecting stenosis and predicting thrombosis in native forearm arteriovenous fistulae for hemodialysis. Am J Kidney Dis. 2003;42(2):331–41.
21. Yang CY, et al. Weight-based assessment of access flow threshold to predict arteriovenous fistula functional patency. Kidney Int Rep. 2022;7(3):507–15.
22. Clapham C, Nicholson J. Ratio scale. In the concise Oxford dictionary of mathematics. Oxford: Oxford University Press; 2009. Available from: https://www.oxfordreference.com/view/10.1093/acref/9780199235940.001.0001/acref-9780199235940-e-2380.
23. Aksu G, Güzeller C, Eser T. The effect of the normalization method used in different sample sizes on the success of artificial neural network model. Int J Assess Tools Educ. 2019;6:170–92.
24. Park S, et al. Measurement noise recommendation for efficient Kalman filtering over a large amount of sensor data. Sensors (Basel). 2019;19(5):1168.
25. Chawla NV, et al. SMOTE: Synthetic Minority Over-sampling Technique. J Artif Intell Res. 2002;16:321–57.
26. Hu Z, et al. Improving Convolutional Neural Network Expression via Difference Exponentially Linear Units. J Phys Conference Series. 2020;1651(1):012163.
27. Ravisutha Sakrepatna S. Understanding 1D Convolutional Neural Networks Using Multiclass Time-Varying Signals. 2018.
28. Agarap AF. Deep learning using rectified linear units (relu). arXiv preprint arXiv. 2018;1803:08375.
29. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv. 2014;1412:6980.
30. Jie L, Jiahao C, Xueqin Z, Yue ZH, Jiajun LI. One-hot encoding and convolutional neural network based anomaly detection. J Tsinghua Univ (Science and Technology). 2019;59(7):523–9.
31. Keras. Probabilistic losses. 2023; Available from: https://keras.io/api/losses/probabilistic_losses/#categorical_crossentropy-function.
32. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol. 2010;5(9):1315–6.
33. Antonio Gulli SP. Deep Learning with Keras. 2017.
34. Jijo B, Mohsin Abdulazeez A. Classification based on decision tree algorithm for machine learning. J Appl Sci Technol Trends. 202AD;2:20–8.
35. Saritas M, Yasar A. Performance analysis of ANN and Naive Bayes Classification Algorithm for Data Classification. Int J Intell Syst Appl. 2019;7:88–91.
36. Srivastava D, Bhambhu L. Data classification using support vector machine. J Theor Appl Inf Technol. 2010;12:1–7.
37. Shokrzade A, et al. A novel extreme learning machine based kNN classification method for dealing with big data. Expert Syst Appl. 2021;183:115293.

## Publisher's Note