CLINICAL KIDNEY JOURNAL

ckj

Clinical Kidney Journal, 2023, vol. 16, no. 3, 560–570

https://doi.org/10.1093/ckj/sfac254
Advance Access Publication Date: 6 December 2022
Original Article

ORIGINAL ARTICLE

A deep learning algorithm to quantify AVF stenosis and predict 6-month primary patency: a pilot study

Jae Hyon Park[1,*], Jongjin Yoon [1,*], Insun Park[2], Yongsik Sim[1], Soo Jin Kim[3], Jong Yun Won[1] and Kichang Han[1]

[1]Department of Radiology, Yonsei University College of Medicine, Seoul, Republic of Korea, [2]Department of Anesthesiology and Pain Medicine, Seoul National University Bundang Hospital, Seongnam, Republic of Korea and [3]Department of Surgery, Yonsei University College of Medicine, Seoul, Republic of Korea

*These authors contributed equally to this work.
Correspondence to: Kichang Han; E-mail: wowsaycheese@yuhs.ac

ABSTRACT

Background. A deep convolutional neural network (DCNN) model that predicts the degree of arteriovenous fistula (AVF) stenosis and 6-month primary patency (PP) based on AVF shunt sounds was developed, and was compared with various machine learning (ML) models trained on patients' clinical data.

Methods. Forty dysfunctional AVF patients were recruited prospectively, and AVF shunt sounds were recorded before and after percutaneous transluminal angioplasty using a wireless stethoscope. The audio files were converted to melspectrograms to predict the degree of AVF stenosis and 6-month PP. The diagnostic performance of the melspectrogram-based DCNN model (ResNet50) was compared with that of other ML models [i.e. logistic regression (LR), decision tree (DT) and support vector machine (SVM)], as well as the DCNN model (ResNet50) trained on patients' clinical data.

Results. Melspectrograms qualitatively reflected the degree of AVF stenosis by exhibiting a greater amplitude at mid-to-high frequency in the systolic phase with a more severe degree of stenosis, corresponding to a high-pitched bruit. The proposed melspectrogram-based DCNN model successfully predicted the degree of AVF stenosis. In predicting the 6-month PP, the area under the receiver operating characteristic curve of the melspectrogram-based DCNN model (ResNet50) (≥0.870) outperformed that of various ML models based on clinical data (LR, 0.783; DT, 0.766; SVM, 0.733) and that of the spiral-matrix DCNN model (0.828).

Conclusion. The proposed melspectrogram-based DCNN model successfully predicted the degree of AVF stenosis and outperformed ML-based clinical models in predicting 6-month PP.

LAY SUMMARY

Auscultation is an effective method of screening arteriovenous fistula (AVF) stenosis through the presence of high-pitched bruit, but it is also subjective and qualitative. This study investigated the feasibility of an auscultation-based deep learning (DL) model for predicting the precise degree of AVF stenosis and 6-month primary patency of AVF. Our auscultation-based DL model successfully predicted the degree of AVF stenosis observed in angiography, and performed better than models constructed using patients' clinical data in predicting 6-month primary patency.

Received: 4.8.2022; Editorial decision: 8.11.2022

© The Author(s) 2022. Published by Oxford University Press on behalf of the ERA. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

560

INTRODUCTION

An arteriovenous fistula (AVF) is the preferred type of access in patients undergoing hemodialysis [1]. Stenosis frequently causes AVF dysfunction and requires repeated percutaneous transluminal balloon angioplasty to establish patency [2, 3]. However, primary patency (PP) rates for AVF are reported to be 72% at 6 months [4] and improved patency has been reported with endovascular intervention.

Auscultation performed using a stethoscope is an effective method of screening for AVF stenosis, as the presence of stenosis can be detected through the sound of abnormal blood flow, referred to as high-pitched bruit. The latest guidelines by the National Kidney Foundation Kidney Disease Outcome Quality Initiative [5] also recommend regular physical examination, which includes auscultation, by health practitioners with moderate quality of evidence.

However, while auscultation is non-invasive, simple and convenient, without the need for expensive equipment or skilled operators as compared with Doppler ultrasound (DUS), diagnosis based on AVF shunt sound can be subjective. Moreover, even a trained practitioner cannot quantify the severity of stenosis based on bruit alone, which makes it difficult to assess whether AVF qualifies for endovascular or surgical treatment based on bruit.

Recently, technical advances have allowed the recording of auscultation using a digital stethoscope by electronic intensification of sound and sharing of recorded sound via Bluetooth transmission [6]. Sound feature extraction, pattern recognition and building of efficient learning pipelines using artificial intelligence (AI) have allowed machine learning (ML) algorithms to detect abnormalities in the sounds of certain organs such as the lungs [7, 8], and these techniques have already been adopted by digital stethoscopes [9, 10].

In regards to AVF, a few studies have explored the use of ML algorithms to analyze AVF shunt sounds [11, 12]. While Ota *et al.* previously classified shunt sounds using a convolutional neural network (CNN), the categories of shunt sounds were not evaluated in relation to the degree of stenosis [11]. Wang *et al.* [12], also developed a model that binomially classifies AVF shunt sounds into stenotic and non-stenotic AVF. However, they used a radial basis function network, the use of which has waned because of the advent of CNN, and evaluated the stenosis using DUS, which is known to be inferior to digital subtraction imaging (DSA) in detecting the stenosis of the venous outflow tract [13]. Overall, no study has fully investigated the performance of the ML algorithm in quantifying the degree of AVF stenosis from shunt sounds or its performance in predicting the 6-month PP.

Previously, we successfully developed an auscultation-based melspectrogram deep CNN (DCNN) model for predicting the presence of significant (≥50%) AVF stenosis [14]. The objective of this pilot study was to evaluate the feasibility of the auscultation-based melspectrogram DCNN model for predicting the precise degree of AVF stenosis and 6-month PP. In addition, performance of the auscultation-based melspectrogram DCNN model in predicting the 6-month PP was compared with that of various ML-based clinical models derived from patients' clinical data.

MATERIALS AND METHODS

Study patients

This single-center, prospective study was approved by the institutional review board of Severance Hospital (2020-2715-009). Patients with dysfunctional autologous AVFs referred for percutaneous transluminal angioplasty (PTA) were assessed for eligibility. Dysfunctional AVF was defined as (i) alterations in the pulse; (ii) failure of the fistula to collapse when the arm is elevated; (iii) inability to achieve the target dialysis blood flow; and (iv) prolonged hemostatic time [5]. Informed consent was obtained from all the patients. The patient inclusion criteria were as follows: (i) presence of native AVF at least 60 days before the procedure that had been used for dialysis for at least 8 of 12 sessions during a 4-week period, ensuring fistula maturity; and (ii) ≥50% stenosis documented on the fistulogram. The exclusion criteria were as follows: (i) thrombosed AVF; and (ii) age below 18 years. Forty end-stage renal disease patients were recruited from November 2020 to August 2021, and study patients were from a previous study [14]. The flowchart of the study is shown in Fig. 1. Information on DSA, PTA and quantification of stenosis in DSA images is provided in the **Supplementary data**.

Recording of AVF shunt sounds

AVF shunt sounds were recorded by placing a wireless electronic stethoscope (Stemoscope, Hulu Devices, San Diego, CA, USA) on top of the venous access 1–2 cm distal to the anastomosis site for 10–15 s. Shunt sounds were recorded before and after PTA, and saved as ".wav" audio-file. A total of 80 AVF shunt sounds (40 before PTA and 40 after PTA) were recorded. The data underlying this article will be shared on reasonable request to the corresponding author.

Data preprocessing and feature extraction

Since the length of each audio file varied owing to manual recording, the audio file was trimmed and padded to a length of 6 s using the Python library *Librosa* [15]. A melspectrogram was used for feature extraction, as it is one of the most widely used methods for audio-data representation [16, 17]. To obtain the melspectrogram, an audio file was first mapped from the time domain to the frequency domain using a short-time Fourier transform (STFT) with a window length of 25 ms and a stride length of 10 ms. The frequency was then converted to mel-scale and amplitude-to-color dimensions by mel-filters to generate a melspectrogram, which represents the short-term power spectrum of the sound. Each melspectrogram was normalized and resized to a resolution of 128 × 128 with three channels on the X-axis, Y-axis and the color representing the time, frequency (Hz) and magnitude of amplitude, respectively.

Data augmentation

The synthetic minority over-sampling technique (SMOTE) algorithm [18] was used to generate synthetic melspectrograms from existing neighboring melspectrograms to best represent
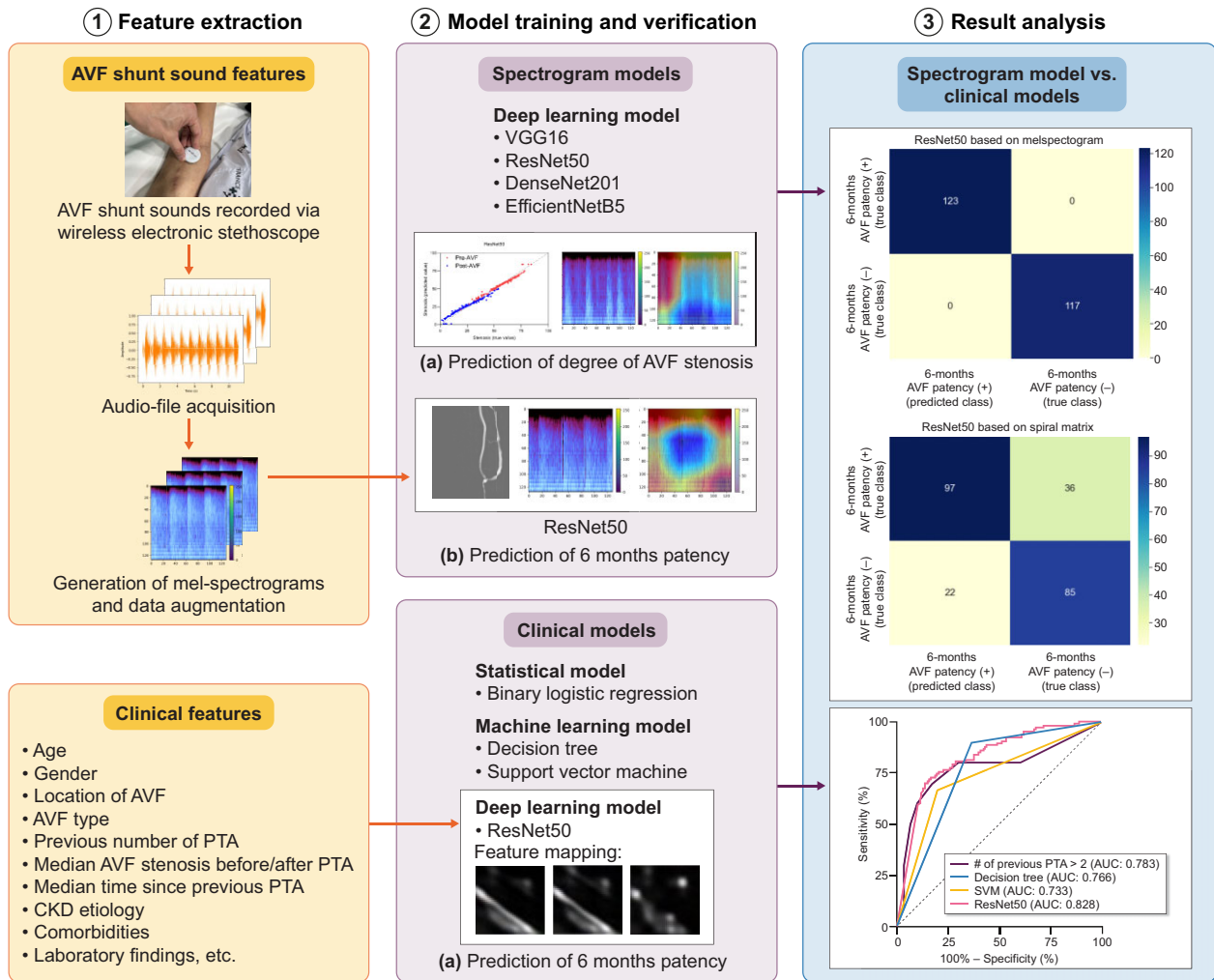
Figure 1: Study flow chart.

real-world data that may be obtained in a clinical setting, which were then augmented or over-sampled 25 times to build the DCNN model to predict the percentage of AVF stenosis. Additionally, they were augmented 60 and 20 times for melspectrograms of patients without and with 6-month PP, respectively, to build the DCNN model predicting the 6-month PP (Fig. 2).

DCNN structure

In this study, we employed widely used CNN architectures, including VGG16 [19], ResNet50 [20], DenseNet201 [21] and EfficientNetB5 [22] to build regression models to predict the degree of AVF stenosis. Two fully connected layers using a rectified linear unit (ReLU) as the activation function were added after the Conv-pool layers with 2048 and 2048 neurons, and two dropout layers (rate = 0.5) were also added for regularization and to avoid overfitting of the model after the first and second dense layers. Finally, for binary classification, a final layer with one neuron, using the softmax activation function, was added. The DCNN models were initialized from the *ImageNet* weights and compiled with mean square error (MSE) as a loss function and a root mean square propagation (RMSprop) [23] optimizer with a learning rate of 0.0001. The models were trained with batch sizes of 24 and 200 epochs, respectively. AVF shunt sounds obtained

before PTA and after PTA were used to construct "pre-AVF" and "post-AVF" models, respectively. Each dataset was randomly divided into training, validation and test sets using split ratios of 70%, 10% and 20%, respectively (Fig. 2**A**). Among the four CNN architectures used to predict the degree of "post-AVF" stenosis, the best CNN architecture was selected for predicting the 6-month PP based on melspectrograms of AVF shunt sounds obtained after PTA. The split ratio used to predict the degree of AVF stenosis was also used to predict the 6-month PP (Fig. 2**B**). Gradient-weighted class activation mapping (Grad-CAM) [24] was used to produce visual explanations for decisions from the DCNN models.

Implementation

All codes were written and run on Google Colab (https://colab.research.google.com, n.d.), which provides 12GB of RAM and an NVIDIA Tesla K80 GPU. Python 3.10.4 was used along with the Python libraries *Numpy*, *Pandas*, *Scikit-learn*, *Tensorflow* and *Keras*.

Data acquisition

To construct clinico-demographic models for predicting 6-month PP based on patients' clinic-demographic and laboratory
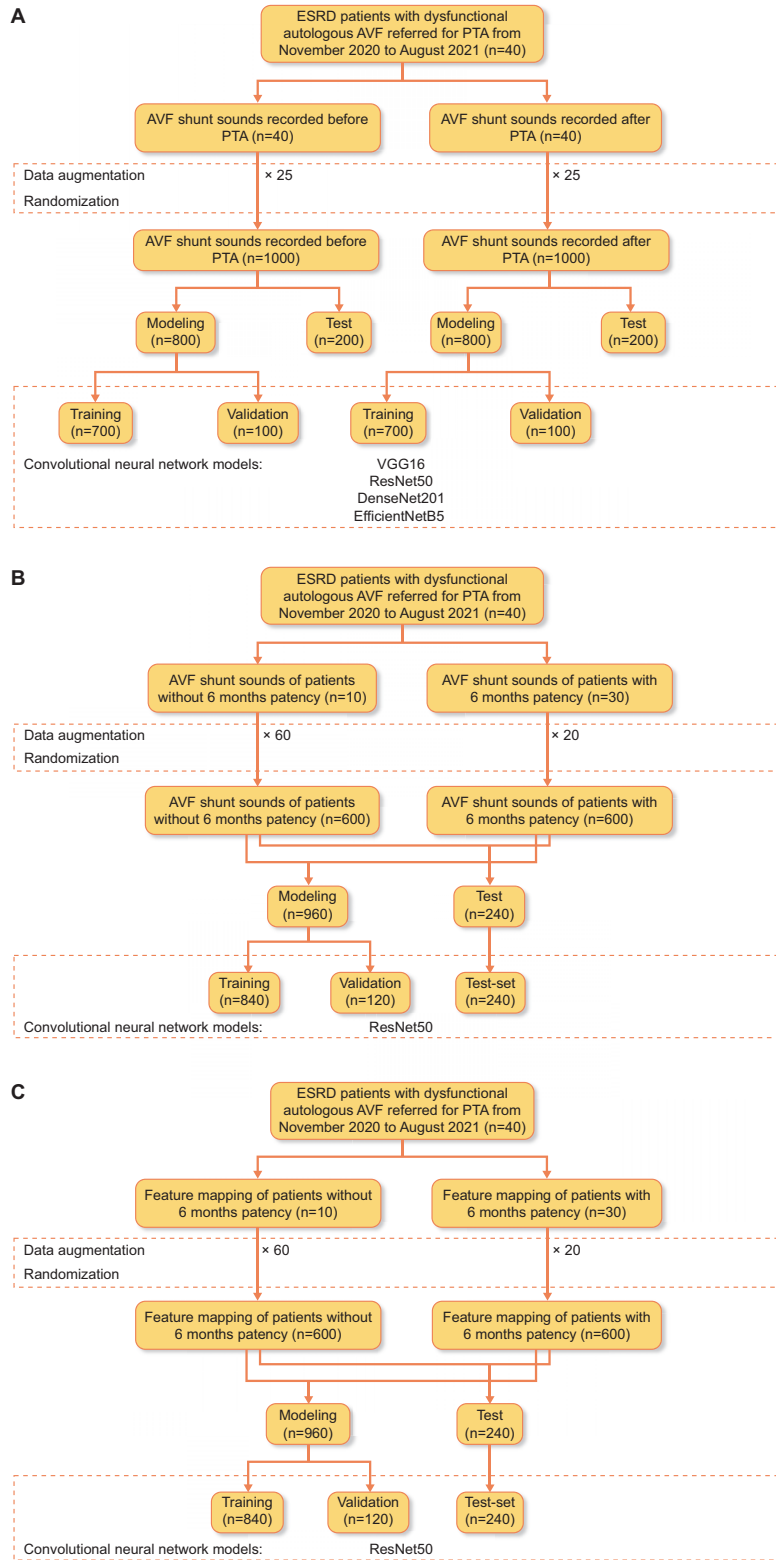
Figure 2: Study flow chart for (**A**) CNN model predicting the degree of AVF stenosis; (**B**) CNN model predicting 6-month PP from AVF shunt sound; and (**C**) CNN model predicting 6-month PP using feature mapping (i.e. spiral-matrix) of patients' clinical data.

findings (from now on "clinical models"), patients' baseline characteristics including age, sex, chronic kidney disease etiology, comorbidities, type and location of AVF, previous PTA history, and laboratory results were obtained from electronic medical records.

Construction of clinical models

Before constructing clinical models, visual exploration of data was performed using two dimensionality reduction techniques: principal component analysis (PCA) and t-distributed stochastic neighboring entities (t-SNE); the former focuses heavily on linear algebra, while the latter is a probabilistic technique [25].

Clinical models for predicting the 6-month PP have been developed using logistic regression (LR), decision tree (DT) [26], and support vector machines (SVM) [27]. In addition, a DCNN model (ResNet50) was trained using a spiral-matrix-arranged patient's clinicodemographic and laboratory findings [28]. In the case of LR, binary LR with backward elimination was used to determine the predictors of 6-month PP. DT, a hierarchical model with decision rules used to predict outcomes from a set of input variables, was trained using a maximum depth of three layers and the importance of the features was calculated. For SVM, a type of supervised ML algorithm was used to classify data points by maximizing the margin between classes in a high-dimensional space, and 3-degree polynomial clinical data was used to train the binary SVM classifier. For the LR, DT and SVM models, the clinical data of a total of 40 patients was used to construct the models and contingency tables.

For spiral-matrix-arranged clinical data (from now on "spiral matrix"), we constructed a 6×6 matrix of patients' numerical clinical data and converted *NumPy* array data into Python Imaging Library grayscale images. Images were then augmented 60 times for patients without 6-month PP and 20 times for patients with 6-month PP for the normalization of the imbalanced dataset. For these images, width and height shift, shear, and zoom with a range of 0.2 and "nearest" fill mode were used. The best CNN architecture used for predicting 6-month PP based on melspectrogram was also used for model prediction based on a spiral matrix with training, validation and test split ratios of 70%, 10% and 20%, respectively (Fig. 2**C**).

Evaluation metrics

Four standard metrics were used to evaluate the proposed method for predicting the degree of AVF stenosis: mean absolute error (MAE), MSE, root-mean-square error (RMSE) and R-squared value (R^2). In brief, MSE is the average square value between the true and predicted values; RMSE is the percentage difference between the true and predicted values; and MAE is the percentage difference between the predicted values.

The performance of the proposed model in predicting 6-month PP was evaluated using a confusion matrix, precision (i.e. positive predictive value), accuracy, recall (i.e. hit rate, sensitivity or true positive rate), F-1 score, and area under the receiver operating characteristics curve (AUROC).

Statistical analyses

The Mann–Whitney U test was used for continuous variables, and the χ^2 or Fisher's exact test was used for categorical variables. All statistical analyses were performed using Google Colab or SAS software (version 9.4) software (SAS Institute, Inc.). Statistical significance was defined as a two-sided *P*-value of <.05.

RESULTS

Study patients' characteristics

The patient characteristics are summarized in Table 1. Of the total patients, 22 (55%), 16 (40%) and 2 (5%) had brachial-cephalic, radial-cephalic and brachial-basilic AVFs, respectively. The median degrees of AVF stenosis were 59% and 23% before and after PTA, respectively. The median time between AVF formation and PTA was 839 days; 70% of the total patients had received two or fewer PTAs and 90% had received five or fewer PTAs, with a median time of 386 days since the previous PTA. The technical success rate was 100%; however, 25 (63%) patients showed <30% AVF stenosis after PTA, while 30 patients (75%) showed 6-month PP.

Performance of melspectrogram-based DCNN models for predicting the degree of AVF stenosis

Four CNN architectures were used to construct DCNN models to quantitatively predict the degree of AVF stenosis. The evaluation metrics of the DCNN models are listed in Table 2. MAE and RMSE were the lowest, and R^2 values were the highest in the following order: VGG16, ResNet50, DenseNet201 and EfficientNetB5, for both pre-AVF and post-AVF models. While all other DCNN models showed an $R^2 \geq 0.90$, the R^2 of the EfficientNetB5 pre-AVF and post-AVF models were 0.82 and 0.85, respectively. While VGG16, ResNet50 and DenseNet201 showed comparable MAE, MSE and RMSE, only the Grad-CAM heatmap of ResNet50 accurately reflected the areas in the melspectrogram where the greatest difference in amplitude and frequency occurred as a result of stenosis (Fig. 3). In contrast, Grad-CAM heatmaps of VGG16 and DenseNet201 pointed to areas in the melspectrogram, irrespective of the degree of stenosis. Moreover, the MSE, MAE and RMSE of DenseNet201 and EfficientNetB5 pre-AVF and post-AVF models increased at epochs higher than the optimized epoch, and thus showed signs of over-fitting (**Supplementary data, Figs S1 and S2**). Thus, the ResNet50 model was considered to best predict the degree of AVF stenosis, and the ResNet50 architecture was used to construct the DCNN model for predicting the 6-month PP.

Performance of clinical models for predicting 6-month PP of AVF

The results of the χ^2-test and univariate analysis of clinical parameters for the 6-month PP are summarized in **Supplementary data, Tables S1 and S2**. Patients without a 6-month PP showed a higher number of previous PTA (4.5 vs 1.0, *P* = .008), lower hemoglobin levels (g/dL) (11.2 vs 11.7, *P* = .038) and higher low-density lipoprotein cholesterol levels (mg/dL) (85.2 vs 61.6, *P* = .022) than those with a 6-month PP. Technical success (<30% post-PTA stenosis) was not associated with the 6-month PP (*P* = .715), and no difference in the degree of residual stenosis was found between patients with and without 6-month PP (*P* = .873). In the multivariate LR analysis, the number of previous PTAs was the only significant predictor of loss of 6-month PP (odds ratio 1.7, 95% confidence interval 1.1–2.6, *P* = .016) (**Supplementary data, Table S3**). The sensitivity and specificity of predicting 6-month PP based on the number of previous PTAs are shown in **Supplementary data, Table S4**. Cases with more than two previous PTAs showed a highest Youden index of 0.53, with a sensitivity of 70% and specificity of 83.3%.

Table 1: Clinico-demographic characteristics of study patients.

Age (years)[a]	62.0 (53.0–69.8)
Sex	
Male/female, *n* (%)	24 (60)/16 (40)
BMI[a]	22.7 (21.0–25.1)
Location of AVF, *n* (%)	
Forearm	16 (40)
Upper arm	24 (60)
AVF type, *n* (%)	
Brachial-cephalic	22 (55)
Radial-cephalic	16 (40)
Brachial-basilic	2 (5)
Location of stenosis, *n* (%)	
Juxta-anastomotic vein	21 (53)
Cephalic arch	12 (30)
Cannulation zone	7 (18)
Median AVF stenosis (%) before PTA	59.1 (49.1–65.2)
Median AVF stenosis (%) after PTA	23.4 (15.1–36.4)
Median time between AVF formation date and PTA (days)[a]	839.5 (529.0–1351.5)
Median flow (ml/min) before PTA[a]	491.4 (246.2–772.9)
Median flow (ml/min) after PTA[a]	899.6 (654.9–1339.5)
Median follow-up period since PTA to last follow-up (days)[a]	450.0 (351.8–525.8)
Number (*n*) of previous PTAs, *n* (%)	
$n \leq 2$	28 (70)
$2 < n \leq 5$	8 (20)
$5 < n \leq 8$	2 (5)
$8 < n$	2 (5)
Median time since previous PTA (days)[a]	386.5 (170.0–791.3)
6-month PP, *n* (%)	30 (75)
CKD etiology, *n* (%)	
Diabetic nephropathy	26 (65)
Hypertensive nephropathy	3 (8)
C1q nephropathy	2 (5)
RPGN	1 (3)
ADPKD	1 (3)
Unknown	7 (18)
Comorbidities, *n* (%)	
Type 2 diabetes mellitus	26 (65)
Hypertension	26 (65)
Coronary artery occlusive disease	10 (25)
Hyperlipidemia	9 (23)
Heart failure	4 (10)
Peripheral artery occlusive disease	2 (5)
Smoker or ex-smoker	11 (28)

[a]The data is presented as median values. Data in parentheses are 25th percentile and 75th percentile.
ADPKD, autosomal dominant polycystic kidney disease; BMI, body mass index; CKD, chronic kidney disease; RPGN, rapidly progressive glomerulonephritis

Table 2: Evaluation metrics for regression models.

	VGG16		ResNet50		DenseNet201		EfficientNetB5	
	Pre-AVF	Post-AVF	Pre-AVF	Post-AVF	Pre-AVF	Post-AVF	Pre-AVF	Post-AVF
MAE	0.018	0.011	0.028	0.043	0.044	0.037	0.059	0.066
MSE	0.001	0.000	0.001	0.003	0.003	0.002	0.005	0.007
RMSE	0.024	0.017	0.036	0.051	0.055	0.048	0.072	0.081
R^2	0.981	0.993	0.956	0.940	0.896	0.947	0.820	0.845

Before using ML to construct a clinical model for predicting the 6-month PP, the dataset was first visualized using PCA and t-SNE techniques, but no clusters were visually observed (**Supplementary data, Fig. S3**). Three ML techniques (LR, DT and SVM) and spiral-matrix-based DCNN (from now on "spiral-matrix DCNN") were used to construct clinical models. The top four important features of DT are shown in **Supplementary data, Fig. S4**. Unlike the multivariate LR model, platelet count was the most important feature, followed by the previous number of PTA, prolongation time and serum blood urea nitrogen level.
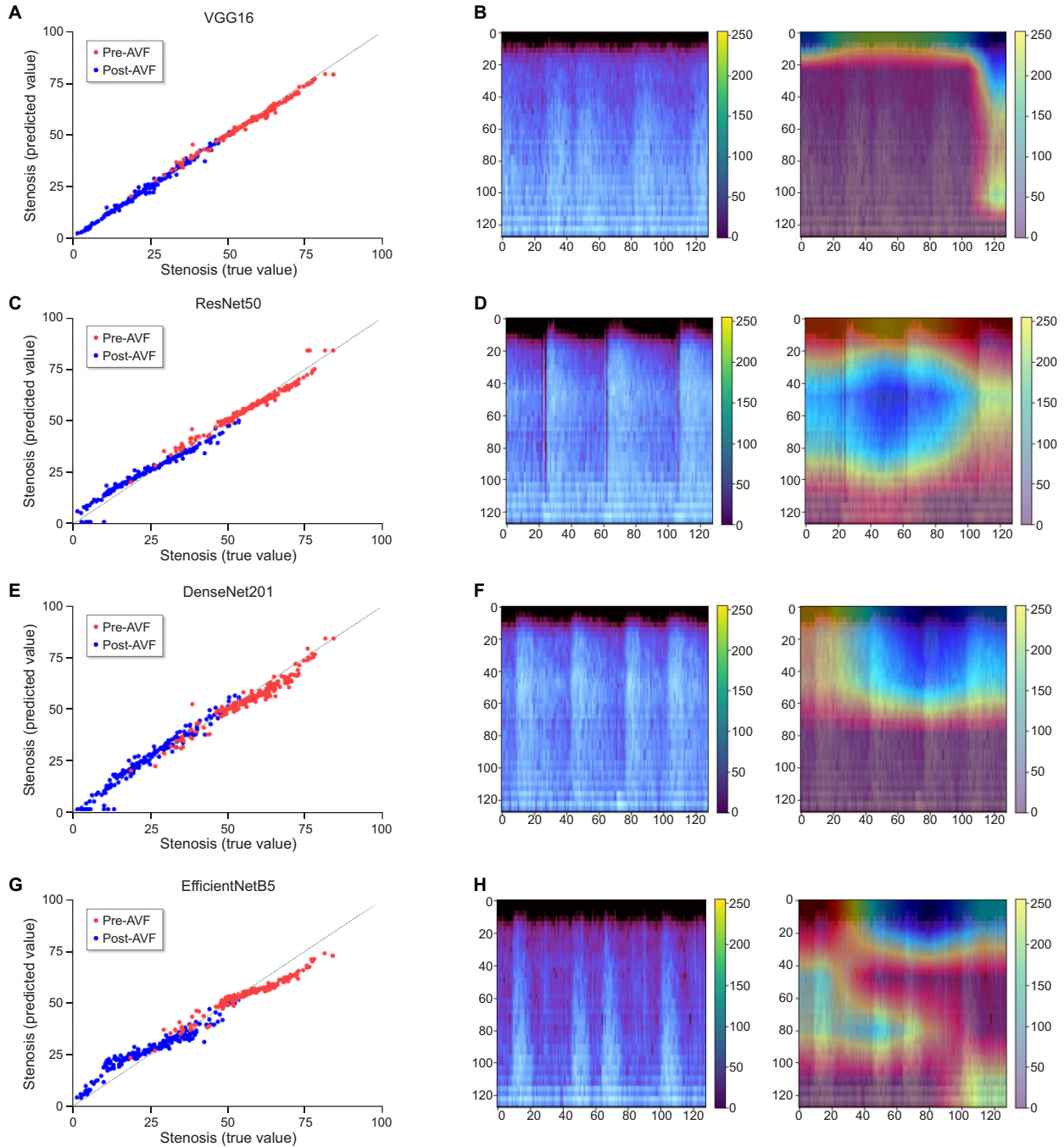
Figure 3: Regression plots for stenosis prediction based on melspectrogram using VGG16, ResNet50, DenseNet201 and EfficientNetB5 CNN models (**A**, **C**, **E** and **G**) and gradient-weighted class activation mapping (Grad-CAM) calculated from the final 2D convoluted layers of VGG16, ResNet50, DenseNet201 and EfficientNetB5 CNN models (**B**, **D**, **F** and **H**) used for stenosis prediction.

The receiver operating characteristic (ROC) curves of the LR, DT, SVM and spiral-matrix DCNN (ResNet50) models are shown in Fig. 4. The AUROC was highest in the following order: spiral-matrix DCNN (0.828), LR (0.783), DT (0.766) and SVM (0.733). Among the DCNN models utilizing the same ResNet50 architecture, the performance of the melspectrogram-based model was superior to that of the spiral-matrix-based model. The accuracy of the spiral-matrix DCNN model in the validation set was lower than that of the training set, whereas the loss of this model was higher in the validation set than in the training set, thus showing signs of under-fitting (**Supplementary data, Fig. S5**). In contrast, the melspectrogram-based DCNN model for predicting the 6-month PP of the validation set showed a similar level of accuracy and loss as that of the training set at approximately 80 epochs (**Supplementary data, Fig. S6**). Moreover, at an epoch >55, the performance of the melspectrogram-based DCNN
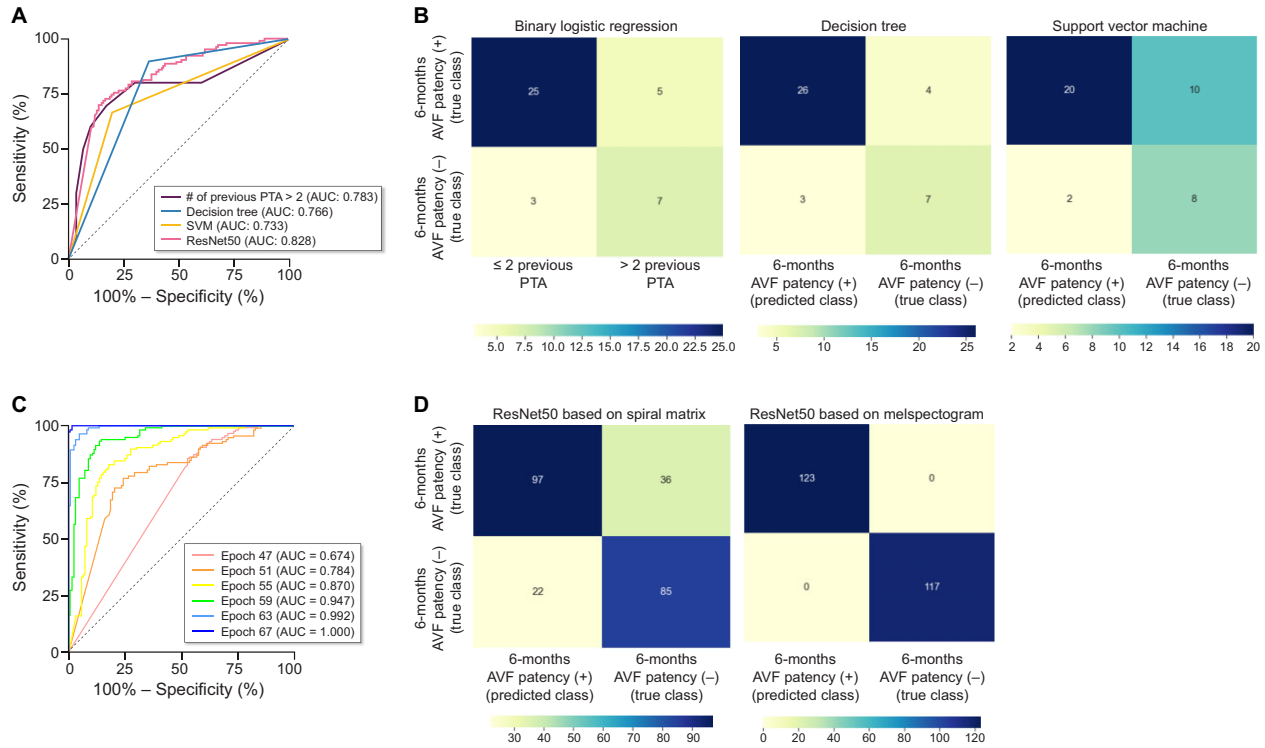
Figure 4: (**A**) Receiver operating characteristic (ROC) curves of various ML models based on clinical features (binary LR, DT and SVM, respectively) and ResNet50 CNN model based on spiral matrix from clinical features. (**B**) Contingency tables of various ML models based on clinical features including binary LR, DT and SVM. (**C**) ROC curves of ResNet50 CNN model for 6-month AVF PP based on melspectrogram. (**D**) Confusion matrix of ResNet50 CNN model for 6-month AVF PP based on spiral matrix from clinical features and from melspectrogram.

Table 3: Precision, recall and F-1 score of the ResNet50 CNN model based on the spiral matrix vs ResNet50 CNN model based on the melspectrogram.

	Precision	Recall	F-1 score
ResNet50 CNN model based on spiral matrix (i.e. clinical features)	0.82	0.73	0.77
ResNet50 CNN model based on melspectrogram	1.0	1.0	1.0

model surpassed that of the spiral-matrix DCNN model, with an AUC of 0.870 (Fig. 4). Overall, comparing these two models, the precision, recall and F-1 score of the melspectrogram-based DCNN model outperformed those of the spiral-matrix DCNN model (Table 3).

Correlation of melspectrogram pattern with the degree of AVF stenosis and 6-month PP

With a more severe degree of AVF stenosis, the melspectrogram showed a greater magnitude of amplitude at mid-to-high frequency, mainly in the systolic phase, corresponding to the high-pitched systolic bruit. Thus, when comparing the melspectrogram of the stenotic AVF with that of the non-stenotic AVF, differences in the melspectrogram were mostly apparent in the magnitude of the high-pitched and low-pitched bruit. In patients with significant ($\geq$50%) AVF stenosis without technical success of PTA, the magnitudes of the high-pitched and low-pitched bruits were similar. In contrast, in patients with significant ($\geq$50%) stenosis where technical success of PTA was

achieved the magnitude of the high-pitched bruit decreased, while that of the low-pitched bruit remained similar or slightly increased (Fig. 5). Furthermore, the Grad-CAM heatmap of the melspectrogram-ResNet50 model highlighted the borders between the high-pitched and low-pitched bruits for predicting the 6-month PP (Fig. 5).

DISCUSSION

The proposed melspectrogram-based DCNN model successfully predicted the degree of AVF stenosis. Of the four different CNN architectures used to construct this model, the ResNet50 model showed no signs of overfitting and predicted the degree of AVF stenosis with a low RMSE (0.036–0.051) and a high R^2 value (0.940–0.956). Using AVF shunt sounds obtained after PTA, a melspectrogram-based DCNN model for predicting the 6-month PP was also constructed. This model outperformed any clinical model constructed using conventional statistics (i.e. binary LR), ML (i.e. DT and SVM) or spiral-matrix DCNN.
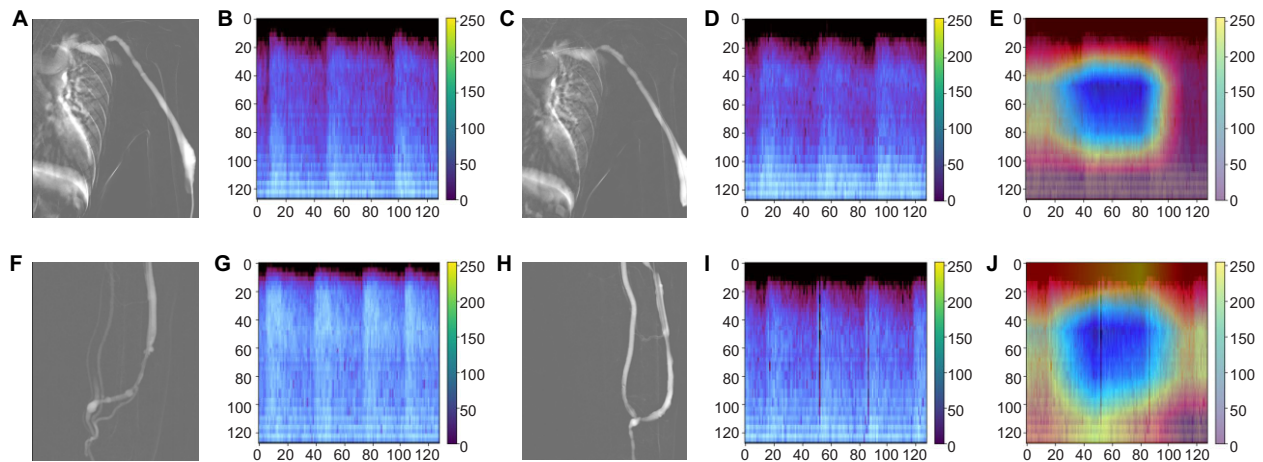
Figure 5: (A–E) A 61-year-old male with brachiocephalic AVF: (**A**) DSA before PTA showing about 65.4% stenosis at cephalic arch, (**B**) melspectrogram of AVF sound before PTA, (**C**) DSA after PTA showing about 53.7% stenosis at cephalic arch, (**D**) melspectrogram of AVF sound after PTA, (**E**) gradient-weighted class activation mapping (Grad-CAM) calculated from the final 2D convoluted layers of ResNet50 CNN used for 6-month PP. This patient showed failed technical success and was referred for another PTA within 6 months. (F–J) A 62-year-old male with radiocephalic AVF: (**F**) DSA before PTA showing about 55.7% stenosis at cephalic vein, (**G**) melspectrogram of AVF sound before PTA, (**H**) DSA after PTA showing 13.7% stenosis at cephalic vein, (**I**) melspectrogram of AVF sound after PTA, (**J**) gradient-weighted class activation mapping (Grad-CAM) calculated from the final 2D convoluted layers of ResNet50 CNN used for 6-month PP. This patient showed successful technical success and was not referred for another PTA within 6 months.

The high diagnostic performance of the melspectrogram-based DCNN model for predicting the degree of AVF stenosis suggests that the findings from the auscultation of AVF shunt sounds correlates well with that of actual angiography. Qualitatively, the melspectrogram of a more stenotic AVF showed a greater magnitude of amplitudes at mid-to-high frequency in the systolic phase, which is consistent with a high-pitched bruit accompanied by systolic accentuation. In contrast, melspectrograms of <50% AVF stenosis showed a greater magnitude of low-pitched bruit. Except for EfficientNetB5, all other CNN architectures adequately correlated melspectrograms to the degree of AVF stenosis, with $R^2 \geq 0.90$. Previously, Glangetas *et al.* proposed the idea of developing a portable autonomous stethoscope integrated with AI [29] that can be easily converted into a smartphone accessory [30]. While this idea was first introduced to classify lung sounds, using the proposed model, AVF shunt sounds can be used to non-invasively monitor the degree of AVF stenosis by general health practitioners, including dialysis staff members, who may not specialize in nephrology, vascular surgery or intervention.

Controversies still exist as to which clinical factors affect the 6-month PP. While age [31, 32], sex [33, 34] and comorbidities such as diabetes [35] have been suggested as risk factors for loss of patency, other studies [36–38] have reported conflicting results. These conflicting results make the accurate prediction of the loss of PP challenging. In this study, the previous number of PTA was a significant risk factor according to both multivariate analysis and the DT classifier. This finding was consistent with that of a previous study indicating that patients with a stenosis and more than one previous PTA underwent primary PTA significantly sooner than those who did not [39]. The simplest clinical model of more than two previous PTAs showed a higher accuracy than either the DT or SVM models. In this study, while more than two previous PTA showed the highest accuracy, more than three previous PTAs showed 90% specificity for 6-month PP. This result was similar to that of a previous study by Elramah *et al.*, who reported a history of more than three previous PTAs to be

associated with diminished PP [40]. However, although this clinical model may be statistically sound, the mere assumption of primary failure based on the number of prior PTA can be dangerous in clinical practice.

As an alternative method, the feasibility of a melspectrogram-based DCNN model for predicting the 6-month PP was explored. The melspectrogram-based DCNN model successfully predicted the 6-month PP and outperformed the clinical models, including the spiral-matrix DCNN model. Grad-CAM heatmaps highlighted the areas in the melspectrogram corresponding to the borders between the high-pitched and low-pitched frequency bruit. This may imply that the degree of restoration in the magnitude of the normal low-pitch bruit and persistence of systolic accentuation may be involved in predicting primary failure. Although no difference in the degree of residual AVF stenosis was noted between patients with and without 6-month PP, unidimensional measurement of residual stenosis on DSA images may be less sensitive to residual blood flow turbulence that may be captured in the AVF shunt sound.

This study had several limitations. First, this pilot study included a small number of patients, all of whom had venous outflow obstruction and underwent interventions by a single interventional radiologist, which may have contributed to a selection bias. The primary focus of this study was to assess the feasibility of the DCNN model in predicting the degree of AVF stenosis and 6-month PP. A subsequent study, which included patients without AVF dysfunction, is underway to evaluate whether these models can be used to differentiate or screen patients in need of intervention from those who do not. Second, because of the paucity of AVF shunt sounds, the SMOTE algorithm was used to generate a synthetic melspectrogram. Third, clinical data from a small number of patients may have resulted in only a few statistically significant predictors of 6-month PP and may have contributed to the underestimated performance of clinical models. Fourth, melspectrograms provide only one type of visual representation of audio data, and for simplicity, the use of other representations such as harmonic-percussive spectrograms or

scattergrams have not been explored [41]. Fifth, DCNN models were not constructed for different AVF type and location of stenosis due to the small number of patients. Lastly, measurements from DSA were used as a reference instead of other imaging modalities such as DUS, as DSA is the current gold standard for assessing vascular stenosis [42].

In conclusion, the melspectrogram-based DCNN model successfully predicted the degree of AVF stenosis and outperformed ML-based clinical models in predicting the 6-month PP.

SUPPLEMENTARY DATA

Supplementary data are available at *ckj* online.

FUNDING

AUTHORS' CONTRIBUTIONS

J.H.P., J.Y., I.P. and K.H were responsible for conceptualization. J.H.P. and J.Y. were responsible for methodology. J.H.P., I.P. and J.Y. were responsible for formal analysis and draft preparation. Y.S., S.J.K., and J.Y.W. were responsible for resources and review and editing of the manuscript. J.H.P. and J.Y. were responsible for visualization. K.H. was responsible for supervision. All authors have read and agreed to the published version of the work.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study will be shared upon reasonable request to the corresponding author.

CONFLICT OF INTEREST STATEMENT

The authors have no competing interests to declare that are relevant to the content of this work.

REFERENCES

1. National Kidney Foundation. Clinical practice guidelines and clinical practice recommendations for 2006 updates: hemodialysis adequacy, peritoneal dialysis adequacy and vascular access. *Am J Kidney Dis* 2006;**48**:S1e322.
2. Turmel-Rodrigues L, Pengloan J, Baudin S *et al*. Treatment of stenosis and thrombosis in haemodialysis fistulas and grafts by interventional radiology. *Nephrol Dial Transplant* 2000;**15**:2029–36.
3. Clark TW, Hirsch DA, Jindal KJ *et al*. Outcome and prognostic factors of restenosis after percutaneous treatment of native hemodialysis fistulas. *J Vasc Interv Radiol* 2002;**13**:51–9.
4. Huber TS, Carter JW, Carter RL *et al*. Patency of autogenous and polytetrafluoroethylene upper extremity arteriovenous hemodialysis accesses: a systematic review. *J Vasc Surg* 2003;**38**:1005–11.
5. Lok CE, Huber TS, Lee T *et al*. KDOQI clinical practice guideline for vascular access: 2019 update. *Am J Kidney Dis* 2020;**75**:S1–164.
6. Swarup S, Makaryus AN. *Digital stethoscope: Technology update. Med Devices (Auckl)* 2018;**11**:29.
7. Altan G, Kutlu Y, Allahverdi N. Deep learning on computerized analysis of chronic obstructive pulmonary disease. *IEEE J Biomed Health Inform* 2019;**24**:1344–50.
8. Aykanat M, Kılıç Ö, Kurt B *et al*. Classification of lung sounds using convolutional neural networks. *EURASIP J Image Video Process* 2017;**2017**:1–9.
9. Fernandez-Granero MA, Sanchez-Morillo D, Leon-Jimenez A. Computerised analysis of telemonitored respiratory sounds for predicting acute exacerbations of COPD. *Sensors* 2015;**15**:26978–96.
10. Murphy RL, Vyshedskiy A, Power-Charnitsky V-A *et al*. Automated lung sound analysis in patients with pneumonia. *Respir Care* 2004;**49**:1490–7.
11. Ota K, Nishiura Y, Ishihara S *et al*. Evaluation of hemodialysis arteriovenous bruit by deep learning. *Sensors* 2020;**20**:4852.
12. Wang H-Y, Wu C-H, Chen C-Y *et al*. Novel noninvasive approach for detecting arteriovenous fistula stenosis. *IEEE Trans Biomed Eng* 2014;**61**:1851–7.
13. Doelman C, Duijm LE, Liem YS *et al*. Stenosis detection in failing hemodialysis access fistulas and grafts: comparison of color Doppler ultrasonography, contrast-enhanced magnetic resonance angiography, and digital subtraction angiography. *J Vasc Surg* 2005;**42**:739–46.
14. Park JH, Park I, Han K *et al*. Feasibility of deep learning-based analysis of auscultation for screening significant stenosis of native arteriovenous fistula for hemodialysis requiring angioplasty. *Korean J Radiol* 2022;**23**:949–58.
15. McFee B, Raffel C, Liang D *et al*. librosa: audio and music signal analysis in python. *Proceedings of the 14th Python in Science Conference*. Citeseer: 2015;18–25.
16. Palanisamy K, Singhania D, Yao A. Rethinking CNN models for audio classification. arXiv preprint arXiv:2007.11154 2020.
17. Sehgal A, Kehtarnavaz N. A convolutional neural network smartphone app for real-time voice activity detection. *IEEE Access* 2018;**6**:9017–26.
18. Chawla NV, Bowyer KW, Hall LO *et al*. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 2002;**16**:321–57.
19. Simonyan K, Zisserman A, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 2014.
20. He K, Zhang X, Ren S *et al*. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016;770–8.
21. Huang G, Liu Z, Van Der Maaten L *et al*. Densely connected convolutional networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017;4700–8.
22. Tan M, Le Q. Efficientnet: rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning*. PMLR: 2019;6105–14.
23. Kingma DP, Ba J. *Adam: A method for stochastic optimization*. arXiv preprint arXiv:1412.6980 2014.
24. Selvaraju RR, Cogswell M, Das A *et al*. Grad-cam: visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision*. 2017;618–26.
25. Pareek J, Jacob J. Data compression and visualization using PCA and T-SNE. *Advances in Information Communication Technology and Computing*. Bikaner, India: Springer, 2020;327–37.
26. Pradhan B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Comput Geosci* 2013;**51**:350–65.

27. Chen H-L, Yang B, Wang G *et al*. Support vector machine based diagnostic system for breast cancer using swarm intelligence. *J Med Syst* 2012;**36**:2505–19.

28. Zhang B, Qi S, Pan X *et al*. Deep CNN model using CT radiomics feature mapping recognizes EGFR gene mutation status of lung adenocarcinoma. *Front Oncol* 2021;**10**: 3318.

29. Glangetas A, Hartley M-A, Cantais A *et al*. Deep learning diagnostic and risk-stratification pattern detection for COVID-19 in digital lung auscultations: clinical protocol for a case–control and prospective cohort study. *BMC Pulm Med* 2021;**21**:1–8.

30. Vasudevan RS, Horiuchi Y, Torriani FJ *et al*. Persistent value of the stethoscope in the age of COVID-19. *Am J Med* 2020;**133**:1143–50.

31. Hodges TC, Fillinger MF, Zwolak RM *et al*. Longitudinal comparison of dialysis access methods: risk factors for failure. *J Vasc Surg* 1997;**26**:1009–19.

32. Miller PE, Carlton D, Deierhoi MH *et al*. Natural history of arteriovenous grafts in hemodialysis patients. *Am J Kidney Dis* 2000;**36**:68–74.

33. Ifudu O, Mayers JD, Cohen LS *et al*. Correlates of vascular access and nonvascular access-related hospitalizations in hemodialysis patients. *Am J Nephrol* 1996;**16**:118–23.

34. Feldman HI, Kobrin S, Wasserstein A. Hemodialysis vascular access morbidity. *J Am Soc Nephrol* 1996;**7**:523–35.

35. Monroy-Cuadros M, Yilmaz S, Salazar-Bañuelos A *et al*. Risk factors associated with patency loss of hemodialysis vascular access within 6 months. *Clin J Am Soc Nephrol* 2010;**5**:1787–92.

36. Windus DW, Jendrisak MD, Delmez JA. Prosthetic fistula survival and complications in hemodialysis patients: effects of diabetes and age. *Am J Kidney Dis* 1992;**19**:448–52.

37. Smith GE, Gohil R, Chetter IC. Factors affecting the patency of arteriovenous fistulas for dialysis access. *J Vasc Surg* 2012;**55**:849–55.

38. Prischl FC, Kirchgatterer A, Brandstätter E *et al*. Parameters of prognostic relevance to the patency of vascular access in hemodialysis patients. *J Am Soc Nephrol* 1995;**6**:1613–8.

39. Bountouris I, Kristmundsson T, Dias N *et al*. Is repeat PTA of a failing hemodialysis fistula durable? *Int J Vasc Med* 2014;**2014**:369687.

40. Elramah M, Boujelbane L, Yevzlin AS *et al*. Dialysis access venous stenosis: treatment with balloon angioplasty 30-second vs. 1-minute inflation times. *Hemodial Int* 2015;**19**:108–14.

41. Nanni L, Costa YM, Aguiar RL *et al*. Ensemble of convolutional neural networks to improve animal audio classification. *EURASIP J Audio Speech Music Process* 2020;**2020**:1–14.

42. Lin Y-P, Wu M-H, Ng Y-Y *et al*. Spiral computed tomographic angiography–a new technique for evaluation of vascular access in hemodialysis patients. *Am J Nephrol* 1998;**18**:117–22.