



# Article Application of a Deep Learning Neural Network for Voiding Dysfunction Diagnosis Using a Vibration Sensor

Yuan-Hung Pong <sup>1,2</sup>, Vincent F.S. Tsai <sup>1,2,†</sup>, Yu-Hsuan Hsu <sup>3</sup>, Chien-Hui Lee <sup>4</sup>, Kun-Ching Wang <sup>5</sup>, and Yu-Ting Tsai <sup>3,6,7,\*,†</sup>

- <sup>1</sup> Department of Urology, Ten Chen Hospital, Taoyuan 326, Taiwan; pong9101095@yahoo.com.tw (Y.-H.P.); ntubala@gmail.com (V.F.S.T.)
- <sup>2</sup> Department of Urology, National Taiwan University Hospital, College of Medicine, National Taiwan University, Taipei 10617, Taiwan
- <sup>3</sup> Master's Program of Electro-Acoustics, Feng Chia University, Taichung City 407, Taiwan; seanpomelo@gmail.com
- <sup>4</sup> Department of Mechanical Engineering, National Yunlin University of Science and Technology, Douliu 64002, Taiwan; fcu7302@gmail.com
- <sup>5</sup> Department of Mechanical and Computer-Aided Engineering, Feng Chia University, Taichung City 407, Taiwan; kunching.wang@gmail.com
- <sup>6</sup> Bachelor's Program in Precision System Design, Feng Chia University, Taichung City 407, Taiwan
- <sup>7</sup> Hyper-Automation Laboratory, Feng Chia University, Taichung City 407, Taiwan
- Correspondence: yuttsai@fcu.edu.tw
- + These authors contributed equally to this work.

Abstract: In a clinical context, there are increasing numbers of people with voiding dysfunction. To date, the methods of monitoring the voiding status of patients have included voiding diary records at home or urodynamic examinations at hospitals. The former is less objective and often contains missing data, while the latter lacks frequent measurements and is an invasive procedure. In light of these shortcomings, this study developed an innovative and contact-free technique that assists in clinical voiding dysfunction monitoring and diagnosis. Vibration signals during urination were first detected using an accelerometer and then converted into the mel-frequency cepstrum coefficient (MFCC). Lastly, an artificial intelligence model combined with uniform manifold approximation and projection (UMAP) dimensionality reduction was used to analyze and predict six common patterns of uroflowmetry to assist in diagnosing voiding dysfunction. The model was applied to the voiding database, which included data from 76 males aged 30 to 80 who required uroflowmetry for voiding symptoms. The resulting system accuracy (precision, recall, and f1-score) was around 98% for both the weighted average and macro average. This low-cost system is suitable for at-home urinary monitoring and facilitates the long-term uroflow monitoring of patients outside hospital checkups. From a disease treatment and monitoring perspective, this article also reviews other studies and applications of artificial intelligence-based methods for voiding dysfunction monitoring, thus providing helpful diagnostic information for physicians.

Keywords: voiding dysfunction; urodynamics; artificial intelligence; mel-frequency cepstrum coefficient

# 1. Introduction

As a result of concurrent developments in healthcare and technology, artificial intelligence technologies have made exceptional contributions to improving healthcare quality. For instance, in medical imaging, deep learning models have been developed for predicting bladder cancer [1] or other urinary tract tumors [2]. Relevant studies have demonstrated that these models can lessen excessive diagnoses and treatment sessions, alleviate pain in patients, and reduce unnecessary treatment costs. Regarding clinical studies on voiding dysfunction, according to the statistics compiled by Berry et al. [3], the prevalence of voiding dysfunction increases with age. The incidence of storage and voiding male lower



Citation: Pong, Y.-H.; Tsai, V.F.S.; Hsu, Y.-H.; Lee, C.-H.; Wang, K.-C.; Tsai, Y.-T. Application of a Deep Learning Neural Network for Voiding Dysfunction Diagnosis Using a Vibration Sensor. *Appl. Sci.* **2022**, *12*, 7216. https://doi.org/10.3390/ app12147216

Academic Editors: Dijana Petrovska-Delacrétaz and Mounim A. El Yacoubi

Received: 23 May 2022 Accepted: 14 July 2022 Published: 18 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). urinary tract symptoms report [4] about 51% and 26%, respectively. Interestingly, it is 18% of men who reported the combination of storage and voiding symptoms. Typical symptoms of voiding dysfunction include urinary frequency, urgency, urge incontinence, obstructive sensation, and even urinary retention. These conditions could result in recurring urinary tract infections and lead to kidney failure in severe cases if not managed appropriately. In clinical practice, there are many people with voiding dysfunction, and current approaches to monitoring the voiding status of patients include voiding diary records at home or urodynamic examinations at hospitals [5]. The former is less objective and is prone to missing data, while the latter is an invasive procedure that requires costly equipment, thus reducing the frequency of measurements.

In the context of monitoring voiding dysfunction, medical professionals expect to achieve their task goals through fully contact-free approaches [6]. Studies on the assistive monitoring of voiding dysfunction, such as by Bright et al. [7], utilized ultrasound to estimate the bladder weight of male patients and analyzed the association between bladder weight and voiding dysfunction. Kuo et al. [8] used ultrasound to measure the detrusor wall thickness of patients. They found a correlation between the detrusor wall thickness of male patients and overactive bladder syndrome, while no such correlation was found in females. These findings were consistent with those of Robinson et al. [9], who investigated 128 female patients whose urodynamic examination findings were either ambiguous or whose symptoms could not be explained by the examination findings alone. After measuring the bladder wall thickness of patients through ultrasound, the authors noted a marginal correlation between bladder wall thickness and voiding dysfunction among female patients.

In the study by Choi et al. [10], only 7.2% of the 1415 female patients sought urologic treatment for lower urinary tract symptoms. Among them, voiding symptoms were more common than storage symptoms according to urodynamic examinations. The low visiting rate of female patients might be due to the lack of a convenient and non-invasive diagnostic detection of voiding dysfunction. Similarly, because of data acquisition difficulties, this study targeted men needing assistive monitoring of voiding dysfunction to serve as participants. Jerry Blaivas et al. [11] attached a small-size vibration sensor that transforms vibration signals into uroflow during urination to the dossal site of the penis. The device assists in in-home monitoring of patients' voiding status. It is a new technology for repeated outpatient uroflow monitoring.

The remaining sections of this text are as follows: Section 2 concludes the related works of machine learning technologies applied for voiding symptoms identification and dysfunction diagnoses of clinical research, Section 3 describes the materials and methodology of the proposed system, Section 4 introduces the data analysis methods and model development results and discussion, and Section 5 presents the conclusions and directions for future research.

# 2. Related Works

Artificial intelligence (AI) has been widely used to study and treat various diseases. The study of Zeeshan Hameed et al. [12] illustrated that AI diagnosis is feasible in urology, regardless of disease severity—from benign prostatic hyperplasia to critical diseases such as urothelial and prostate cancers. Therefore, using AI technology to detect, assess, and treat urological diseases is superior to all the existing traditional methods. Eun et al. [13] indicated the possibility of the modern application and development of AI in urological research, especially in the diagnosis and treatment of urological diseases. Machine learning algorithms will improve the accuracy of predictions for various bladder diseases, including interstitial cystitis, bladder cancer, and reproductive urology. The application of AI leads to a higher accuracy of disease diagnosis and prediction models to monitor accurate decision making in medical treatment. Qureshi et al. [14] used general neural network (ANN) and convolutional neural network (CNN) methods to effectively solve the problem of female urinary incontinence diagnosis and management. Jin et al. [15] combined MFCC with a

convolutional neural network (CNN) and developed a psychoacoustic model to measure automotive transmission noise perceived by humans. Cuocolo et al. [16] used machine learning (ML) to read prostate magnetic resonance imaging (MRI). This study compares the ML method with traditional methods. Many clinical applications of MRI do not segment glands, which limits their effectiveness in detecting, diagnosing, and locating cancer. In this study, the particular novel ML model incorporating ML and deep learning (DL) is proposed to highlight the differences between MRI applications in the prostate treatment pipeline. However, the validation and reproducibility of clinical applicability must be improved. Jin et al. [17] generally used non-invasive uroflowmetry to diagnose lower urinary tract symptoms (LUTS) and the patient's health status. This study established an AI system that used acoustic features for LUTS assessment and was validated with a long short-term memory (LSTM) deep learning algorithm. The results were similar to clinical urological measurements. In the future, the system could be developed for home testing devices to realize the functions of remote continuous monitoring and diagnosis.

In clinical research, Enshaeifar et al. [18] indicated that urinary tract infection (UTI) is one of the top five reasons for hospital admissions of dementia patients (about 9% of all admissions of dementia patients in the UK). Based on that, we worked with clinical physicians on the machine learning algorithms designed to detect UTI and integrated internet of things (IoT) technology, home sensory devices, and machine learning techniques to monitor the health and well-being of people with dementia. This will allow us to provide more effective and advanced care for the patients and reduce the number of hospital admissions. Karmonik et al. [19] used four random forest machine learning methods to study multiple sclerosis (MS) in 27 females with neurogenic lower urinary tract dysfunction (NLUTD) and voiding dysfunction (VD). The relative importance of brain regions with urination-initiated reduced functional connectivity (FC) was part of the research. Micturating people and VD patients showed significantly different voiding initiation networks, and machine learning was able to observe and identify differences in brains. Lee et al. [20] developed a smartphone-based approach to acoustic uroflowmetry to achieve accuracy and reliability and to compare with conventional uroflowmetry. A prospective validation experiment with 128 subjects concluded that the use of acoustics to analyze uroflowmetry is comparable to conventional analyses of uroflowmetry. This allows patients to perform urine flow measurements in the clinic or at home and provides a reliable and low-cost test. Dawidek et al. [21] used audio-based uroflowmetry as a novel alternative method. The subjects in this experiment were 44 male patients. Urinary flow can be confirmed by sound measurement. With a smartphone application, screening and monitoring can be performed. This approach can be applied to urological disorders and can be used outside the clinic. Helou et al. [22] proposed and evaluated a novel mobile acoustic flow measurement (sonouroflowmetry, SUF) method to overcome the inconvenience of the traditional non-invasive uroflowmetry (UF) clinical test. The proposed method estimated urinary flow rate in the diagnosis of lower urinary tract dysfunction (lower urinary tract dysfunction, LUTD) using mobile phone-recorded sound signals. In this method, by linearly mapping the total sound energy to the total voided volume, the sound energy curve was converted into a flow rate curve, which allows for an estimation of the flow rate over time. Evaluation of this experiment on data from 44 healthy young men showed a high degree of similarity between UF and SUF flow rates, with a mixed-effects model correlation coefficient of 0.993 and a mean root mean square error of 2.37 mL/s.

This study was based on a smart healthcare recommendation system for human voiding dysfunction, as shown in Figure 1a, with the target population being men with voiding dysfunction. Unlike past methods of using microphones for urinary flow sound capturing, we directly use the attached vibration sensor, which can avoid receiving background sound noise from the environment. The accelerometer was installed on the side of the funnel-shaped urine bucket in the clinic, as shown in Figure 1b. When the bucket is used, the instrument will simultaneously detect the urine flow and the vibration signal. Our goal is to develop a smart, contact-free monitoring system for voiding dysfunction to shed preliminary light on patients' existing symptoms. The combination of urodynamics and the principles of artificial intelligence facilitates data collection and patient status monitoring, thus providing diagnostic information that enables both physicians and patients to achieve favorable outcomes.







(b)

**Figure 1.** (a) The study flowchart and framework. (b) The experimental vibration accelerometer attached to the urinal bucket.

# 3. Method Descriptions

# 3.1. System Architecture

The architecture of the proposed system is shown in Figure 2. The system consists of three major components: (1) a vibration signal and deep learning-based voiding pattern classifier; (2) an accelerometer and signal extractor that performs real-time vibration signal collection and testing; and (3) a graphical user interface that displays the output. The first step developed a vibration signal dataset and trained and tested the model through deep learning, followed by data storage and deployment of the offline model. The final step performed real-time testing and displayed the results. The components of the system are described in the following sections.



Figure 2. The components and architecture of the prototype system.

# 3.2. System Hardware

A prototype system was developed to test the deep learning model in real-time. The system's hardware components include a urinary flow meter, a Kistler triaxial accelerometer (8688A50), a signal extractor (NI 9234), and a mobile user device that displays real-time feedback information. The system usage procedure is described as follows: When a patient urinates, the impacting urine creates vibrations on the accelerometer, which are first converted into vibration signals registered by the accelerometer and then into digital signals by the signal extractor. The signals are transmitted wirelessly to the mobile device, and the results are tabulated and displayed.

# 3.3. Data Sets

The real-time monitoring system for voiding dysfunction classification was developed by training and testing the deep learning model using the collected datasets, named Ten Chen Medical Group (TCMG) voiding database. The following subsections describe the dataset details, signal preprocessing, and model training and testing.

The TCMG datasets used for deep learning model training and testing [23,24] were collected using measures and standards approved by the institutional review board (IRB) of the cooperating hospital. The hospital used the developed system to acquire patients' uroflowmetry data. The TCMG voiding database included data from 76 males aged 30 to 80 who required uroflowmetry [25] for voiding symptoms. The exclusion criteria were small bladder capacity (<50 mL), external genital malformations, and inability to complete uroflowmetry. For each participant, at least one or two uroflow measurements were performed, and each was recorded for no longer than 30 s. An average of approximately 24.67 seconds of signal length was collected per patient. The total signal length in the database was 31.25 min. Each patient's data was annotated with his diagnosis through a uroflowmetry test [25] for voiding dysfunction diagnosis. The uroflowmetry is a key diagnostic information for voiding dysfunction. The physicians labeled all the voiding patterns and clinical impressions. There were six typical voiding patterns labeled from 0 to 5 (0 indicates normal voiding while the others indicate abnormal patterns). A total of 38 patients were diagnosed with decrease, 4 patients with flattened, 5 patients with intermittent, 9 patients with saw-tooth, 2 patients with tall and peak, and 18 patients with normal.

As shown in Figure 1, the six typical voiding patterns are labeled classified as: 0 (normal flow, which is the bell-shaped pattern with an adequate peak flow rate); 1 (Tall and peak flow, i.e., the urine flow rate reaches a high flow peak quickly and decline soon); 2 (decreased flow, i.e., a prolonged urine flow time with a low peak flow rate presented); 3 (saw-tooth flow, i.e., the urine flow rate fluctuates greatly but does not drop to zero); 4 (intermittent flow, i.e., the urine flow rate drops to zero and rise again later); and 5 (flattened flow, i.e., the urine flow rate is gradual and constant until the end). In this study, the physicians visually compared the voiding patterns based on the raw vibration signal data. For example, for Patient 39, the physician labeled the raw data as "decreased flow", i.e., a prolonged urine flow time with a low peak flow rate was presented, which is consistent with the observed raw data. Thus, there is a correlation between the datasets and the patients' urinary features.

#### 3.4. Data Preprocessing

The root mean square (RMS) [26] of a sensor-generated vibration signal was used to assess the vibration level in the data preprocessing. As shown in Figure 3, the raw signals underwent pre-emphasis filtering through a Butterworth filter, were converted into the frequency domain via a fast Fourier transform (FFT) [27], and their features were extracted through the mel scale. The final step generated the spectrogram of the signals.

#### 3.4.1. RMS Processing

The RMS value is a reference for assessing the magnitude of vibrations. It is the mean vibrational energy within a given period, calculated as follows:

$$\text{RMS} = \sqrt{\frac{1}{T} \sum_{n=0}^{T} X(n)^2}$$
(1)

where *T* is the period, and *X* is the accelerometer vibration signal. This equation yields the RMS of a raw signal, and this value is used for comparisons with the uroflowmetry chart used initially for diagnosis. The patterns in both charts should be consistent with each other.



Figure 3. Data preprocessing procedure.

# 3.4.2. MFCC Processing

This project converts the subject's periodic breathing audio data into a characteristic spectrum sequence using mel-frequency cepstrum [28]. This method is based on a linear transformation of the logarithmic energy spectrum of the mel frequency scale, which approximates the human auditory system.

The raw input signals were windowed such that each segmented window had a length of 2048. The input signals were subjected to pre-emphasis filtering the unwanted high frequency components through a Butterworth low-pass filter [29]. When the pre-emphasis filtering of the signals was complete, they were converted from the time domain to the frequency domain via an FFT, and the frequencies were further converted to units of mels [30,31]. The mel scale is a non-linear scale unit based on the pitches perceived by the human ear. The formula equation for converting between the mel scale and frequency is as follows:

$$y(k) = \log_{10} \left( \sum_{k=0}^{10000} |x(k)|^2 B_m(k) \right)$$
(2)

where *k* is the frequency grid point, x(k) is the frequency spectrum of auditory signal in time domain, and y(k) is the mel spectrum after applying the function of  $B_m(k)$  as the mask function of the mel-frequency cepstrum, which corresponds to mel frequency scale  $f_m(k)$  as

$$f_m(k) = 2595 \log_{10} \left( 1 + \frac{k}{700} \right) \tag{3}$$

Then, the principle of mel triangular filtering is based on the use of a certain number (m) of triangular filters, and the interval between each f(m) changes with the value of m; the interval decreases when m decreases, and vice versa. The frequency response of a triangular filter [32] is defined in the following equation:

$$B_m(k) = \begin{cases} 0, \ k < f_m(k-1) \ \text{and} \ k > f_m(k) \\ \frac{k - f_m(k-1)}{f_m(k) - f_m(k-1)}, \ f_m(k-1) \ \le k \le f_m(k) \\ \frac{-k + f_m(k+1)}{-f_m(k) + f_m(k+1)}, \ f_m(k) \ \le k \le f_m(k+1) \end{cases}$$
(4)

Using the inverse discrete cosine transform to obtain the MFCC parameter value by,

$$y_{mfcc}(n) = \frac{1}{k_N} \sum_{k=1}^{k_N} |\mathbf{y}(k)|^2 \cos\left(\frac{\pi n(k-0.5)}{k_N}\right)$$
(5)

A mel triangular filter smooths the frequencies and reduces the harmonic effects while highlighting the raw signal features as  $y_{mfcc}(n)$ , where n is set to 32 in this case. Additionally, the derived spectrogram  $y_{mfcc}(n)$  improves the learning speed of a CNN by reducing the amount of data to be processed.

#### 3.4.3. Data Analysis

This study transformed the original physiological vibration signal through the MFCC spectrums to display the energy distribution of the physiological vibration signal in the time–frequency domain. This conversion allows the characteristics of various physiological vibration signals to become more prominent and easier to be directly observed by the human eye. Among them, red is the color response when the energy of the time–frequency diagram is higher, and blue is the lower energy. The y-axis of the time spectrum is normalized and displayed, and the frequency interval is divided into 1 Hz to 150 Hz. The example analysis diagram (as comparatively shown in the normal flow in Figure 4a and the tall and peak flow in Figure 4b) shows that the normalized resonance frequency of the measuring cup is around the 50 Hz to 100 Hz. In normal urination, the energy area of the spectrum is continuous and concentrated; on the contrary, in abnormal conditions, as in the case of the tall and peak flow, it indicates that the patient's voiding velocity rises quickly and declines later. Furthermore, no significant resonance in frequency was observed in the decreased, saw-tooth, intermittent, and flattened flow cases as shown in Figure 4c-f, in which multiple impulses can be seen on the time axis according to the RMS signals. However, it is evident based on the time axis that the intervals between and patterns of urination differ among all cases.

On the other hand, as shown in Figure 4c,e, the patient's initial RMS value is marked by Qmax, and there are small signal impulses. In reality, the saw-tooth (Figure 4d) and intermittent flows (Figure 4e) are presented as multiple and instantaneous occurrences of strained urination, which is a symptom of compensated abdomen forcing due to outlet obstruction or detrusor muscle weakness. However, in comparison with Figure 4d, the intermittent flow in Figure 4e shows more multiple and instantaneous occurrences of strained urination. Therefore, even though both types of flow have similar features, the peak frequency in the same time interval differs throughout the urination process. The flattened flow in Figure 4f suggests that insufficient energy was applied during urination. The resonance energy generated by the measuring cup is low, and the energy sustained is minimal compared to the other cases. Due to the insufficient energy applied for urination, the length of urination is comparatively longer, which reflects the characteristic of this vibration.

#### 3.5. Deep Convolutional Neural Network

The model developed in this study as shown in Figure 5 was based on a CNN architecture [33,34] consisting of convolutional layers with batch normalization layers [35] and pooling layers, a rectified linear unit (ReLU) as activation function layers [36], fully connected layers, and then finally with a Softmax function for classification purposes [37]. The difference between a CNN and a traditional artificial neural network [38] is that the former combines a convolutional layer and a pooling layer to perform down-sampling, thus increasing the feature extraction process and subsequently reducing the time required to train the neural network. Each hidden layer is expressed as the following sections:



**Figure 4.** Plotting of RMS signal and MFCC spectrum of cases of six uroflowmetry patterns, including normal case. (a) The normal flow; (b) The tall and peak flow; (c) The decrease flow; (d) The saw-tooth flow; (e) The intermittent flow; (f) The flattened flow.



Figure 5. The architecture of purposed convolutional neural network.

#### 3.5.1. Convolutional Layer

There are two types of convolutional operations: one-dimensional (1-D) and twodimensional (2-D). This study adopted the second approach because 2-D spectrograms were input into the neural network model. In image feature extraction using convolutional layers, multiple convolutional kernels, or filters [39] slide and convolve over the input image in a defined order. After completing the operation, a feature map is generated, serving as the next layer's input data. The advantage of a convolutional layer is that it retains the original order information of an image, and the neurons in a convolutional layer can only connect with some of the neurons in the preceding layer. The neurons in the same convolutional layer share parameters, significantly reducing the number of parameters required for training.

$$Y(i,j) = \sum_{m} \sum_{n} w(m,n) X(i+m-\theta,j+n-\theta)$$
(6)

where *X* represents the input data, *Y* represents the output results,  $\theta$  is the stride, *w* is the convolution kernel, and *n* and *m* are the one-time convolution kernel sliding lengths, respectively.

# 3.5.2. MaxPooling Layer

A MaxPooling layer primarily reduces the complexity of a feature map and uses a kernel-based approach to pick the largest value of the feature map. A pooling layer differs from a convolutional layer because it does not require learning parameters and is simpler to operate. The equation of applying MaxPool can be represented as

$$Y = \mathbf{MaxPool}(i_d, j_d) = \max\{X(i + \Delta, j + \Delta\}$$
(7)

More detailed features can be extracted through a pooling layer while also effectively reducing the number of parameters needed for subsequent training as well as preventing overfitting [40,41]. Common pooling layer approaches include max pooling and average pooling. The latter effectively preserves background noise but is prone to feature blurring. The former extracts the maximum values from a feature map and effectively filters background noise, making it more suitable for processing spectrograms. Hence, the max pooling layer processing method was used in this study to highlight voiding dysfunction spectrum features.

#### 3.5.3. Fully Connection Layer

When the MaxPooling layer extracts the largest characteristics from the convolution results, the Full Connection layer is then used for compiling the features extracted to form the final output. If the two-dimensional convolution results are substituted into the Full Connection layer, the two-dimensional matrix generated after the convolution must be

converted into a one-dimensional vector, expressed as,  $y_n = \text{reshape}[\text{MaxPool}_{last}(i, j)]$ , n = 1, 2, 3, ..., N and N is the vector length after conversion. Then, the full connection layer can be expressed as

$$Y_m = \operatorname{RELU}\left(\sum_m w_m X_m + b_m\right) \tag{8}$$

where weighting *w* and bias *b* are the updating parameters of the Full Connection layer, and the ReLU is the activation function as RELU( $\emptyset$ ) = max(0, $\emptyset$ ). For the classification of six classes, the Softmax function, as  $Y_{class} = \exp[z_j] / \sum_{i} \exp[z_i]$ , which maps the multiple

classification results of the model's final prediction output to the elements in each vector is located within (0, 1), and the sub-intervals represent the probability distribution of each classification.

For supervised learning, this study used a set of MFCC spectra and their corresponding voiding patterns labels to define the classification purpose of the convolutional neural network. To train the proposed model, the Adam (adaptive moment estimation) [42] iterative search algorithm was used to search for the neural network coefficient to minimize the error of the model predictions and ground truth (clinically case label).

#### 4. Results and Discussion

### 4.1. Model Evaluation

We used statistical indicators to assess the prediction model and further described it using binary classification. *P* represents positive dataset values, and *N* represents negative dataset values. True positive (TP) and true negative (TN) indicate that the predicted values are consistent with actual values, while false positive (FP) and false negative (FN) indicate that the predicted values differed from actual values [43,44]. In this study, the four following indicators were used for assessment:

Accuracy: The ratio of datasets correctly predicted by the classifier to the total number of datasets.

$$Accuracy = \frac{TP + FN}{TP + TN + FP + FN}$$
(9)

Precision: The ratio of correctly predicted positives to the total number of predicted positives in the sample.

$$Precision = \frac{TP}{TP + FP}$$
(10)

Recall: The ratio of correctly predicted positives to the total number of positives in the sample.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{11}$$

F1-score: A measure of the precision of an algorithm by simultaneously including precision and recall.

$$F1-score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$
(12)

Based on the preceding descriptions, accuracy is an indicator that directly measures the overall identification results but cannot reflect the actual status of each category. However, in the training process, most of the datasets in the major categories (Labels 0 and 1) could learn more information and features than the minor categories (Labels 2 to 5). This issue may reduce the identification capacity of the minor categories despite the high overall accuracy in the model evaluation. Since some classes of data have significant quantitative imbalances, we used random oversampling, which involves randomly selecting examples from the minority class, with replacement, and adding them to the training dataset but not for the testing dataset (unseen dataset). In this study, naive random oversampling was

used for balancing and augmenting data. To avoid overfitting, naive random oversampling was used to generate new samples for balancing and augmenting data from the minority classes (2, 3 and 5), which is underrepresented.

The classification model for voiding dysfunction patterns was developed by dividing the TCMG database into training and testing datasets. K-Fold cross-validation resulted in statistically significant results. In K-Fold cross-validation, the original sample is randomly divided into k equal-sized subsamples. Among the k subsamples, one subsample is reserved as validation data for testing the model, and the remaining k - 1 subsamples are used as training data. The cross-validation process is then repeated k times, and each of the k subsamples is used only once as validation data. The k results can then be averaged to produce a single estimate. In this study, we used 10 observations for k = 2. After the data were shuffled, a total of 5 models were trained and tested. Then, the validation results were averaged over the rounds to estimate the model's predictive performance.

Estimating the performance of the model can be resolved using the three indicators (precision, recall, and F1 score), thus highlighting their importance. One of the representative CNN results is shown in Table 1. On this basis, all indicators were used for model assessment, with the results shown in Table 1. The testing results revealed that the model recall accuracy was 98.19% for the weighted average and 98.09% for the macro average. The precision, recall, and F1 score of Labels 3 and 5 exceeded 99%, suggesting a higher classifying ability than the other labels. The result of the K-Fold cross-validation is provided in Table 2. As shown, the average accuracy of three indicators as assessed through K-Fold cross-validation was between 97.98% and 98.25%, indicating that the CNN model's predictive performance is significant.

Table 1. The performance of urination patterns in data sets and evaluation indicators.

Label	Category	Number of Person	Precision	Recall	F1-Score	Support Data (Length in Seconds)
0	Normal	18	0.9647	0.9932	0.9787	86.17 s
1	Decrease	38	0.9937	0.9876	0.9906	183.45 s
2	Flattened	4	0.9979	0.9844	0.9911	16.07 s
3	Intermittent	5	0.9228	0.9977	0.9588	29.15 s
4	Saw-tooth	9	0.9989	0.9248	0.9604	47.65 s
5	Tall and peak	2	1.0000	0.9974	0.9987	13.05 s
W	Macro-average accura /eighted-average accu	cy racy	0.9797 0.9826	0.9809 0.9819	0.9797 0.9819	375.54 s

Table 2. Result of K-Fold cross-validation for the CNN model.

K-Fold	Mac	cro-Average Accu	racy	Weighted-Average Accuracy					
	Precision	Recall	F1-Score	Precision	Recall	F1-Score			
1	0.9797	0.9809	0.9797	0.9826	0.9819	0.9819			
2	0.9760	0.9798	0.9772	0.9805	0.9797	0.9797			
3	0.9783	0.9823	0.9797	0.9831	0.9822	0.9822			
4	0.9866	0.9711	0.9787	0.9806	0.9805	0.9805			
5	0.9897	0.9849	0.9872	0.9881	0.9880	0.9879			
Average	0.9821	0.9798	0.9805	0.9830	0.9825	0.9824			

Table 3 shows comparisons between the performance of three different types of neural networks, using the same data settings as Table 1, in which CNN and recurrent neural network (RNN) with LSTM architecture had more than 0.9 accuracy, while the simple three-layer hidden layer had a precision of more than 0.9. In this study, the vibration spectrum was used as the input. For a general ANN, due to the lack of convolutional layers, we first flattened each spectral signal into one dimension and input it to a pure full connection

layer with 512 neurons; then we added an output full connection layer with a Softmax activation function for classification. As a result, the ANN model achieved a precision of 0.6873. In the case of multi-dimensional approaches, the LSTM and ANN models are difficult to detect effectively the features displayed by the frequency because there is no convolution method to extract the important features of the MFCC spectrum. However, LSTM has a good comparison of before and after information detection for time series, so the accuracy still has a certain feasible level. However, compared to ANN's approach of only having a purely full connection layer, its accuracy is much lower.

Table 3	. The	performance	ce of	different	neural	networl	k l	earni	ing 1	metl	nod	s wi	th ev	zalu	atio	on i	ind	icat	ors
---------	-------	-------------	-------	-----------	--------	---------	-----	-------	-------	------	-----	------	-------	------	------	------	-----	------	-----

Method	Precision	Recall	F1-Score	Remark	
RNN	0.9026	0.9010	0.9018	Inspired by LSTM [17]	
ANN	0.6873	0.6626	0.6658	Inspired by [14]	
CNN	0.9826	0.9819	0.9819	Proposed	
T . 1 TAT ! 1 .	1				1

Noted: Weighted-average accuracy.

To validate the classifiable nature of the collected data, the AI model developed in this study was also subjected to data clustering through uniform manifold approximation and projection (UMAP) data visualization [45]. The testing results provided in Tables 1 and 2 and the results of the UMAP data visualization analysis show that this measure indeed enhances data classification. The vibration signals in the study were converted into MFCC spectrum features. The UMAP approach was used to convert these spectrum features onto an N-to-2D plane (where N was selected to be two in this case), and the data labels were expressed in colors to visualize the data so that the model's performance could be easily observed, as shown in Figure 6a,b. As a result, as shown in Figure 6a, each class feature is rough and overlapping before training. After training, each label of data can be distinctly categorized as shown in Figure 6b; it is observed that different clusters can be segmented, thus confirming the classifiable nature of the datasets.



**Figure 6.** Segmentation of the raw data of the UMAP regions: (**a**) raw data, (**b**) the data extracted from the final layer output, where 0–5 represent the urinary pattern label.

#### 4.2. Real-Time Implementation

Based on the real-time urine flow monitoring, this study introduced the model to identify real-time voiding situations, and the experiment will identify and diagnose several cases based on time series. As shown in Figure 7, the red labels represent actual values in the bottom figure, while the blue curve represents model-predicted values. Sections in

which the colors overlap indicate that the model had accurately predicted the classification. Because most of the lines overlapped, we can conclude that the model has excellent predictive power.



**Figure 7.** Expanded samples and labels in time series to verify the accuracy of the model applied to real-time detection.

Although the RMSs of the vibration signals sometimes are similar, the predictions were different compared to the ground truth labels in the time series. For the physiciandiagnosed symptoms, a voiding pattern is usually classified as a primary or a secondary diagnosis in a clinical diagnosis. However, most patients had composite symptoms because of the lack of one-type clinical data. There are slight identification errors that occurred in each category. For example, the model determined the primary diagnosis and secondary diagnosis for patients with their primary diagnosis being bladder outlet obstruction (BOO) as decrease and intermittent flow cases, respectively. The average prediction was consistent with the physician's diagnostic label for minor prediction abnormalities. However, the application of time series is very likely to be an application example of home health care.

#### 5. Conclusions and Future Work

This study has proposed an innovative assistive diagnostic system that uses existing testing and system development methods to generate contact-free vibration signals from men with voiding dysfunction. Based on the high accuracy rate, there are some possible clinical applications. The at-home and continuous monitoring of voiding conditions can provide clinicians with important information on voiding drug effects or pre- and post-prostate surgery voiding conditions, e.g., alpha-blockers for prostate enlargement disease or transurethral resection of prostate surgery. Otherwise, the AI model with a time recording system makes the voiding diary more convenient and delicate.

Nonetheless, this study has certain limitations. One limitation is the lack of female data due to the difficulties with female voiding recording. We may collect female voiding data during the urodynamic examination. In addition, the AI model needs to be calibrated when collecting signals in different media, e.g., the uroflowmetry plastic bottle or a toilet bowl. We may be able to develop a correction coefficient database based on the vibration signals of different media analyses. Composite diagnostic criteria can easily lead to model identification errors in clinical diagnoses. The relevant datasets must be broadened to classify the primary symptoms more accurately. We can use the UMAP method to visualize and understand how well clustered each category is. The results of the visualization analysis show that the CNN enhanced the performance of the classification of different voiding symptoms. These results need to be taken with caution, as they were obtained on a rather small and unbalanced database.

The proposed monitoring system is low-cost, small-sized, and provides real-time results, making it suitable for patients undergoing at-home care. It also achieves real-time remote and continuous symptom monitoring. In the future, the data collected by the system will assist in the diagnosis of voiding dysfunction-related diseases and facilitate the integration of large databases and multiple sources of information, producing algorithms

that improve voiding dysfunction diagnoses. These algorithms can help physicians to more effectively gain access to diagnostic information, reduce the probabilities of diagnostic errors, and compensate for existing shortcomings in clinical diagnosis.

Author Contributions: Conceptualization, Y.-H.P., V.F.S.T. and Y.-T.T.; investigation, Y.-H.P. and Y.-H.H.; methodology, V.F.S.T. and Y.-T.T.; analysis, Y.-H.P., K.-C.W., C.-H.L. and Y.-H.H.; writing—original draft preparation, Y.-H.P. and Y.-H.H.; writing—review and editing, Y.-H.P., C.-H.L., V.F.S.T. and Y.-T.T.; funding acquisition, V.F.S.T. and Y.-T.T., V.F.S.T. and Y.-T.T. contributed equally. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors gratefully acknowledge the Ministry of Science and Technology of Taiwan for financial support under Grant Nos. MOST 109-2625-M-005-006-MY2, MOST 110-2622-E-035-006-, and MOST 109-2221-E-035-001-MY2 and the joined project of Feng Chia University and Ten Chen Medical Group.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Ten Chan General Hospital (protocol No. 107003 and date of approval: 21 November 2018).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the IRB statement.

Conflicts of Interest: The authors declare no conflict of interest.

# References

- Woerl, A.-C.; Eckstein, M.; Geiger, J.; Wagner, D.C.; Daher, T.; Stenzel, P.; Fernandez, A.; Hartmann, A.; Wand, M.; Roth, W.; et al. Deep learning predicts molecular subtype of muscle-invasive bladder cancer from conventional histopathological slides. *Eur. Urol.* 2020, 78, 256–264. [CrossRef]
- Nojima, S.; Terayama, K.; Shimoura, S.; Hijiki, S.; Nonomura, N.; Morii, E.; Okuno, Y.; Fujita, K. A deep learning system to diagnose the malignant potential of urothelial carcinoma cells in cytology specimens. *Cancer Cytopathol.* 2021, 129, 984–995. [CrossRef]
- Berry, S.J.; Coffey, D.S.; Walsh, P.C.; Ewing, L.L. The development of human benign prostatic hyperplasia with age. J. Urol. 1984, 132, 474–479. [CrossRef]
- Fusco, F.; Creta, M.; Trama, F.; Esposito, F.; Crocetto, F.; Aveta, A.; Mangiapia, F.; Imbimbo, C.; Capece, M.; La Rocca, R.; et al. Tamsulosin plus a new complementary and alternative medicine in patients with lower urinary tract symptoms suggestive of benign prostatic hyperplasia: Results from a retrospective comparative study. *Arch. Ital. Urol. Androl.* 2020, 92. [CrossRef] [PubMed]
- Drake, M.J.; Lewis, A.L.; Young, G.J.; Abrams, P.; Blair, P.S.; Chapple, C.; Glazener, C.M.; Horwood, J.; McGrath, J.S.; Noble, S.; et al. Diagnostic assessment of lower urinary tract symptoms in men considering prostate surgery: A noninferiority randomised controlled trial of urodynamics in 26 hospitals. *Eur. Urol.* 2020, *78*, 701–710. [CrossRef] [PubMed]
- 6. Jalbani, I.K.; Ather, M.H. The accuracy of three-dimensional bladder ultrasonography in determining the residual urinary volume compared with conventional catheterisation. *Arab J. Urol.* **2014**, *12*, 209–213. [CrossRef] [PubMed]
- 7. Bright, E.; Pearcy, R.; Abrams, P. Ultrasound estimated bladder weight in men attending the uroflowmetry clinic. *Neurourol. Urodyn.* **2011**, *30*, 583–586. [CrossRef]
- 8. Kuo, H.-C.; Chen, Y.-C.; Chen, C.-Y.; Chancellor, M.B. Transabdominal ultrasound measurement of detrusor wall thickness in patients with overactive bladder. *Tzu Chi Med. J.* 2009, *21*, 129–135. [CrossRef]
- 9. Robinson, D.; Anders, K.; Cardozo, L.; Bidmead, J.; Toozs-Hobson, P.; Khullar, V. Can ultrasound replace ambulatory urodynamics when investigating women with irritative urinary symptoms? *BJOG Int. J. Obstet. Gynaecol.* **2002**, *109*, 145–148. [CrossRef]
- Choi, Y.S.; Kim, J.C.; Lee, K.S.; Seo, J.T.; Kim, H.-J.; Yoo, T.K.; Lee, J.B.; Choo, M.-S.; Lee, J.G.; Lee, J.Y. Analysis of female voiding dysfunction: A prospective, multi-center study. *Int. Urol. Nephrol.* 2013, 45, 989–994. [CrossRef]
- 11. Blaivas, J.; Benedon, M.; Weinberger, J.; Rozenberg, Y.; Ravid, L.; Vapnek, J. PD21-11 the dynamic urine vibration halter: A new outpatient device for remote patient monitoring of uroflow. *J. Urol.* **2015**, *193*, e475. [CrossRef]
- Hameed, B.M.Z.; Dhavileswarapu, A.V.L.S.; Raza, S.Z.; Karimi, H.; Khanuja, H.S.; Shetty, D.K.; Ibrahim, S.; Shah, M.J.; Naik, N.; Paul, R.; et al. Artificial Intelligence and Its Impact on Urological Diseases and Management: A Comprehensive Review of the Literature. J. Clin. Med. 2021, 10, 1864. [CrossRef]
- 13. Eun, S.J.; Kim, J.; Kim, K.H. Applications of artificial intelligence in urological setting: A hopeful path to improved care. *J. Exerc. Rehabil.* **2021**, *17*, 308–312. [CrossRef]

- Qureshi, A.; Mathur, A.; Alshiek, J.; Shobeiri, S.A.; Wei, Q. Utilization of Artificial Intelligence for Diagnosis and Management of Urinary Incontinence in Women Residing in Areas with Low Resources: An Overview. Open J. Obstet. Gynecol. 2021, 11, 403–418. [CrossRef]
- 15. Jin, S.; Wang, X.; Du, L.; He, D. Evaluation and modeling of automotive transmission whine noise quality based on MFCC and CNN. *Appl. Acoust.* **2021**, *172*, 107562. [CrossRef]
- 16. Cuocolo, R.; Cipullo, M.B.; Stanzione, A.; Ugga, L.; Romeo, V.; Radice, L.; Brunetti, A.; Imbriaco, M. Machine learning applications in prostate cancer magnetic resonance imaging. *Eur. Radiol. Exp.* **2019**, *3*, 35. [CrossRef]
- Jin, J.; Chung, Y.; Kim, W.; Heo, Y.; Jeon, J.; Hoh, J.; Park, J.; Jo, J. Classification of Bladder Emptying Patterns by LSTM Neural Network Trained Using Acoustic Signatures. Sensors 2021, 21, 5328. [CrossRef]
- Enshaeifar, S.; Zoha, A.; Skillman, S.; Markides, A.; Acton, S.T.; Elsaleh, T.; Kenny, M.; Rostill, H.; Nilforooshan, R.; Barnaghi, P. Machine learning methods for detecting urinary tract infection and analysing daily living activities in people with dementia. *PLoS ONE* 2019, 14, e0209909. [CrossRef]
- Karmonik, C.; Boone, T.; Khavari, R. Data-Driven Machine-Learning Quantifies Differences in the Voiding Initiation Network in Neurogenic Voiding Dysfunction in Women with Multiple Sclerosis. *Int. Neurourol. J.* 2019, 23, 195–204. [CrossRef]
- Lee, Y.J.; Kim, M.M.; Song, S.H.; Lee, S. A Novel Mobile Acoustic Uroflowmetry: Comparison With Contemporary Uroflowmetry. Int. Neurourol. J. 2021, 25, 150–156. [CrossRef]
- Dawidek, M.T.; Singla, R.; Spooner, L.; Ho, L.; Nguan, C. Clinical validation of an audio-based uroflowmetry application in adult males. *Can. Urol. Assoc. J.* 2022, 16, E120–E125. [CrossRef]
- 22. El Helou, E.; Naba, J.; Youssef, K.; Mjaess, G.; Sleilaty, G.; Helou, S. Mobile sonouroflowmetry using voiding sound and volume. *Sci. Rep.* **2021**, *11*, 11250. [CrossRef]
- Shkolyar, E.; Jia, X.; Chang, T.C.; Trivedi, D.; Mach, K.E.; Meng, M.Q.-H.; Xing, L.; Liao, J.C. Augmented bladder tumor detection using deep learning. *Eur. Urol.* 2019, 76, 714–718. [CrossRef]
- Yang, Y.; Zou, X.; Wang, Y.; Ma, X. Application of deep learning as a noninvasive tool to differentiate muscle-invasive bladder cancer and non–muscle-invasive bladder cancer with CT. *Eur. J. Radiol.* 2021, 139, 109666. [CrossRef]
- 25. Kelly, C.E. Evaluation of voiding dysfunction and measurement of bladder volume. Rev. Urol. 2004, 6, S32–S37.
- 26. Nakhaie Jazar, G.; Alkhatib, R.; Golnaraghi, M.F. Root mean square optimization criterion for vibration behaviour of linear quarter car using analytical methods. *Veh. Syst. Dyn.* 2006, 44, 477–512. [CrossRef]
- Gudivada, A.A.; Sudha, G.F. STQCA-FFT: A fast fourier transform architecture using stack-type QCA approach with power and delay reduction. J. Comput. Sci. 2022, 60, 101594. [CrossRef]
- Milner, B.; Shao, X. Clean speech reconstruction from MFCC vectors and fundamental frequency using an integrated front-end. Speech Commun. 2006, 48, 697–715. [CrossRef]
- Acharya, A.; Das, S.; Pan, I.; Das, S. Extending the concept of analog Butterworth filter for fractional order systems. *Signal Process*. 2014, 94, 409–420. [CrossRef]
- Abduh, Z.; Nehary, E.A.; Wahed, M.A.; Kadah, Y.M. Classification of heart sounds using fractional fourier transform based mel-frequency spectral coefficients and traditional classifiers. *Biomed. Signal Process. Control* 2020, 57, 101788. [CrossRef]
- Iqtidar, K.; Qamar, U.; Aziz, S.; Khan, M.U. Phonocardiogram signal analysis for classification of Coronary Artery Diseases using MFCC and 1D adaptive local ternary patterns. *Comput. Biol. Med.* 2021, 138, 104926. [CrossRef] [PubMed]
- Kui, H.; Pan, J.; Zong, R.; Yang, H.; Wang, W. Heart sound classification based on log Mel-frequency spectral coefficients features and convolutional neural networks. *Biomed. Signal Process. Control* 2021, 69, 102893. [CrossRef]
- 33. Ambroa, E.M.; Pérez-Alija, J.; Gallego, P. Convolutional neural network and transfer learning for dose volume histogram prediction for prostate cancer radiotherapy. *Med. Dosim.* **2021**, *46*, 335–341. [CrossRef] [PubMed]
- Milani, M.; Abas, P.E.; De Silva, L.C.; Nanayakkara, N.D. Abnormal heart sound classification using phonocardiography signals. Smart Health 2021, 21, 100194. [CrossRef]
- Amin, J.; Sharif, M.; Anjum, M.A.; Raza, M.; Bukhari, S.A.C. Convolutional neural network with batch normalization for glioma and stroke lesion detection using MRI. *Cogn. Syst. Res.* 2020, 59, 304–311. [CrossRef]
- Parisi, L.; Neagu, D.; Ma, R.; Campean, F. Quantum ReLU activation for Convolutional Neural Networks to improve diagnosis of Parkinson's disease and COVID-19. *Expert Syst. Appl.* 2022, 187, 115892. [CrossRef]
- Vo, D.M.; Nguyen, D.M.; Lee, S.-W. Deep softmax collaborative representation for robust degraded face recognition. *Eng. Appl. Artif. Intell.* 2021, 97, 104052. [CrossRef]
- Sun, Y.; Cui, Y.; Huang, Y.; Lin, Z.J. SDMP: A secure detector for epidemic disease file based on DNN. *Inf. Fusion* 2021, 68, 1–7. [CrossRef]
- Hesser, D.F.; Mostafavi, S.; Kocur, G.K.; Markert, B. Identification of acoustic emission sources for structural health monitoring applications based on convolutional neural networks and deep transfer learning. *Neurocomputing* 2021, 453, 1–12. [CrossRef]
- 40. Mutasa, S.; Sun, S.; Ha, R. Understanding artificial intelligence based radiology studies: What is overfitting? *Clin. Imaging* **2020**, 65, 96–99. [CrossRef]
- Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 2014, 15, 1929–1958.
- Aslan, M.F.; Unlersen, M.F.; Sabanci, K.; Durdu, A. CNN-based transfer learning–BiLSTM network: A novel approach for COVID-19 infection detection. *Appl. Soft Comput.* 2021, 98, 106912. [CrossRef]

- 43. Abhishek, A.; Jha, R.K.; Sinha, R.; Jha, K. Automated classification of acute leukemia on a heterogeneous dataset using machine learning and deep learning techniques. *Biomed. Signal Process. Control* **2022**, *72*, 103341. [CrossRef]
- Aslan, N.; Koca, G.O.; Kobat, M.A.; Dogan, S. Multi-classification deep CNN model for diagnosing COVID-19 using iterative neighborhood component analysis and iterative ReliefF feature selection techniques with X-ray images. *Chemom. Intell. Lab. Syst.* 2022, 224, 104539. [CrossRef]
- McInnes, L.; Healy, J.; Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv* 2018, arXiv:1802.03426.