Sensing and Bio-Sensing Research 52 (2026) 100999

Contents lists available at ScienceDirect

Sensing and Bio-Sensing Research

journal homepage: www.elsevier.com/locate/sbsr

Noninvasive intracranial hypertension detection using machine-learning of cerebral blood flow velocity waveforms

Miaomiao Wei [a,*,1], Solventa Krakauskaite [e,1], Ryan Mercer [b], Jiaguo Lin [d], Laimonas Bartusis [e], Fabien Scalzo [c]

[a] School of Information and Communication Engineering, Zhongyuan University of Technology, Zhengzhou, China
[b] University of California, Riverside (UCR), USA
[c] Department of Neurology and Computer Science, University of California, Los Angeles (UCLA), USA
[d] Keck Data Science Institute, Pepperdine University, Malibu, USA
[e] Health Telematics Science Institute, Kaunas University of Technology, Kaunas, Lithuania

ARTICLE INFO

Keywords:
CBFV
Intracranial hypertension detection waveform analysis
Machine learning noninvasive assessment

ABSTRACT

The monitoring of intracranial pressure (ICP) is crucial in the clinical management of various cerebral diseases and injuries, including head trauma, hydrocephalus, intracranial tumors, and cerebral edema. It could also play a broader role, for example, in the clinical management of stroke. The objective of monitoring is to prevent intracranial hypertension (IH) from causing further brain damage, which can be irreversible. However, current technology for ICP monitoring is invasive and typically requires pressure probes to be inserted through the skull, which is associated with potential complications. This procedure is reserved for the most severe clinical conditions, potentially overlooking IH injuries in other cases. To address this issue, we propose a non-invasive framework for IH detection that analyzes the morphology of cerebral blood flow velocity (CBFV) waveforms using non-invasive transcranial Doppler (TCD) ultrasound, thereby identifying and preventing IH injuries without the need for invasive procedures. Such a non-invasive framework could help detect IH injuries outside neurointensive care units and help provide timely treatment. The proposed methodology was evaluated on a cohort of 89 patients treated for various ICP-related conditions. Compared to previous frameworks based on semi-supervised learning of specific CBFV metrics, we found that using the raw waveform as input to a machine learning model improves the area under the ROC curve (AUC) to 96%. This is a significant improvement, as, by comparison, the pulsatility index (PI) achieved much lower accuracy in detecting IH at 59%.

1. Introduction

Intracranial pressure (ICP) measurement, an invasive procedure, is critical for patients with brain injuries at risk of cerebral ischemia and herniation caused by elevated ICP episodes. It is usually acquired by drilling a hole and placing a catheter beneath the skull. It can be used to guide medication or to normalize ICP through drainage, cerebrospinal fluid shunts, etc. Although this provides timely information, the invasive procedure is not only expensive but can also lead to other complications such as bleeding and infection.

Beyond the Neuro-ICU setting, knowing the ICP is also helpful in other scenarios, but the cost and risks usually preclude its use. Therefore, it is of great clinical necessity to develop a reliable, non-invasive method for predicting elevated ICP [1,2].

The analysis of cerebral blood flow velocity (CBFV) using transcranial Doppler monitoring provides a promising solution for this purpose. According to the Ursino model, ICP acts as the external pressure surrounding cerebral vessels; its elevation compresses the vasculature, thereby increasing vascular resistance and reducing compliance. These mechanical changes function as physiological modulators that shape the arterial pressure pulse into the observed CBFV waveform. Consequently, the pulse morphology of CBFV—such as its peak contours and decay

---

* Corresponding author.
E-mail address: mmwei@zut.edu.cn (M. Wei).

[1] This document is the result of the research project funded by the National Natural Science Foundation of China with Grant number 62301624, the Scientific Research Projects of Higher Education Institutions of Henan Province of China with Grant number 23A510012, and the Research Council of Lithuania under Grant Agreement No. SV5-40.

https://doi.org/10.1016/j.sbsr.2026.100999
Received 3 January 2026; Received in revised form 9 March 2026; Accepted 10 March 2026

patterns—serves as a deterministic surrogate for the state of intracranial pressure.

Some groups have used CBFV metrics, measured with a TCD, as estimators, including the following: maximum velocity, diastolic velocity, mean flow velocity, pulsatility index, resistance index [3,4], ratio of pulse slopes (RPS) [5], and peak appearance time (PAT) [6]. However, these still require further validation before meeting the requirements for clinical use.

Non-invasive intracranial pressure (nICP) assessment methods can be categorized into machine-learning and non-machine-learning approaches. Among these, non-machine learning-based non-invasive ICP (nICP) assessment primarily relied on mechanistic physiological models and empirical ultrasound indices. The most prevalent approach utilizes TCD-based indices such as the Pulsatility Index; however, its diagnostic reliability remains controversial [7], with studies reporting an Area Under the Curve (AUC) as low as 0.62 for detecting intracranial hypertension [8] due to its high sensitivity to systemic hemodynamic fluctuations [9]. More sophisticated non-ML approaches involve physiological modeling, such as the pseudo-Bayesian framework proposed by Imaduddin et al. [10], which achieved a root-mean-square error (RMSE) of 3.7 mmHg. Despite their high interpretability, these model-based methods are inherently constrained by the necessity of a pre-defined ICP prior to establish a baseline. This reliance on statistical priors may inadvertently bias the estimations, potentially masking individualized pathological variation. Furthermore, traditional morphological analyses, while capable of identifying key pulse features [11], often struggle with signal artifacts and inter-patient physiological variability, limiting their robustness in real-time clinical applications.

Recent advancements in non-invasive ICP monitoring have highlighted the potential of machine learning techniques applied to physiological signals. For instance, research by Megjhani et al. [12] utilized a domain adversarial neural network to estimate non-invasive ICP from blood pressure, electrocardiogram, and cerebral blood flow velocity data, reporting a mean median absolute error of 3.88 mmHg. Similarly, a study by Nair et al. [13] demonstrated that deep learning models, specifically temporal convolutional networks, could predict ICP waveforms from extracranial signals such as arterial blood pressure, electrocardiography, and photoplethysmography, achieving a median mean absolute error of 1.60 mmHg, offering a promising direction for non-invasive ICP monitoring. In the realm of photoplethysmography (PPG), studies have explored its efficacy in non-invasive ICP estimation. Bradley and Kyriacou [14] evaluated the use of PPG signals combined with classical machine learning methods to estimate continuous ICP values in patients with traumatic brain injury, highlighting the potential of PPG in this context.

These developments underscore a growing consensus on the value of multimodal approaches to noninvasive intracranial pressure (ICP) monitoring, which combines various physiological signals and advanced machine learning techniques to enhance the accuracy and reliability of ICP assessments. However, these methods focus on methodology and multimodal data while neglecting research into the purity and characteristics of single data sources.

Detection of intracranial hypertension is a classification problem that involves identifying elevated intracranial pressure (ICP) segments from standard ICP measurements. The traditional approach to such classification problems involves training a specific classifier using only labeled samples, a process known as supervised learning. The primary drawback of this method is its inability to leverage unlabeled samples, even when useful information can be learned from them to improve classification accuracy. The presence of unlabeled samples may stem from various reasons, such as the high cost of labeling all samples or the ambiguity of providing binary labels. For instance, one approach would be to label CBFV waveforms as IH samples if the corresponding ICP exceeds 20 mmHg, an accepted threshold for considering elevated ICP. A supervised learning algorithm is then employed to construct the classifier. Previously, it was considered that this approach might be overly rigid, making the detection of actual IH states heavily dependent on the relevance of using 20 mmHg as the threshold. However, selecting a different threshold is also challenging. If the threshold is set too high or too low, it may lead to missed IH diagnoses or generate excessive false positives.

To address issues such as missing labels or ambiguous thresholds in intracranial hypertension (IH) detection, semi-supervised learning approaches [8,14] can be effectively employed. Within the semi-supervised learning framework, model training does not require labeling all samples. Instead, it optimizes classifiers by simultaneously utilizing limited labeled data alongside vast amounts of unlabeled data. Current semi-supervised learning techniques encompass diverse types, including generative models, self-training, collaborative learning, Transductive Support Vector Machines (TSVM), and graph-based approaches [15,16]. Among them, graph-based learning techniques have drawn large attention due to their excellent performance and successful applications in visual and speech recognition [17,18]. In particular, the Spectral Regression technique proposed by Cai et al. has successfully overcome several drawbacks of conventional graph-based semi-supervised learning techniques by integrating regression with spectral graph analysis [19].

Kim et al. first proposed a semi-supervised learning framework for non-invasive intracranial pressure detection, introducing cerebral blood flow velocity (CBFV) morphological features extracted via the MOCAIP algorithm as model inputs. Compared to supervised learning, this approach improved detection accuracy by 10%, demonstrating that morphological features extracted by MOCAIP guide the model to focus more effectively on the intrinsic dynamic characteristics of the signal. Studies by Hu et al. [20] and Scalzo et al. [21] also demonstrated that key morphological information and detailed physiological characteristics provided by MOCAIP metrics can effectively enhance the detection capability of conditions such as intracranial hemorrhage. MOCAIP is capable of extracting 128 morphological indicators [22] from ICP and CBFV signals in each one-minute segment. These indicators quantify the pulse shape by calculating the amplitude of each characteristic point, the time intervals between them, the slopes of the rising and falling edges, and other relevant features. Likewise, MOCAIP is a fully automated pulse morphology analysis algorithm that enhances ICP signal quality, identifies artifact ICP pulses, and accurately detects the three distinct peaks of ICP pulses.

Our previous research has focused on quantitatively studying the CBFV and ICP pulse morphology by integrating a series of novel shape features [11]. Several of these features were highly correlated in an inter-subject manner. This discovery highlights the potential to develop alternative, noninvasive methods for ICP assessment by utilizing morphological analysis of CBFV waveforms. However, the necessity of data integrity, the cornerstone of such morphological analysis, and its impact on the robustness of results remain under-discussed in the existing literature.

To overcome the limitations of current noninvasive intracranial pressure (ICP) assessment techniques, we propose a noninvasive diagnostic tool for detecting intracranial hypertension (IH) that utilizes transcranial Doppler measurements of cerebral blood flow velocity (CBFV) at the middle cerebral arteries (MCAs).

In this paper, we revisit the design of machine-learning-based noninvasive ICP estimation by evaluating methods for processing the input signal before incorporating it into a machine-learning model.

Our goal is to evaluate whether machine learning models based on raw CBFV signals can outperform those established manually. Starting with time-aligned CBFV and ICP data, the CBFV samples were labeled based on their ICP pressure values, using a standard threshold. One of four different pre-processing methods was then applied to the CBFV data. Two preprocessing approaches used MOCAIP feature extraction, while two used raw waveforms. Both classification and regression machine learning models were selected based on their suitability for the application. Models were chosen that are capable of capturing moderate complexity, as morphological features are more intricate than simple curve fitting. Additionally, the relatively small dataset size was deemed unsuitable for more complex models such as convolutional neural networks (CNNs). Once trained on the preprocessed data, the classification models generate a binary decision indicating either normal or high pressure, while the regression models produce a continuous estimate of the ICP pressure in millimeters of mercury (mmHg).

The principle of noninvasive IH detection using CBFV data is illustrated in Fig. 1. There is a correlation between the ICP waveform and the CBFV waveform, which has been identified in many papers [5,6]. By leveraging this correlation, a patient's intracranial hypertension can be detected using machine learning methods on CBFV data. It can be expressed as a classification problem or a regression problem. For classification, the machine learning algorithm outputs either "IH" or "not IH". For regression, the machine learning algorithm outputs the probability that the patient has intracranial hypertension.

2. Methods

This section describes the methodology we proposed for noninvasive IH detection using the CBFV waveform. The subsequent subsections provide a detailed description of various aspects of the study, including the dataset employed, the different data pre-processing steps adopted, the machine learning algorithms, and the experimental framework.

2.1. Dataset acquisition

The dataset originates from the University of California, Los Angeles (UCLA) Medical Center and was approved by the institutional review board (IRB) for use in this study. This is a retrospective study on patients who were being treated for various ICP related conditions, including idiopathic IH, Chiari syndrome, and slit ventricle patients with clamped shunts. The dataset comprises 638 episodes collected from 89 patients in which ICP monitoring was indicated and performed concurrently with TCD recording. Among these recorded episodes, 548 are normal and 90 are hypertensive (ICP $>$ 20 mmHg).

Invasive ICP was monitored continuously for clinical management using either intraventricular catheters (for TBI patients) or intraparenchymal microsensors (for NPH patients). Simultaneously, cerebral hemodynamics were assessed daily by diagnostic medical sonographers using transcranial Doppler (TCD). CBFV signals were obtained from the middle cerebral artery (MCA) ipsilateral to the ICP monitoring site using hand-held 2 MHz probes. Due to the manual nature of the ultrasound assessment, recording sessions typically lasted 3 to 5 min each.

All physiological signals, including CBFV, invasive ICP, and arterial blood pressure (from bedside GE monitors), were archived via a mobile cart equipped with a PowerLab data acquisition system (ADInstruments, Colorado Springs, CO). The system performed high-fidelity analog-to-digital conversion at a sampling rate of 400 Hz. The synchronized datasets were archived in Chart™ binary format, ensuring precise temporal alignment between the intracranial pressure variations and the corresponding pulse waveforms for subsequent semi-supervised learning analysis.

An expert researcher retrospectively identified IH episodes and annotated the time of onset of elevation, elevation plateau, and invasive cerebrospinal fluid drainage in each patient's recording. Ten patients were excluded due to signal drop and artifacts that prevented the expert from identifying IH episodes with high confidence. Each segment included 20 min of data, capturing the transition from a normal state ($0 \sim 20mmHg$) to an elevated ICP ($> 20mmHg$). The segments were time-aligned to contain 15 min of data before the plateau and 5 min after.

The temporal alignment of 15 min pre-onset and 5 min post-onset was strategically chosen to model the dynamic evolution of intracranial dynamics. The 15-min 'normal' segment serves as a steady-state baseline, allowing the semi-supervised algorithm to capture the progressive deterioration of pulse compliance before the $20mmHg$ threshold is crossed. To ensure the purity of the 'high ICP' labels, the expert review process prioritized the exclusion of transient spikes (< 1 min) that lacked the characteristic rounding of the CBFV waveform. This 20-min standardized windowing also ensures that the segments remain unaffected by rapid clinical interventions that typically follow sustained hypertensive episodes.

Before normalization, the labels were generated based on each beat's pressure level: 1 for levels above 20 mmHg and 0 for those below, resulting in a binary classification problem.

2.2. Pre-processing

Since the training of a machine learning model relies heavily on the quality of its training data, pre-processing steps were chosen for their ability to normalize the data. One set of steps was performed to compare with similar work on predicting ICP levels from CBFV waveforms [8].

2.2.1. Beat extraction

To make a meaningful comparison with past works, the same unprocessed binary data files were used as in [8]. In our comparisons, the original MOCAIP algorithm was used to extract individual beats from

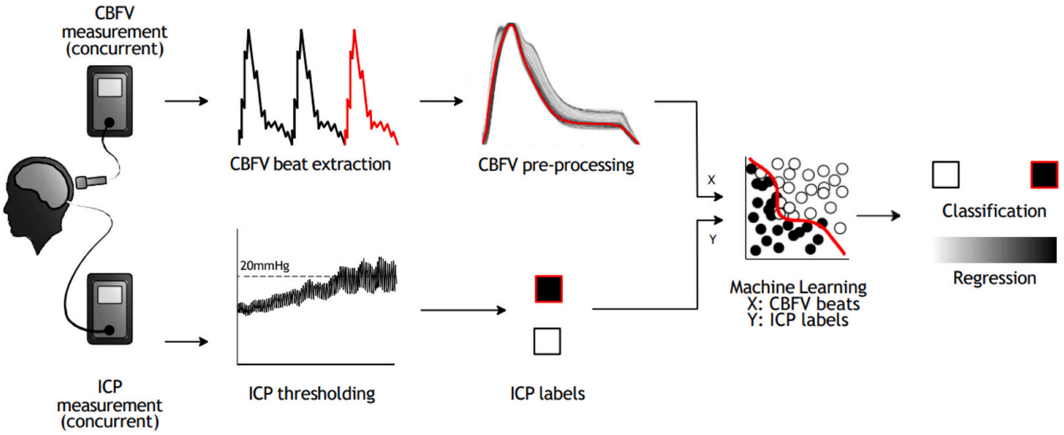


Fig. 1. A step-by-step visualization of CBFV and ICP acquisition and analysis necessary to create a machine learning model. CBFV is measured noninvasively using a TCD, while ICP is measured invasively using a catheter inserted into the right frontal lobe of the brain. The patient recordings are extracted, manually reviewed, and annotated to mark the onset time of IH. The beat data is then extracted from these annotated recordings and pre-processed. The average ICP level is thresholded and assigned a binary label. The pre-processed CBFV and ICP data are then used in a machine learning algorithm. Through experimentation, it has been demonstrated that machine learning offers a clinically relevant approach for the noninvasive detection of IH using CBFV data.

CBFV data, then 128 morphological features for each pulse beat were obtained.

2.2.2. Labeling strategy

Because the labels expected by the supervised and semi-supervised SR-KDA algorithm cannot be negative, labels were assigned as 0 for normal pressure and 1 for high pressure. Otherwise, if labels were assigned as −1 or 1, side effects might have biased the results. The premise of this work is that CBFV waveforms will change together with ICP.

While the morphological features were initially extracted from individual pulse beats using the MOCAIP algorithm, the final unit of analysis for our classification model was a 1-min segment. This hierarchical approach was adopted to enhance the robustness of the estimation against beat-to-beat variability and transient artifacts. Specifically, for each 1-min window, the 128 metrics derived from all identified pulses were averaged to produce a single representative feature vector. The ground-truth label for each segment was then assigned based on the mean value of the invasive ICP recording within that same minute, using 20 mmHg as the threshold for intracranial hypertension. Consequently, each sample processed by the proposed framework represents a stable 1-min physiological state rather than an isolated pulse.

2.2.3. Fixed heartbeat resampling

The learning models used require a fixed-length input. CBFV waveforms were first resampled to a fixed length because their length depends on the patient's heart rate, which is highly variable. Each beat sample $\overrightarrow{S}_i \in S$ was resampled to a vector of 400 values $\overrightarrow{Z}_i$. A left shift was then performed to align the beats. We define the alignment point to be the minimum of each beat waveform.

$$start = \underset{i}{\operatorname{argmin}}\left(\overrightarrow{Z}_i\right) \tag{1}$$

and perform a circular shift to set this point as the first element of each beat vector, where n is the length of $\overrightarrow{Z}_i$.

$$\overrightarrow{N}_i = \overrightarrow{Z}_{(\mathrm{i+start})\mathrm{mod}(\mathrm{n})} \tag{2}$$

2.2.4. AUC normalization

AUC Normalization is performed to eliminate amplitude-scale errors caused by equipment loosening, patient movement, or coughing, which often introduces transient artifacts that manifest as sudden shifts in signal magnitude. To mitigate this impact, each sample $\overrightarrow{X}_i$ is normalized so that its AUC equals 1.

$$\overrightarrow{X}_{i,j} = \frac{\overrightarrow{X}_{i,j}}{\sum_{j=1}^{J}\overrightarrow{X}_{i,j}} \tag{3}$$

In this way, we effectively suppress these non-physiological amplitude fluctuations and ensure that the subsequent feature extraction of machine learning methods focuses on the intrinsic morphology of waveforms, which is more robust for detecting intracranial hypertension.

2.2.5. Mean and STD normalization

Because the beats are now all a uniform length and time-aligned to their minimum value, differences between each time feature can be emphasized through mean and standard deviation normalization. The mean of corresponding time points $\overrightarrow{X}_j$

$$\overline{\overrightarrow{X}_j} = \frac{\sum_i \overrightarrow{X}_{i,j}}{I} \tag{4}$$

is used to center the data.

$$\overrightarrow{X}_{i,j} = \overrightarrow{X}_{i,j} - \frac{\overrightarrow{X}_{i,j}}{\overline{\overrightarrow{X}_j}} \tag{5}$$

The standard deviation σ_j of corresponding time points $\overrightarrow{X}_j$

$$\sigma_j = \sqrt{\frac{1}{I-1}\sum_{i=1}^{I}\left(\overrightarrow{X}_{i,j} - \overline{\overrightarrow{X}_j}\right)^2} \tag{6}$$

was then used for normalization.

$$\overrightarrow{X}_{i,j} = \overrightarrow{X}_{i,j} - \sigma_j \tag{7}$$

2.2.6. 0 and 1 normalization

Some models, such as SR-KDA and SVM, require the features in a sample to be in the range [0,1].

This range has a negligible impact on other models with less constrained input requirements.

$$\overrightarrow{X}_{i,j} = \frac{\overrightarrow{X}_{i,j} - \min\left(\overrightarrow{X}_{:,j}\right)}{\max\left(\overrightarrow{X}_{:,j}\right) - \min\left(\overrightarrow{X}_{:,j}\right)} \tag{8}$$

2.2.7. Beat clustering

A final clustering step is performed to remove noise by grouping similar beats and retaining common features. There are various sources of noise, including a patient coughing and a sensor coming loose. A hierarchical clustering method is used to first separate CBFV beats into groups based on morphological distance, then average the groups within each 1-min segment. The distance d between any two samples $\overrightarrow{X}_i$ and $\overrightarrow{X}_j$ is calculated as

$$d\left(\overrightarrow{X}_i, \overrightarrow{X}_j\right) = \sum_{k=1}^{\left|\overrightarrow{X}_i\right|}\left(\overrightarrow{X}_{i,k} - \overrightarrow{X}_{j,k}\right)^2 \tag{9}$$

where the lengths of the samples are uniform. To ensure the purity of the physiological signals, we iteratively compute the Silhouette value for each resulting cluster to assess its internal similarity. The cluster exhibiting the highest Silhouette value, indicating the most similarity and typical physiological patterns, is retained for analysis. Conversely, Atypical clusters, characterized by low similarity and often resulting from noise or artifacts, are discarded. The beats within the dominant, high-cohesion cluster are then averaged to generate a single representative mean pulse for the segment. A deeper explanation of the Silhouette criterion can be found in [23].

2.3. Machine learning models

The following five supervised machine learning models were selected as the best suited for this ICP classification task. The training data contains complexities that can't be modeled with linear regression, but the dataset is not large enough to leverage the power of deep learning models. The default MATLAB methods were used for the Random Forest, SVM, and Neural Network models. Fig. 1 visualizes the process of obtaining and analyzing CBFV and ICP data to develop a machine learning model for the noninvasive detection of IH.

2.3.1. Random Forest

The Random Forest model applies ensemble learning to the classic decision tree model [24]. By combining individual decision trees, random forest avoids the overfitting problem that typically occurs in a single decision tree. There are several methods for randomizing the combination. The simplest approach is to bag by growing each tree with a random selection of samples from the training set. Another way is random split selection, randomly selected from the K best splits at every node of a single tree. An approach to improve randomization is to weight samples in the training set. There are also several methods for randomizing the subspace where a subset of features is selected to grow trees.

A random forest based on bagging is adopted for comparison in this paper. Suppose that there are an ensemble of trees $t_1(x)$, $t_2(x)$, …, $t_m(x)$, m is the number of trees in random forest. The training set is ($x_i \in X$, $y_i \in Y$),$i = 1, 2, \ldots, n$, which are randomly drawn from distribution Z. For each tree, ($s < n$) samples are selected withreplacement from the total training set. $t_1(x)$, $t_2(x)$, …, $t_m(x)$ are trained independently on their own training sets. Then, the prediction for a new sample is the average of the predictions from all trained regression trees.

$$\widehat{t} = \frac{1}{m}\sum_{j=1}^{m} t_j(x') \tag{10}$$

For classification, it is made by the majority vote of all trained classification trees.

2.3.2. SR-DA

Spectral regression discriminant analysis (SR-DA) is an efficient method for performing linear discriminant analysis that replaces eigenvector computations with several regularized least squares problems. By utilizing a spectral graph, SD-RA transforms discriminant analysis into a regression problem. The regression process of SD-RA is to find a transformation matrix which maps the data $x_i \in X$ to labels $\widehat{y}_i \in \widehat{Y}$, through minimizing the sum of the square error:

$$\Phi = \sum_{i,j=1}^{n} \left(\widehat{y}_i - \widehat{y}_j\right)^2 W_{i,j} \tag{11}$$

where $W \in \mathbb{R}^{n\times n}$ is the n weight matrix of the adjacent graph to denote the similarity of the input samples. $W_{i,j}$ is the edge weight between $\overrightarrow{X}_i$ and x_j. If x_i and x_j belong to the same class, $0 < w_{i,j} < 1$, otherwise, $W_{i,j} = 0$.

Further, eigen decomposition is made by

$$Wv = \lambda Dv \tag{12}$$

Where D is a diagonal matrix whose diagonal element $D_{i,j} = \sum_j W_{i,j}$, the row or column sums of the symmetrical matrix W. $v = [v_1, v_2, \ldots, v_n]$ includes n eigenvectors corresponding to eigen values $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_n]$.

Then, the projection vectors are obtained by

$$\widehat{\gamma} = argmin_{\gamma} \sum_{i=1}^{n} \left(\gamma^T x_i - y_i^j\right)^2 \tag{13}$$

Where y_i^j is the i-th element of v_j. The transformation matrix is $\gamma = [\gamma_1, \gamma_2, \ldots, \gamma_n]$

2.3.3. SR-KDA

Spectral Regression is still a linear regression algorithm, which limits its application to nonlinear discriminant analysis. However, this limitation can be overcome using a "kernel trick". By kernel projection, SD-KRA mapped the data samples into a feature space. The nonlinear mapping can be expressed as:

$$\phi : \mathbb{R}^n \rightarrow \mathscr{F} \tag{14}$$

The corresponding Eigen-formula becomes:

$$KWK\mathrm{v} = \lambda KK\mathrm{v} \tag{15}$$

The kernel we adopted is the Gaussian Radial Basis Function (RBF) kernel

$$K\left(x_i, x_j\right) = e^{\left(-\frac{\left\|x_i - x_j\right\|^2}{2\sigma^2}\right)} \tag{16}$$

Then, the spectral regression becomes to minimize:

$$\min_{f\in\mathscr{F}} \sum_{i=1}^{n} \left(f(x_i) - y_i\right)^2 \tag{17}$$

Where, $\mathrm{f}(\mathrm{x}) = \sum_{\mathrm{i}=1}^{\mathrm{n}} v_i K(\mathbf{x}, x_i)$ represents the projection function. This technique is regarded as spectral regression kernel discriminant analysis (SR-KDA) [25].

2.3.4. Neural network

A neural network is a powerful machine learning algorithm that mimics the workings of the human brain and the information transfer mechanisms among biological neurons. It has demonstrated excellent performance in recognizing handwriting, speech, voice, and movement, as well as image and video analysis, and system control. It is typically composed of an input layer, one or more hidden layers, and an output layer. As the number of layers increases, deep neural networks (DNNs) are generated. Sometimes, a network includes hundreds of layers. A feedforward net is trained in this paper by minimizing the loss function Φ:

$$\Phi = \sum_{i=1}^{N} e_i^2(x) \tag{18}$$

Where (**x**)is the error between the output **y** of the network and the true value **y** in the training set for input vector **x**, N is the dimension of vector **y**.

The Marquardt-Levenberg algorithm is applied to solve this minimization problem. It is an easy and fast way to train moderate-sized networks. The update equation of a neuron is

$$\Delta x = \left[J^T(x)J(x) + \mu I\right]^{-1} J^T(x)e(x) \tag{19}$$

Where μ is the learning rate. (**x**) is the Jacobian matrix, and J(**x**) is the transposed matrix of J(**x**). During the training process, each time the loss function decreases, μ is multiplied by α, where $0 < \alpha < 1$ is the learning rate. By adjusting the network's parameters, the optimal weights and thresholds are determined when the point of convergence is reached. The prediction for new input samples can then be obtained from the network model.

2.3.5. SVM

The support vector machine (SVM) is a popular machine learning method in pattern recognition and intelligent systems, known for its excellent generalization performance. SVM aims to find a hyperplane to separate positive samples from negative samples. The training set are labeled as$\{x_i, y_i\}$, $i = 1, 2, \ldots, n$,$y\in \{-1,1\}$. Suppose that the set is separable, and all the training samples conform to the following constraints:

$$w\cdot x_i + b \geq +1 \ \mathit{for}\ \mathrm{y}_i = +1$$

$$w\cdot x_i + b \leq +1 \ \mathit{for}\ y_i = -1 \tag{20}$$

The closest points lie on the planes $\mathbf{w}\cdot x_i + b = \pm 1$. Thus margin between two kinds of samples is $=\frac{1}{\|w\|_2}$. The optimal model can be obtained by maximizing the margin:

$$\max_{w,b} \frac{1}{\|w\|_2}, s.t. y_i(w\cdot x_i + b) \geq 1 \ , \ \forall i \tag{21}$$

Then the output of new input samples can be predicted by the trained model with optimized parameters **w** and b.

2.4. Experiments

The primary purpose of our experiments is to investigate whether machine learning models are advantageous for predicting ICP levels from corresponding CBFV data. The secondary purpose is to determine whether the classification accuracy of machine learning models increases when trained on raw waveforms of CBFV beats, as opposed to first extracting features, such as those extracted in the MOCAIP feature extraction algorithm.

The experiments are evaluated using 10-fold cross-validation at the patient level, with AUC-ROC and Precision-Recall metrics computed and confidence intervals reported for each result. The AUC of the ROC curve is known to be affected by the large discrepancy between the number of positive and negative labels. While we do not directly address the class imbalance issue, we introduce an additional accuracy metric based on the area under the precision-recall curve, which is less sensitive to class imbalance and is recommended in this case.

We compare four methodologies to thoroughly investigate and verify the impact of data pre-processing methods on the performance of correctly classifying IH episodes from CBFV waveforms. Experiment one (uncorrected MOCAIP) and experiment two (manually corrected MOCAIP), respectively, evaluated the effectiveness of automated/uncorrected and manually corrected MOCAIP features in providing waveform information and their impact on the performance of five machine learning algorithms. The goal of these two experiments is to verify the adequacy of the original feature extraction method and to assess the consistent effect of the corrected features. Subsequently, the third experiment utilized the clustered CBFV waveform as model input. By comparing the results of the third experiment with those using MOCAIP features as input in experiments 1 and 2, we can verify whether waveform clusters can provide higher accuracy in IH detection. Experiment 4 (normalization and clustering of raw waveforms) further reduces the possibility of losing information by using raw waveform data. The specific details of the four experiments are discussed in detail in the following paragraphs.

2.4.1. Uncorrected MOCAIP

Uncorrected MOCAIP metrics are used to evaluate classifier accuracy using ROC and precision-recall curves. No manual correction was performed, as shown in Fig. 2. The metrics were normalized as described by Eqs. (5) and (7). The entire dataset was then normalized to a range of 0 to 1. This is described by Eq. (8).

2.4.2. Manually corrected MOCAIP

The second experiment was designed to replicate results found in a paper by Kim et al. [8]. Classifier accuracy is evaluated using ROC and precision-recall curves based on manually corrected MOCAIP metrics in Fig. 2. It was unclear how much impact the manually corrected metrics had, so the trial with uncorrected MOCAIP data serves as a reference. The same dataset was used as in Kim et al.'s study, except that the labeling strategy differed.

2.4.3. Normalization of raw beats

The extended set of MOCAIP metrics is somewhat sensitive to outliers because it is based on ratios of basic features. This also tends to amplify the noise. The classification accuracy was evaluated by using the average CBFV waveform of the 1-min segment to be classified. In Fig. 3, each CBFV beat was pre-processed so that it was resampled to a fixed heart-rate (Eqs. (1) and (2)) and normalized so that the area under the CBFV curve was 1 (Eq. (3)), which effectively decouples the morphological shape (the physiological signal) from the absolute magnitude fluctuations. Then, each feature was normalized using mean and standard deviation (Eqs. (5)–(7)), which is employed to suppress disturbances such as electromagnetic interference, resulting in a smoother and more consistent waveform representation for subsequent analysis.

2.4.4. Normalization and clustering of raw beats

In addition to the previously described normalization steps, we further explored a clustering approach to characterize the CBFV waveform of a segment shown in Fig. 4. This was done to mitigate the impact of outliers that might still influence the average waveform calculated using a trimmed mean.

3. Results

Experiments were run using 10-fold cross-validation at the patient level. Classification results are reported as the area under the receiver operating characteristic curve (AUC), which is commonly reported and informative on balanced datasets, and the area under the precision-recall curve (AUPR), which is more informative due to the class imbalance in the dataset. Regression results are evaluated by root mean square error (RMSE), bias, and standard error (std).

The classification results of the six methods are compared in Table 1. Table 1 contains four tables, each corresponding to a different dataset. By comparing the four tables, the classification result of SR-KDA on clustered waveforms is the highest, regardless of the algorithms or datasets. It achieved this by using a 1-min average of beat segments or waveforms, which were then resampled to a fixed heart rate. These segments were normalized to a CBFV AUC curve of 1, feature-normalized using mean and standard deviation, and finally clustered based on segment characteristics. With these pre-processing steps, SR-

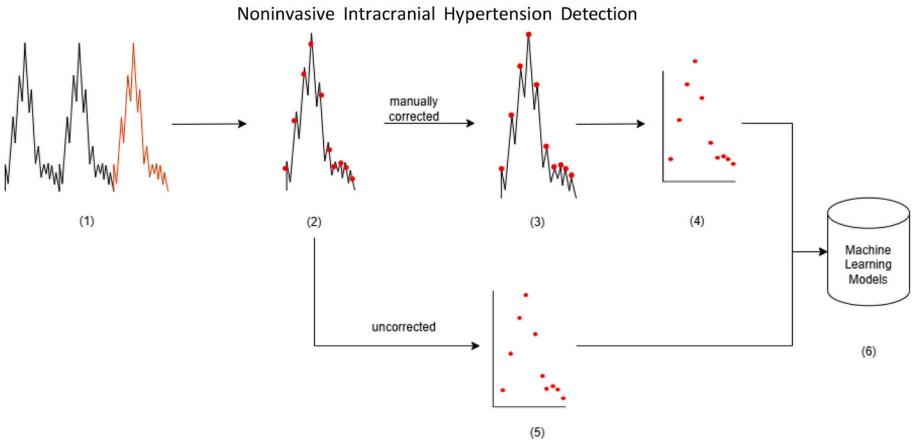


Fig. 2. A visualization of manually corrected MOCAIP and uncorrected MOCAIP preprocessing steps. (1) Raw CBFV waveforms acquired over one-minute segments. (2) MOCAIP metrics detect wave peaks and extract morphological features from the waveforms. (3) Manually correct MOCAIP metrics and extract morphological features from the waveforms. (4) Standardize and normalize manually corrected MOCAIP features to a 0–1 range. (5) Standardize and normalize uncorrected MOCAIP features to a 0–1 range. (6) Input into machine learning models.

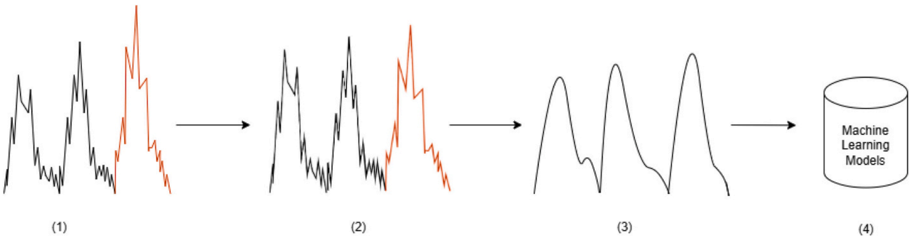


Fig. 3. A visualization of the normalization of the raw beats preprocessing steps. (1) Raw CBFV waveforms acquired over one-minute segments. (2) Normalize the area under the CBFV curve (AUC) to 1. (3) Normalize the features using mean and standard deviation. (4) Input into machine learning models.

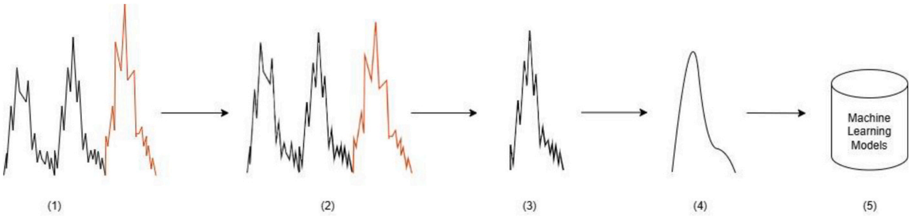


Fig. 4. A visualization of normalization and clustering of raw beats preprocessing steps. (1) Raw CBFV waveforms acquired over one-minute segments. (2) Normalize the area under the CBFV curve (AUC) to 1 for every segment. (3) Cluster the raw beats data and take the average. (4) Normalize the features using mean and standard deviation. (5) Input into machine learning models.

Table 1
Performance results of various pre-processing configurations and classification models. Table (a) shows the area under the ROC curve (AUC-ROC) accuracy of the models using MOCAIP metrics. Table (b) shows the area under the ROC curve (AUC-ROC) accuracy of the models using manually corrected MOCAIP metrics. Table (c) shows the area under the ROC curve (AUC-ROC) accuracy of the models using the (trimmed) 1-min average waveform. Table (d) shows the area under the ROC curve (AUC-ROC) accuracy of the models using a 1-min dominant cluster waveform.

	AUC-ROC (CI)	AUC-PR (CI)
(a) MOCAIP metrics		
Random Forests	**94.61% [92.37%, 96.84%]**	**80.44% [69.25%, 91.62%]**
SR-KDA	94.03% [91.35%, 96.72%]	73.57% [65.56%, 81.58%]
SR-DA	79.92% [76.08%, 83.76%]	46.39% [36.14%, 56.64%]
Neural Network	81.64% [75.29%, 87.98%]	47.13% [38.82%, 56.64%]
SVM	73.44% [68.16%, 78.73%]	39.49% [29.83%, 56.64%]
CNN	60.42% [50.69%, 70.14%]	14.54% [11.84%, 56.64%]
(b) Manual MOCAIP metrics		
	AUC-ROC (CI)	AUC-PR (CI)
Random Forests	**96.07% [94.23%, 97.92%]**	**85.98% [76.53%, 95.44%]**
SR-KDA	93.28% [89.45%, 97.12%]	76.38% [69.67%, 83.09%]
SR-DA	81.11% [76.92%, 85.30%]	53.65% [45.94%, 61.37%]
Neural Network	73.10% [65.96%, 80.24%]	38.93% [29.32%, 61.37%]
SVM	74.12% [70.39%, 77.85%]	45.46% [38.60%, 61.37%]
CNN	56.45% [53.35%, 59.55%]	17.72% [15.09%, 61.37%]
(c) Clustered Waveform		
	AUC-ROC (CI)	AUC-PR (CI)
Random Forests	95.73% [93.66%, 97.79%]	84.95% [77.44%, 92.46%]
SR-KDA	**96.10% [93.43%, 98.77%]**	**88.26% [82.95%, 93.58%]**
SR-DA	80.00% [74.62%, 85.39%]	42.49% [33.69%, 51.28%]
Neural Network	77.36% [69.63%, 85.09%]	41.72% [31.29%, 51.28%]
SVM	73.06% [65.73%, 80.40%]	37.33% [29.02%, 51.28%]
CNN	64.01% [60.73%, 67.28%]	22.12% [20.00%, 51.28%]
(d) Raw Waveform		
	AUC-ROC (CI)	AUC-PR (CI)
Random Forests	95.98% [93.82%, 98.14%]	87.07% [79.39%, 94.75%]
SR-KDA	**96.92% [94.55%, 99.30%]**	**83.88% [77.63%, 90.13%]**
SR-DA	82.60% [78.10%, 87.11%]	45.64% [36.43%, 54.85%]
Neural Network	85.60% [82.66%, 88.54%]	63.47% [53.63%, 54.85%]
SVM	76.26% [70.88%, 81.65%]	40.98% [32.14%, 54.85%]
CNN	60.83% [55.13%, 66.53%]	20.19% [15.27%, 54.85%]

KDA differentiated hypertensive and control episodes thresholded at 20 mmHg with an AUC-ROC of 96.92% ± 2.37% and an AUC-PR of 83.88% ± 6.25%. This is in contrast to the results of the model based on manually corrected MOCAIP metrics, which achieved an AUC of 95.59% ± 2.39% and an AUC-PR of 84.11% ± 10.27%, respectively.

Table 1(a) and Table 1(b) show that Random Forest achieves the best performance on MOCAIP metrics, regardless of whether the metrics are manually corrected. In Table 1(c) and Table 1(d), AUC-ROC and AUC-PR of SR-KDA are larger than those of the other four methods, except for AUC-PR of Random forests on the raw waveform. This inconsistency is because AUC-PR is more sensitive to data imbalance and more reliable than AUC-ROC. However, the AUC-ROC and AUC-PR of SR-KDA on the waveform remain higher than those on the MOCAIP metrics, as indicated by the comparison of the four tables. The AUC-ROC and AUC-PR for SR-KDA have been improved by using waveforms rather than MOCAIP metrics. Furthermore, a comprehensive comparison across all experimental setups shows that the combination of SR-KDA and raw waveforms yields superior classification performance.

Table 2 presents the regression results for different preprocessing configurations and models. In regression tasks, the SR-KDA model again outperformed other models, achieving a root mean square error (RMSE) of 3.45 mmHg, standard deviation (SD) of 0.42 mmHg, bias of − 0.05 mmHg, and mean absolute error (MAE) of 2.13 mmHg. This result also indicates that SR-KDA delivers the optimal regression performance. Although the Random Forest model demonstrated comparable accuracy on the MOCAIP metric (root mean square error of 3.87 mmHg, standard deviation of 0.57 mmHg, bias of 0.003 mmHg, and mean absolute error of 2.56 mmHg), its estimation accuracy remains slightly inferior to that of the SR-KDA model. The consistency between the regression and classification outcomes provides robust evidence that the waveform-based SV-KDA framework significantly outperforms the Random Forest model utilizing MOCAIP metrics.

The ROC curve and precision-recall curve of the six algorithms are compared in Figs. 5 and 6. It can be seen that the curves for random forests and SR-KDA are all above those of other algorithms. Especially in Fig. 6, when waveforms are applied during training for random forests and SR-KDA, the vertices of the curves are closer to the corners of the figure.

Given that a new sample arrives at most every 0.5 s, the SR-KDA model can run in real-time, with a classification time of 0.007 s. While

Table 2

Performance results of various pre-processing configurations and regression models. Table (a) shows the RMSE of the models using various pre-processing configurations. Table (b) shows the bias of the models using various pre-processing configurations. Table (c) shows the standard deviation of the models using various pre-processing configurations.

	MOCAIP	Manual MOCAIP	Clustered Waveform	Raw Waveform
(a) RMSE ± SD(mmHg)				
Random Forests	**3.87 ± 0.57**	4.31 ± 0.55	4.68 ± 0.72	4.33 ± 0.46
SR-KDA	5.93 ± 1.95	5.24 ± 1.36	**4.27 ± 0.73**	**3.45 ± 0.42**
SR-DA	6.57 ± 0.49	7.36 ± 0.76	8.41 ± 0.78	8.04 ± 0.70
Neural Network	7.19 ± 0.83	7.90 ± 1.22	8.20 ± 1.75	7.00 ± 0.92
SVM	6.91 ± 0.49	8.30 ± 1.26	8.84 ± 1.65	8.27 ± 1.28
(b) Bias ± SD(mmHg)				
	MOCAIP	Manual MOCAIP	Clustered Waveform	Raw Waveform
Random Forests	**0.003 ± 0.47**	−0.02 ± 0.31	−0.08 ± 0.49	−0.04 ± 0.44
SR-KDA	0.05 ± 0.99	0.09 ± 0.81	−0.06 ± 0.55	**−0.05 ± 0.28**
SR-DA	−0.02 ± 0.81	0.04 ± 0.61	0.06 ± 0.67	0.02 ± 0.62
Neural Network	−0.11 ± 0.82	−0.13 ± 0.77	0.29 ± 0.69	0.13 ± 0.73
SVM	−0.05 ± 0.79	−0.34 ± 0.63	−1.06 ± 0.79	−0.42 ± 0.66
(c) MAE ± SD(mmHg)				
	MOCAIP	Manual MOCAIP	Clustered Waveform	Raw Waveform
Random Forests	**2.56 ± 0.32**	2.75 ± 0.31	2.92 ± 0.32	2.80 ± 0.29
SR-KDA	3.41 ± 0.41	3.29 ± 0.52	2.88 ± 0.34	**2.13 ± 0.18**
SR-DA	4.88 ± 0.34	5.35 ± 0.47	5.67 ± 0.53	5.70 ± 0.54
Neural Network	5.23 ± 0.39	5.73 ± 0.74	5.53 ± 0.69	5.15 ± 0.60
SVM	5.03 ± 0.43	5.61 ± 0.66	5.70 ± 0.76	5.50 ± 0.68

the total runtime may vary due to pre-processing steps, the classification time remains constant at 0.007 s when using SR-KDA.

4. Discussion

Our findings demonstrate that utilizing raw cerebral blood flow velocity waveforms significantly enhances the detection accuracy of intracranial hypertension compared to traditional methods based on pre-established morphological features. The 10.63% range observed in a 10-fold cross-validation could suggest some degree of overfitting in the classifier due to the low sample size. However, this range is markedly smaller than the 20.54% range observed in the corrected MOCAIP experiment, indicating that our method reduces overfitting compared to models that rely on MOCAIP-derived metrics. Furthermore, the significantly higher minimum precision-recall accuracies highlight improvements in the model's worst-case performance, underscoring its robustness in detecting IH under varied conditions.

It is noteworthy that our experiments employed default settings without parameter fine-tuning. This indicates room for optimization that could further enhance performance. However, parameter fine-tuning introduces the risk of overfitting and must be approached with caution, particularly in medical applications where generalizability is critical.

Interestingly, our attempts to implement both a Convolutional Neural Network (CNN) and Multiple Kernel Learning (MKL) failed to achieve satisfactory convergence. While MKL possesses a multi-layered fusion structure conceptually similar to deep learning, it lacks the robust optimization capabilities of backpropagation in deep learning. Consequently, these complex models tend to overfit to individual noise rather than generalizing across physiological patterns. This comparison highlights the advantage of our semi-supervised scheme, which ensures stable performance in data-constrained clinical scenarios.

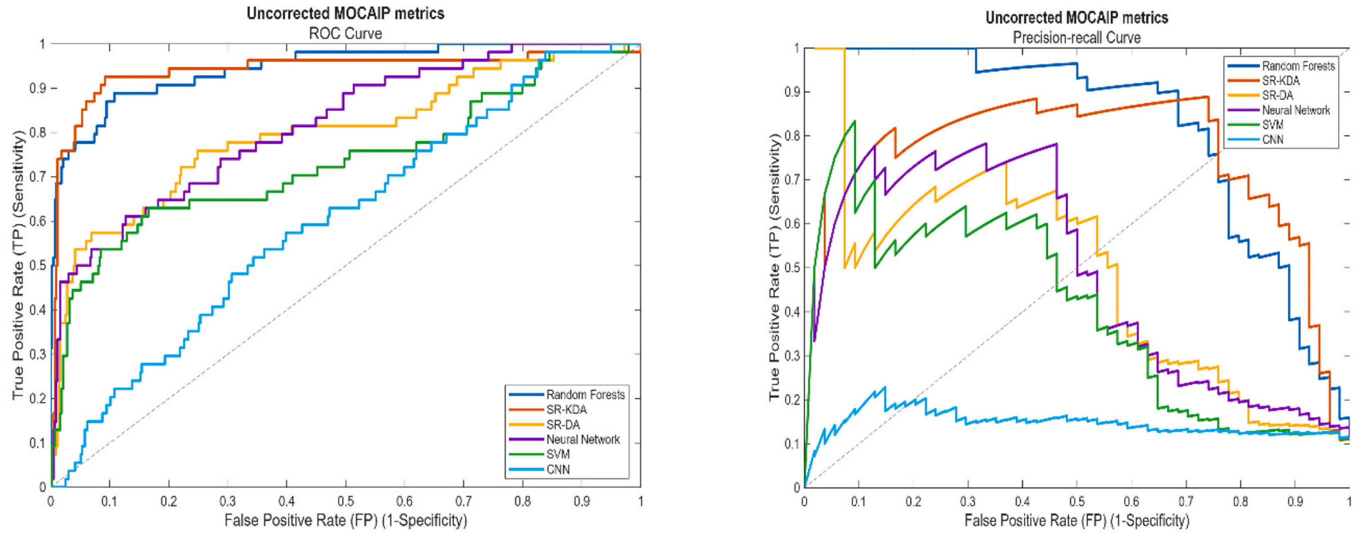
By directly analyzing raw CBFV waveform data using machine learning algorithms, our study not only improved detection accuracy for IH but also demonstrated the model's ability to generalize to complex medical data. Unlike traditional methods that rely on pre-processed features, such as MOCAIP metrics, our approach captures a broader spectrum of hidden patterns and features within the raw waveform data—patterns that are often imperceptible to human intuition. Importantly, this method avoids the information loss that occurs during feature extraction in traditional approaches, thereby preserving clinically relevant information. These diverse and abundant features contribute to the model's improved generalizability while reducing the risk of overfitting. The relatively narrow accuracy range in cross-validation further underscores the stability of our predictive model across different datasets. This generalization capability is particularly important for clinical applications, as overfitting not only compromises the practicality of machine learning models but also increases the risk of misdiagnosis.

4.1. *Relation to contemporary research*

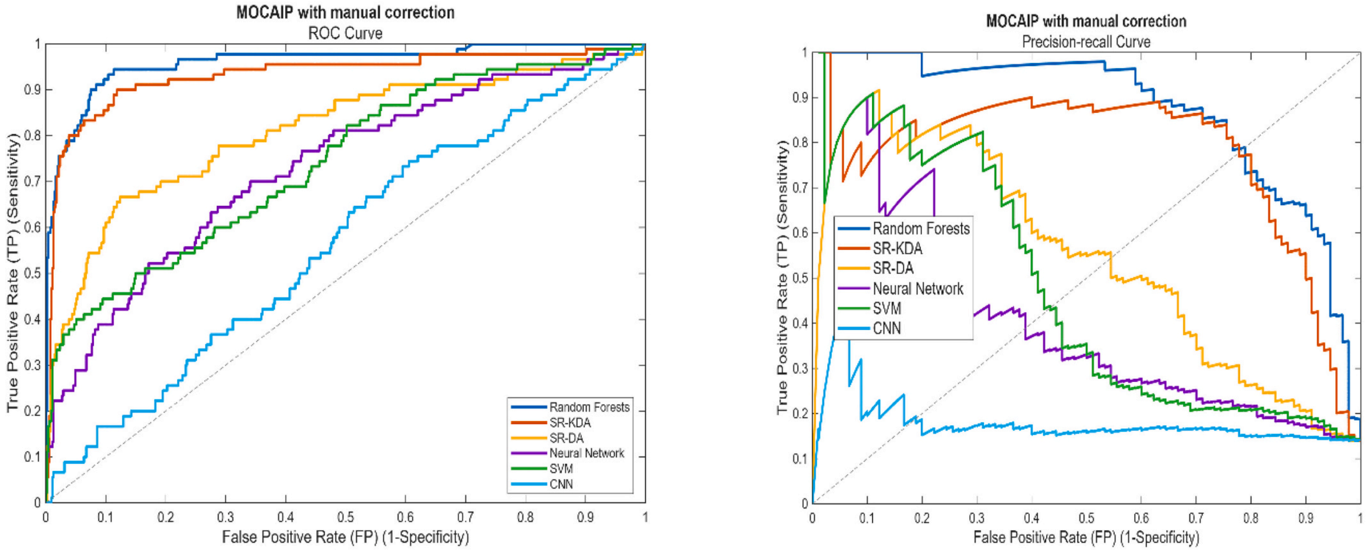
Our findings align with the growing body of literature emphasizing the importance of utilizing raw physiological signals, such as cerebral blood flow velocity (CBFV) waveforms, in non-invasive intracranial pressure (ICP) monitoring. Recent studies have also demonstrated the efficacy of using raw signal data rather than relying on pre-established features, underscoring the value of machine learning models capable of detecting subtle patterns that are not immediately visible to clinicians. In a similar vein, recent work by Nair et al. [13] utilized deep learning models to predict ICP using extravasated signals like electrocardiogram (ECG) and blood pressure without feature extraction.

resulting in improved model robustness and accuracy. Their approach highlighted the ability of raw data to preserve more granular details of physiological states that pre-processed features might miss. The issue of overfitting on small datasets has been well-documented in medical machine learning. Studies like those by Megjhani et al. [12] have explored techniques to mitigate overfitting in ICP prediction models, with a particular focus on regularization methods and cross-validation strategies. The 10.63% range observed in our study falls within the expected performance variance for small clinical datasets, and our approach of using raw waveforms may offer a more robust alternative to feature extraction techniques that often increase model complexity. Importantly, the use of machine learning models on raw CBFV data, rather than traditional metrics such as MOCAIP, has gained momentum in recent studies. A notable example is the work by Bradley and Kyriacou [14], where raw pulse waveforms from photoplethysmography (PPG) were utilized to detect changes in intracranial pressure with higher accuracy than methods that rely on predefined features. Their results align with our own findings, supporting the notion that machine learning models can capture a broader spectrum of information by analyzing raw, unprocessed physiological signals. Our findings contribute to the broader discussion on generalizability in medical machine learning. As evidenced by studies such as those of Ye et al. [26], whose model demonstrated improved generalization across diverse patient populations, ensuring that models generalize effectively to unseen clinical data is crucial in medical settings. Our approach, which involves using raw waveform data and minimizing overfitting through cross-validation, aligns with the broader goal of creating robust and widely applicable models in clinical practice.

When compared to previously proposed non-invasive ICP detection methods, our framework demonstrates significant performance gains. Kim et al. [8], using a semi-supervised learning framework on a similar cohort, reported an AUC of 94%. By contrast, our refined methodology



(a) MOCAIP Metrics - Uncorrected



(b) MOCAIP Metrics - Manually corrected

Fig. 5. ROC and ROC-PR plots after pre-processing steps applied to the CBFV training data. Figure (a) shows the ROC and precision-recall curves when using uncorrected MOCAIP metrics. Figure (b) shows the pROC and precision-recall curves using manually corrected MOCAIP metrics.

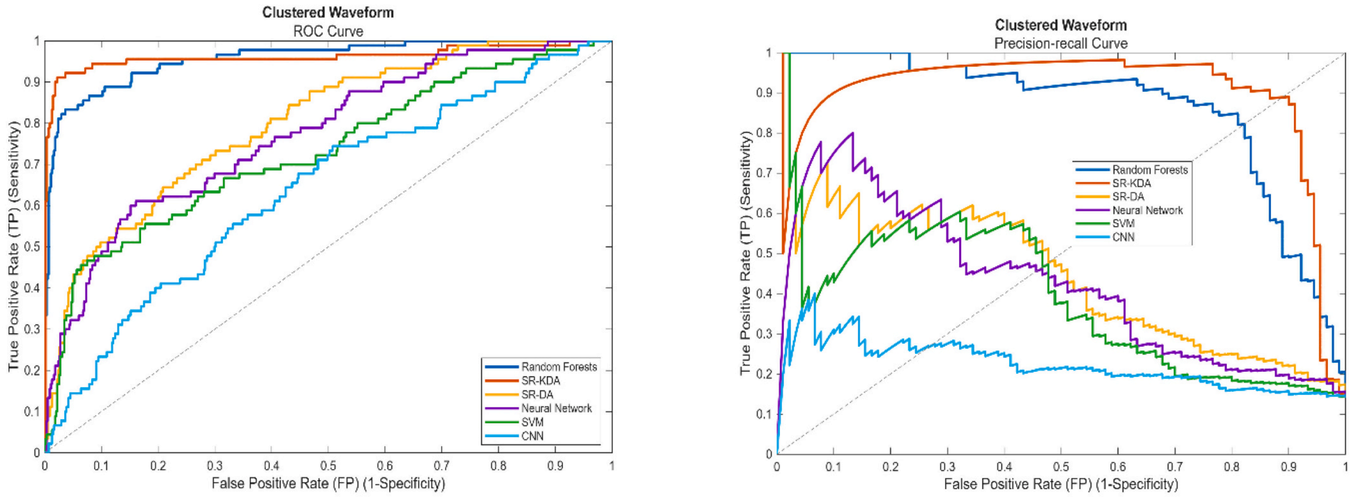
achieved a higher AUC of 96.92%. We attribute this advancement to two critical factors: first, the implementation of a more rigorous three-stage waveform preprocessing pipeline (AUC normalization, standardization, and amplitude scaling), which effectively suppresses non-physiological artifacts; and second, the integration of SR-KDA, which excels at identifying non-linear morphological patterns within the CBFV pulses. This superiority is further highlighted when compared to traditional TCD-based regression models, an AUC of 86% $\geq$ 20 mmHg) reported in [27] and an AUC of 80% reported by Varsos et al. [28]. Our morphologically-driven approach captures the complex pulse-to-pulse dynamics, providing a more reliable clinical surrogate for invasive monitoring.

4.2. *Statistical robustness of our framework*

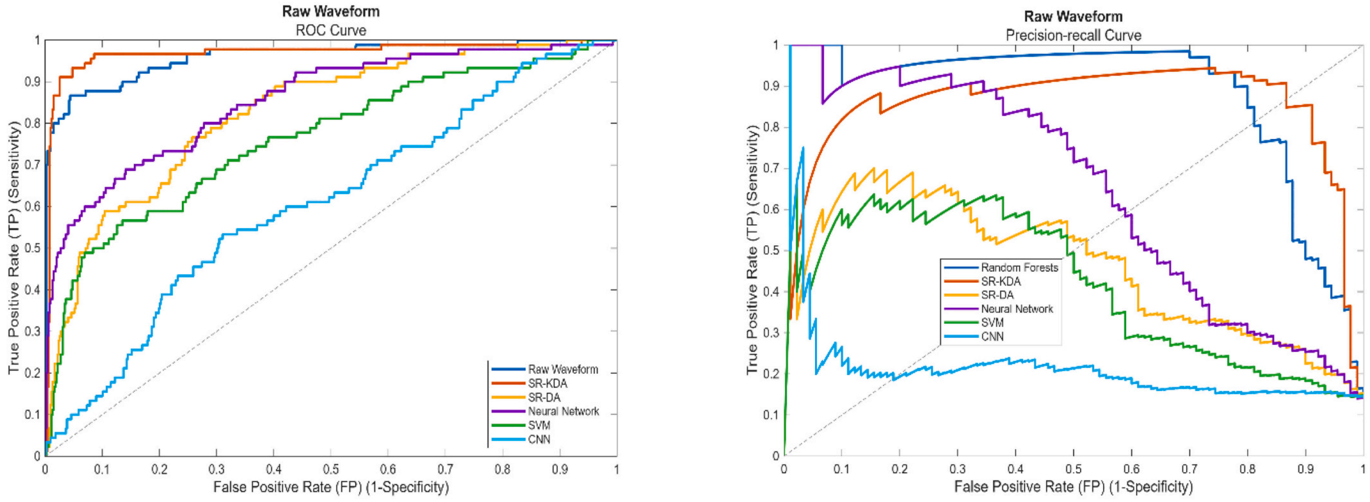
A pivotal finding of this study is that utilizing the full raw CBFV pulse as input yields superior classification accuracy compared to utilizing the 128 MOCAIP morphological landmarks. This suggests that the continuous raw waveform preserves high-fidelity morphological features of cerebral hemodynamics, including non-linear morphological variations that are lost during the discrete feature extraction process of MOCAIP. This mitigates the risk of overfitting to the limited labeled cohort and ensures that the classification logic is based on the stable, recurring patterns of intracranial hemodynamics. Consequently, the model provides a consistent, evidence-based risk assessment that is less sensitive to transient artifacts, thereby satisfying the practical requirements of neurocritical care.

4.3. *Limitations*

Despite the robust classification performance (AUC-ROC: 96.92%) achieved by our proposed framework, certain limitations must be addressed regarding its clinical generalizability. This study is based on a single-center retrospective dataset comprising 89 patients and 638 episodes. While our patient-level balancing and cross-validation strategies were designed to mitigate bias, the high performance observed here



(a) Clustered Waveforms



(b) Raw Waveforms

Fig. 6. Figure (a) shows the ROC and precision-recall curve when using 1-min dominant cluster waveforms. Figure (b) shows the ROC curves when using raw waveforms.

primarily reflects the model's efficacy within this specific cohort. These results may not be directly extrapolatable to broader clinical populations with different age groups, ethnicities, or diverse neuropathologies not represented in our sample. Furthermore, variations in TCD equipment and operator technique across clinical settings could introduce noise not present in this controlled dataset. Therefore, while our findings provide a strong proof-of-concept, large-scale, multi-center, and prospective external validation is essential to confirm the universal applicability and clinical reliability of the proposed methodology before it can be deployed in routine neurocritical care.

Besides, a key methodological consideration in this study is the adoption of a fixed ICP threshold of 20 mmHg for labeling intracranial hypertension. This value was chosen because it represents the standard clinical threshold for intervention according to major neurotrauma guidelines [29,30]. However, the choice of threshold significantly influences both the model's statistical performance and its clinical utility. Mathematically, lowering the threshold would likely increase the model's sensitivity (recall), facilitating earlier detection of rising ICP, but might simultaneously decrease precision by including more borderline cases that do not require immediate intervention. Conversely, a stricter threshold (e.g., 25 mmHg) would yield a more specific model but might fail to provide early warning for evolving pathology. In clinical practice, the transition from normal to high ICP is often a continuum rather than a discrete jump. Future iterations of this framework could benefit from a multi-threshold approach or a continuous estimation model to better capture the 'gray zone' (e.g., 15–25 mmHg), where clinical decision-making is most challenging.

In summary, our study highlights the promise of leveraging raw CBFV waveforms for non-invasive detection of intracranial hemorrhage (IH), offering improved accuracy and generalizability compared to traditional methods. Future research should focus on addressing the challenges of feature interpretability and computational efficiency to facilitate broader clinical adoption.

5. Conclusion

This study presents a non-invasive framework for detecting intracranial hypertension (IH) that utilizes raw cerebral blood flow velocity (CBFV) waveforms as input for machine learning algorithms. Our approach addresses the limitations of traditional invasive monitoring techniques and overcomes the drawbacks of pre-established morphological metrics, such as MOCAIP. By directly analyzing raw CBFV waveforms, the proposed method demonstrates superior accuracy and generalization, as evidenced by the improved area under the receiver operating characteristic curve (AUC-ROC) and precision-recall metrics.

The findings highlight the potential of raw waveform analysis to preserve clinical information and extract hidden features that are otherwise lost in traditional pre-processing methods. This capability not only improves prediction accuracy but also ensures robust model performance across varied datasets, making it particularly suitable for clinical applications.

In conclusion, this study demonstrates the feasibility and effectiveness of using raw CBFV waveforms for non-invasive detection of intracranial hemorrhage (IH), marking a significant step toward safer and more accessible monitoring solutions. Future work should aim to expand the dataset size, improve model interpretability, and explore real-time applications to maximize clinical utility.

CRediT authorship contribution statement

Miaomiao Wei: Writing – review & editing, Writing – original draft, Supervision, Software, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Solventa Krakauskaite:** Writing – review & editing, Writing – original draft, Software, Investigation, Funding acquisition, Data curation, Conceptualization. **Ryan Mercer:** Writing – original draft, Validation, Data curation. **Jiaguo Lin:** Writing – original draft, Visualization, Validation, Software, Formal analysis, Data curation. **Laimonas Bartusis:** Methodology, Investigation, Formal analysis. **Fabien Scalzo:** Writing – review & editing, Supervision, Resources, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

All authors confirm that all procedures were performed in compliance with relevant laws and institutional guidelines. The authors confirm that the privacy rights of human subjects were always observed.

This research was supported in part by the National Natural Science Foundation of China with Grant number 62301624, the Scientific Research Projects of Higher Education Institutions of Henan Province of China with Grant number 23A510012, and the Research Council of Lithuania under Grant Agreement No. SV5-40.

Data availability

The data that has been used is confidential.

References

[1] G. Frigieri, S. Brasil, D. Cardim, et al., Machine learning approach for noninvasive intracranial pressure estimation using pulsatile cranial expansion waveforms, npj Digit. Med. 8 (2025) 57, https://doi.org/10.1038/s41746-025-01463-y.

[2] M. Megjhani, B. Weinerman, T. Alalqum, et al., Application of a time series foundation model to noninvasively estimate intracranial pressure, J. Clin. Monit. Comput. 39 (2025) 1283–1292, https://doi.org/10.1007/s10877-025-01366-z.

[3] G. Hunter, C. Voll, M. Rajput, Utility of transcranial doppler in idiopathic intracranial hypertension, Can. J. Neurol. Sci. 37 (2010) 235–239.

[4] J. Bellner, B. Romner, P. Reinstrup, K.A. Kristiansson, E. Ryding, L. Brandt, Transcranial doppler sonography pulsatility index (pi) reflects intracranial pressure (icp), Surg. Neurol. 62 (2004) 45–51.

[5] A. Ziółkowski, A. Pudełko, A. Kazimierska, Z. Czosnyka, M. Czosnyka, M. Kasprowicz, Analysis of relative changes in pulse shapes of intracranial pressure and cerebral blood flow velocity, Physiol. Meas. 42 (2021) 125004, https://doi.org/10.1088/1361-6579/ac38bf.

[6] A. Ziółkowski, A. Pudełko, A. Kazimierska, A. Uryga, Z. Czosnyka, M. Kasprowicz, M. Czosnyka, Peak appearance time in pulse waveforms of intracranial pressure and cerebral blood flow velocity, Front. Physiol. 13 (2023), https://doi.org/10.3389/fphys.2022.1077966.

[7] A. Behrens, N. Lenfeldt, K. Ambarki, J. Malm, A. Eklund, L.O. Koskinen, Transcranial Doppler pulsatility index: not an accurate method to assess intracranial pressure, Neurosurgery 66 (6) (2010) 1050–1057, https://doi.org/10.1227/01.NEU.0000369519.35932.F2.

[8] S. Kim, R. Hamilton, S. Pineles, M. Bergsneider, X. Hu, Noninvasive intracranial hypertension detection utilizing semisupervised learning, IEEE Trans. Biomed. Eng. 60 (4) (2013) 1126–1133, https://doi.org/10.1109/TBME.2012.2227477.

[9] N. Canac, K. Jalaleddini, S.G. Thorpe, C.M. Thibeault, R.B. Hamilton, Review: pathophysiology of intracranial hypertension and noninvasive intracranial pressure monitoring, Fluids Barriers CNS 17 (1) (2020) 40. Published 2020 Jun 23, https://doi.org/10.1186/s12987-020-00201-8.

[10] M. Imaduddin S, A. Fanelli, F.W. Vonberg, R.C. Tasker, T. Heldt, Pseudo-Bayesian model-based noninvasive intracranial pressure estimation and tracking, IEEE Trans. Biomed. Eng. 67 (6) (2020) 1604–1615, https://doi.org/10.1109/TBME.2019.2940929.

[11] S. Kim, X. Hu, D. McArthur, R. Hamilton, M. Bergsneider, T. Glenn, N. Martin, P. Vespa, Inter-subject correlation exists between morphological metrics of cerebral blood flow velocity and intracranial pressure pulses, Neurocrit. Care. 14 (2) (2011) 229–237, https://doi.org/10.1007/s12028-010-9471-x.

[12] M. Megjhani, K. Terilli, B. Weinerman, D. Nametz, S.B. Kwon, A. Velazquez, S. Ghoshal, D.J. Roh, S. Agarwal, E.S. Connolly Jr., J. Claassen, S. Park, A deep learning framework for deriving noninvasive intracranial pressure waveforms from transcranial Doppler, Ann. Neurol. 94 (1) (2023) 196–202, https://doi.org/10.1002/ana.26682.

[13] S.S. Nair, A. Guo, J. Boen, A. Aggarwal, O. Chahal, A. Tandon, M. Patel, S. Sankararaman, N.J. Durr, T.D. Azad, R. Pirracchio, R.D. Stevens, A deep learning approach for generating intracranial pressure waveforms from extracranial signals routinely measured in the intensive care unit, Comput. Biol. Med. 177 (2024) 108677, https://doi.org/10.1016/j.compbiomed.2024.108677.

[14] G.R.E. Bradley, P.A. Kyriacou, Evaluating the effectiveness of non-invasive intracranial pressure monitoring via near-infrared photoplethysmography using classical machine learning methods, Biomed. Signal Process. Control. 96 (Part B) (2024) 106517, https://doi.org/10.1016/j.bspc.2024.106517.

[15] O. Chapelle, A. Zie, B. Schölkopf, Semi-Supervised Learning, MIT Press, 2006.

[16] X. Zhu, Semi-Supervised Learning Literature Survery, 2005.

[17] A. Kapoor, Y. Qi, H. Ahn, R. Picard, Hyperparameter and kernel learning for graph based semi- supervised classification, in: Advances in Neural Information Processing Systems, 2005.

[18] T. Zhang, R. Audo, Analysis of spectral kernel design based semi-supervised learning, in: Proceedings of the 19th International Conference on Neural Information Processing Systems (NIPS'05), MIT Press, Cambridge, MA, USA, 2005, pp. 1601–1608.

[19] D. Cai, X. He, J. Han, Semi-Supervised Regression Using Spectral Techniques, Department of Computer Science, University of Illinois at Urbana-Champaign, 2006. Tech. Rep. July. UIUCDCSR-2006-2749.

[20] X. Hu, P. Xu, S. Asgari, P. Vespa, M. Bergsneider, Forecasting ICP elevation based on prescient changes of intracranial pressure waveform morphology, IEEE Trans. Biomed. Eng. 57 (5) (2010) 1070–1078, https://doi.org/10.1109/TBME.2009.2037607.

[21] F. Scalzo, R. Hamilton, S. Asgari, S. Kim, X. Hu, Intracranial hypertension prediction using extremely randomized decision trees, Med. Eng. Phys. 34 (8) (2012) 1058–1065, https://doi.org/10.1016/j.medengphy.2011.11.010.

[22] S. Asgari, M. Bergsneider, R. Hamilton, P. Vespa, X. Hu, Consistent changes in intracranial pressure waveform morphology induced by acute hypercapnic cerebral vasodilatation, Neurocrit. Care. 15 (1) (2011) 55–62, https://doi.org/10.1007/s12028-010-9463-x.

[23] L. Kaufman, P. Rousseeuw, Findig Groups in Data: An Introduction to Cluster Analysis., Wiley, New York, 1990.

[24] L. Breiman, Random Forests, Mach. Learn. 45 (1) (2001) 5–32, https://doi.org/10.1023/A:1010933404324.

[25] D. Cai, X. He, J. Han, Speed up kernel discriminant analysis, VLDB J. 20 (2011) 21–33, https://doi.org/10.1007/s00778-010-0189-3.

[26] G. Ye, V. Balasubramanian, J.K. Li, M. Kaya, Machine learning-based continuous intracranial pressure prediction for traumatic injury patients, IEEE J. Transl. Eng. Health Med. 10 (2022) 4901008. Published 2022 Jun 2, https://doi.org/10.1109/JTEHM.2022.3179874.

[27] C. Robba, J. Donnelly, R. Bertuetti, D. Cardim, M.S. Sekhon, M. Aries, P. Smielewski, H. Richards, M. Czosnyka, Doppler non-invasive monitoring of ICP in an animal model of acute intracranial hypertension, Neurocrit. Care. 23 (3) (2015) 419–426, https://doi.org/10.1007/s12028-015-0163-4.

[28] G.V. Varsos, A.G. Kolias, P. Smielewski, K.M. Brady, V.G. Varsos, P.J. Hutchinson, J.D. Pickard, M. Czosnyka, A noninvasive estimation of cerebral perfusion pressure using critical closing pressure, J. Neurosurg. 123 (3) (2015) 638–648, https://doi.org/10.3171/2014.10.JNS14613.

[29] N. Carney, A.M. Totten, C. O'Reilly, J.S. Ullman, G.W. Hawryluk, M.J. Bell, S. L. Bratton, R. Chesnut, O.A. Harris, N. Kissoon, A.M. Rubiano, L. Shutter, R. C. Tasker, M.S. Vavilala, J. Wilberger, D.W. Wright, J. Ghajar, Guidelines for the management of severe traumatic brain injury, Neurosurgery 80 (1) (2017) 6–15, https://doi.org/10.1227/NEU.0000000000001432.

[30] P.M. Kochanek, R.C. Tasker, N. Carney, A.M. Totten, P.D. Adelson, N.R. Selden, C. Davis-O'Reilly, E.L. Hart, M.J. Bell, S.L. Bratton, G.A. Grant, N. Kissoon, K. E. Reuter-Rice, M.S. Vavilala, M.S. Wainwright, Guidelines for the management of pediatric severe traumatic brain injury, third edition: update of the Brain Trauma Foundation guidelines, Pediatr. Crit. Care Med. 20 (3S Suppl 1) (2019) S1–S82, https://doi.org/10.1097/PCC.0000000000001735.