A Real-Time Deep Learning Approach for Inferring Intracranial Pressure from Routinely Measured Extracranial Waveforms in the Intensive Care Unit

Shiker S. Nair,[1] Alina Guo,[1] Joseph Boen,[1] Ataes Aggarwal,[1] Ojas Chahal,[1]
Arushi Tandon,[1] Meer Patel,[1] Sreenidhi Sankararaman,[1] and Robert D. Stevens[2]

[1] *Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA*
[2] *Departments of Anesthesiology and Critical Care Medicine, Neurology, Neurosurgery, and Radiology, Johns Hopkins University School of Medicine, Baltimore, MD, USA*

Abstract - Objective. Intracranial pressure (ICP) is a vital waveform used to assess the conditions of patients with intracranial pathologies, such as traumatic brain injury. The current standard for monitoring ICP requires a catheter to be inserted into the brain, which is extremely invasive and increases the risk of hemorrhage and infection. We hypothesize that ICP can be inferred from extracranial waveforms routinely measured in the Intensive Care Unit (ICU), such as invasive arterial blood pressure (ABP), photoplethysmography (PPG), and electrocardiography (ECG). **Methods.** We extracted 600 hours of simultaneous ABP, EKG, PPG, and ICP waveforms (125 Hz) across 10 different patients from the MIMIC III Waveform Database. These concurrent recordings were broken into 10 second windows to train on long-short term memory (LSTM) and temporal convolutional network (TCN) models to predict ICP using ABP, EKG, and PPG as input features. These models were evaluated in both a single-patient analysis and multi-patient analysis study. **Results.** In the single-patient analysis study, TCN's median mean average error (MAE) test loss was 1.60 mmHg in comparison to a median MAE test loss of 1.88 mmHg for LSTM. TCN was used in the multi-patient analysis study, reporting a test MAE loss of 5.44 mmHg. **Conclusions.** These novel results indicate that it is possible to predict ICP waveforms from extracranial waveforms in a real-time, continuous, and non-invasive manner. This approach could potentially eliminate the risks associated with invasive monitoring and allow for more timely treatment of patients with intracranial pathologies with further follow-up studies.

I. INTRODUCTION

Traumatic brain injury (TBI), a type of brain damage caused by physical trauma to the head, is one of the leading causes of death and disability worldwide [1]. There were approximately 223,135 TBI-related hospitalizations in 2019 and 69,473 TBI-related deaths in 2021 [2]. The main cause of TBI-related death is increased intracranial pressure (ICP) [3]. ICP is defined as the pressure exerted onto the craniospinal compartment of a patient [4]. At rest, a normal supine adult's ICP can range between 7-15 millimeters of mercury (mmHg) [5,6]. Intracranial hypertension is defined as the rise of ICP above 22 mmHg [4]. This is often caused by a rise in the pressure of the cerebral spinal fluid (CSF) or an increase in pressure due to tumors or hemorrhaging within the brain [7].

Monitoring ICP is imperative in patients with TBI. In cases of elevated ICP, immediate treatment is required for the patient (i.e. drainage of cerebrospinal fluid) [4]. Currently, the intraventricular catheter (IVC) is the gold standard for monitoring ICP because of its accuracy ($\pm$ 2 mmHg) [5,8]. Insertion of an IVC requires a hole to be drilled at the base of the skull and the probe to be inserted into the ventricles of the brain. The invasive nature of this procedure poses a risk to the patient's health. The risk of infection for the IVC procedure is 5-14% and the risk of hemorrhage during insertion of the catheter is 5-7% [3,6]. Additionally, insertion of an IVC takes approximately 30-60 minutes, resulting in delayed ICP measurements and furthermore late diagnosis and treatment action [9]. This could result in rapid deterioration of the health of a TBI patient and, eventually, lead to death [4].

There are alternative, less invasive devices on the market for measuring ICP (i.e. subarachnoid bolt, epidural bolt, parenchymal monitor, or transcranial Doppler) [8]. However, they are either not as accurate as the IVC or do not offer real-time, continuous monitoring of ICP [8]. Thus, there remains a need for a less invasive, continuous way of monitoring ICP that is reliable and accurate.

Previous research suggested that ICP is regulated by the vascular system, and signatures in electrical, blood pressure, and blood oxygenation waveforms have been associated with intracranial events [10]-[12]. Cerebral autoregulation may be impaired when ICP is elevated, leading to changes in blood flow, vasoreactivity, and oxygenation. The cerebral vascular response as a whole depends on ICP. Therefore, we hypothesize that ICP can be inferred from extracranial waveform monitoring, such as invasive arterial blood pressure monitoring (ABP), photoplethysmography (PPG), and electrocardiography (ECG).

Monitoring ICP using extracranial waveforms, which are routinely measured in the ICU, would empower clinicians to assess the health of a patient in a less-invasive, cost-effective manner. Such a solution would be easily integrable into the current clinical workflow for monitoring ICP and would alleviate the concerns for infection, hemorrhage, and time delays associated with the invasive nature of the IVC.

II. RELATED WORK

Although previous groups have attempted to develop both physics-based simulations and statistical models, there have been limited efforts to use data-driven models to accomplish this task [13, 14]. This paper looks at using traditional machine learning and deep learning architectures to successfully predict ICP using high-frequency (125 Hz) ABP, PPG, and ECG waveform data in real-time.

Our research presents a novel approach to predicting ICP in patients with TBI that significantly differs from existing studies in the literature. The primary novelty of our approach lies in the usage of only extracranial signals to predict the ICP waveform in quasi-real-time, using deep learning techniques. Similar studies in this field have successfully developed methods to predict the occurrence of hypertensive events or the progression of intracranial hypertension (IH) into a potentially life-threatening condition (LTH) [15]-[18]. Other works have developed deep-learning algorithms to forecast the ICP waveform [19,20]. However, all of these studies rely on using ICP as an input and therefore do not offer the potential of an extracranial, non-invasive methodology for measuring ICP. There is indeed literature for the real-time prediction of ICP using extracranial signals [13,14]; however, these studies do not utilize deep learning in their formula-based methods. A potential weakness with such methods may be a lack of causality, as transfer functions may only capture a statistical relationship between input and output signals, which may inhibit patient specificity. By focusing on non-invasive methods for measuring ICP, our research has the potential for transforming the monitoring and treatment of TBI patients, particularly in resource-limited settings. ICP can become a routinely measured vital sign beyond the intensive care unit (ICU). In addition, real-time waveform recreation with deep learning may offer a novel approach to offering clinicians an unbiased recreation of the waveform for their own assessment as well.

III. DATA AND DEMOGRAPHICS

The publicly accessible MIMIC-III database (v1.4, 2016) is an intensive care database that comprises de-identified data from 46,520 patients and 58,976 admissions, recorded between June 1, 2001, and October 31, 2012, at the Beth Israel Deaconess Medical Center in Boston, USA [21]. The data includes demographics, admission, and discharge notes, International Classification of Diseases - 9th revision (ICD-9) codes for diagnoses, radiological findings, survival data, vital sign recordings, medications, and procedures for each patient [21]. MIMIC-III was selected as the data source because it contains concurrent recordings of high-frequency physiological waveforms at 125 Hz that are relevant to the ICP prediction task. From the available waveforms, we chose model ICP from ABP, ECG, and PPG. By choosing these three waveforms, we had three modalities of predicting ICP and assessing the contributions of each waveform in the modeling tasks. The demographics of the 10 patients evaluated in this study can be seen in Table 1.

Table I. Patient Demographic Data in Study

Total Number of Subjects	n = 10
Age Median [IQR]	56 [33-64]
Sex(% M, % F)	30, 70
Admission Type (Number of Subjects)	Emergency (8) Urgent (2)
Discharge Location (Number of Subjects)	Rehab/Distinct Part Hosp (5) Dead/Expired (3) Home Health Care (1) Home (1)
Diagnosis (Number of Subjects)	Subarachnoid hemorrhage (5) Acute liver failure (2) Traumatic brain injury (1) Intracerebral hemorrhage (1) Closed fracture of base of skull (1)

IV. METHODOLOGY

A. Preprocessing and Cohort Selection

Patients who were at least 18 years old with ECG, PPG, ABP, and ICP recordings for at least five minutes were included in this study. In order to accomplish this, we first assessed whether a patient had recordings of all four waveforms. Because ECG is measured at multiple different leads, we used the most prevalent ECG lead (II) in our patient population of interest. The rationale behind the common usage of ECG lead II for ICP prediction is rooted in its ability to provide a reliable depiction of the mechanical forces engendered by the cardiac system and their subsequent impact on ICP [22]. After identifying patients who have ICP, ABP, ECG Lead II, and PPG data available, we removed patients who had less than five minutes of overlapping data between these four waveforms. This was done based on the knowledge that elevated ICP lasting for five minutes or more may indicate a hypertensive event, according to clinical experts [23]. Therefore, it was important to ensure that we had sufficient data to accurately detect such an event if it occurred.

From the patients that remained, the waveform data (synchronized in time) were sliced into five-minute seg-

ments. To ensure rich data quality, five-minute segments that contained any not a number (NaN) values or physiologically impossible values (ICP < -10 mmHg or > 200 mmHg, ABP < 20 mmHg or > 300 mmHg, not NaN mean heart rate) unique to each waveform were filtered out. To address potential issues with IVC calibration, an additional periodicity filtering step was applied to the ICP signals. This filtering step removed extremely high-frequency and low-frequency waveforms that could lead to inaccuracies in the data. In order to ensure that hypertensive events were adequately represented in the data, patients who had less than one hour of hypertensive recordings or less than 15 hours of total data were excluded from the analysis. For this study, an ICP hypertensive event was defined as an ICP greater than 20 mmHg for at least five minutes [23]. Although some recent studies use a threshold of 22 mmHg for traumatic brain injuries, our study includes a heterogeneous population of patients who underwent ICP monitoring for various clinical purposes [4]. Therefore, we opted to use a threshold of 20 mmHg (Figure 1).

B. Determining Window Length for ICP Prediction Task

To ensure that the ICP prediction task was as close to real-time as possible, a window size of 10 seconds was chosen. To assess the amount of information lost as the window size decreased, a dynamic time warping analysis (DTW) was conducted [24]. The patient data was divided into windows of 10 seconds, 20 seconds, 30 seconds, and 60 seconds. The DTW score for each of the extracranial waveforms (ABP, ECG, and PPG) and ICP was computed between adjacent windows. This resulted in four scores (one for each waveform) for each pair of adjacent windows. The Spearman correlation between the DTW scores for each waveform and ICP was calculated (ABP/ICP, ECG/ICP, and PPG/ICP) in order to assess if changes in the extracranial waveforms would be reflected in ICP.

The results showed that all window sizes had a weak positive correlation between the extracranial waveforms and ICP (Table 2). The Spearman coefficient increased with increasing window length, but the correlation was similar across all window sizes. This justified the choice of a 10-second window size to replicate real-time analysis.

C. Model Selection and Architecture

The goal of this study is to predict ICP from ABP, ECG, and PPG waveforms for a 10 second window. Given this is a multivariate time series prediction problem, we implemented six different deep learning models which are fit to accomplish this task and have been used

Table II. Dynamic Time Warping Analysis for Assessing Information Loss Across Decreasing Window Sizes

Window Size (s)	ICP/ABP	ICP/PPG	ICP/ECG
10	0.39	0.29	0.25
20	0.37	0.30	0.29
30	0.40	0.33	0.31
60	0.42	0.37	0.35

in previous studies in for similar problem spaces (Figure 2):

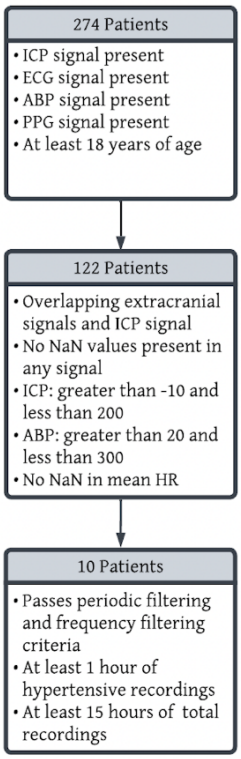
- Recurrent Neural Network (RNN) [25,26]
- Gated Recurrent Unit (GRU) [27]-[30]
- Long-Short Term Memory (LSTM) network [19], [30- 32]
- Temporal Convolutional Networks (TCN) [33]-[36]
- Volume-Net (VNET) [37]
- Transformer [38,39]

These six models vary in sophistication and complexity. We currently have validated and tested on only TCN and LSTM. In general, adding more layers improved the predictions for each model, however, this increased the risk of overfitting. Rigorous hyperparameter tuning was applied for each model, specifically finding an optimal learning rate, batch size, and number of nodes for each layer. The final algorithms tested were a 6-layer TCN model and 7 layer LSTM model.

D. Normalization and Evaluation Criteria

Both features (ABP, ECG, and PPG) and labels (ICP) waveforms were independently normalized for the prediction tasks. Normalization was completed using a standard scaling function (subtracting by mean and dividing by the standard deviation of the waveform) which was applied to all the data before splitting it into a training and testing set. While it is more common practice to normalize the training set and apply this normalization scheme to the test set, the small number of patients used in this analysis (n=10) did not allow for a generalizable normalization function in the training set that would properly transfer over to the testing set. As a result, all the data was normalized before splitting into a training and testing set. Predictions were then unscaled back into a clinically interpretable ICP using the inverse scaling function.

Different loss functions were assessed when optimizing each model and it was determined that the log hyperbolic cosine (log cosh) loss was the best optimization function. It is defined as the logarithm of the hyperbolic cosine




```
Result: Final list of patients that meet all inclusion/exclusion criteria
patients ← dataset of 274 patients with ICP, ECG, PPG, and ABP
  present and at least 18 years of age;
filtered_patients ← empty list;
for patient in patients do
    icp_signal ← get_icp_signal(patient);
    extracranial_signals ← get_extracranial_signals(patient);
    overlapping_signals ← get_overlapping_signals(icp_signal,
      extracranial_signals);
    if len(overlapping_signals) < 5 minutes then
        continue;
    end
    if any(signal_value is NaN for signal_value in overlapping_signals)
      then
        continue;
    end
    if any(signal_value < -10 or signal_value > 200 for signal_value in
      icp_signal) then
        continue;
    end
    abp_signal ← get_abp_signal(patient);
    if any(signal_value < 20 or signal_value > 300 for signal_value in
      abp_signal) then
        continue;
    end
    hr_signal ← get_hr_signal(patient);
    if any(signal_value is NaN for signal_value in hr_signal) then
        continue;
    end
    filtered_patients.append(patient);
end
hypertensive_patients ← empty list;
for patient in filtered_patients do
    icp_signal ← get_icp_signal(patient);
    if mean(icp_signal) ≤ 20 then
        continue;
    end
    if total_recording_hours(patient) < 15 then
        continue;
    end
    if hypertensive_recording_hours(patient) < 1 then
        continue;
    end
    hypertensive_patients.append(patient);
end
periodic_and_frequency_filtered_patients ← empty list;
for patient in hypertensive_patients do
    if at_least_one_hour_hypertensive_recording(patient) and
      at_least_15_hours_total_recording(patient) and
      frequency_filtering(patient) then
        periodic_and_frequency_filtered_patients.append(patient);
    end
end
final_patients ← empty list;
for patient in periodic_and_frequency_filtered_patients do
    final_patients.append(patient);
end
return final_patients;
```

Algorithm 1: Inclusion/exclusion criteria for patient selection

FIG. 1: Cohort Selection Workflow (Left) and Pre-Processing Methodology (Right)

of the difference between the true and predicted values (Equation 1). The logcosh loss function is a smooth approximation of the MAE loss function, allowing it to be easily optimized through gradient descent methods. Furthermore, logcosh is known to converge faster than other loss functions such as MAE, overall making it an ideal function for deep learning tasks [40]-[42].

$$(1) \quad Logcosh(t) = \sum_{p \in P} log(cosh(p - t))$$

E. Single Patient Analysis

The first experiment conducted looked at the performance of each of the six models within a single patient for each of the 10 patients in the study. The goal of this experiment was to ascertain that each model was capable of predicting ICP waveforms from the extracranial waveforms in the simplest context: within a single patient. This removes any notion of generalizability and enables the comparison of each model in a simpler context before considering inter-patient generalization. Data were split into 70% train/validation set and 30% test set. The mean and median logcosh for training and testing datasets were reported for each model along with the standard deviation. Using this model information, we selected the top model(s) to run a multi-patient analysis. Additional follow-up analysis was performed to understand the strengths and weaknesses of each model and trends in error and prediction.

F. Multi-Patient Analysis

To ensure that the top models are generalizable to new patients, a multi-patient analysis was conducted using 6 patients for training, 2 for validation, and 2 for the test set. This approach ascertains that the models are evaluated on data from patients that are distinct from those used in the training set, thus providing a robust assessment of model performance. The mean and median logcosh values for the training and testing datasets were reported for each model, to provide a measure of model performance accuracy and consistency. Furthermore, additional follow-up analyses were performed to understand trends across the top models and across different patients in the validation and testing sets. This analysis is crucial for ensuring that the models are generalizable and can be applied to new patients in clinical practice, where the algorithm will be trained on different patients than the patients it will be used on in the ICU.

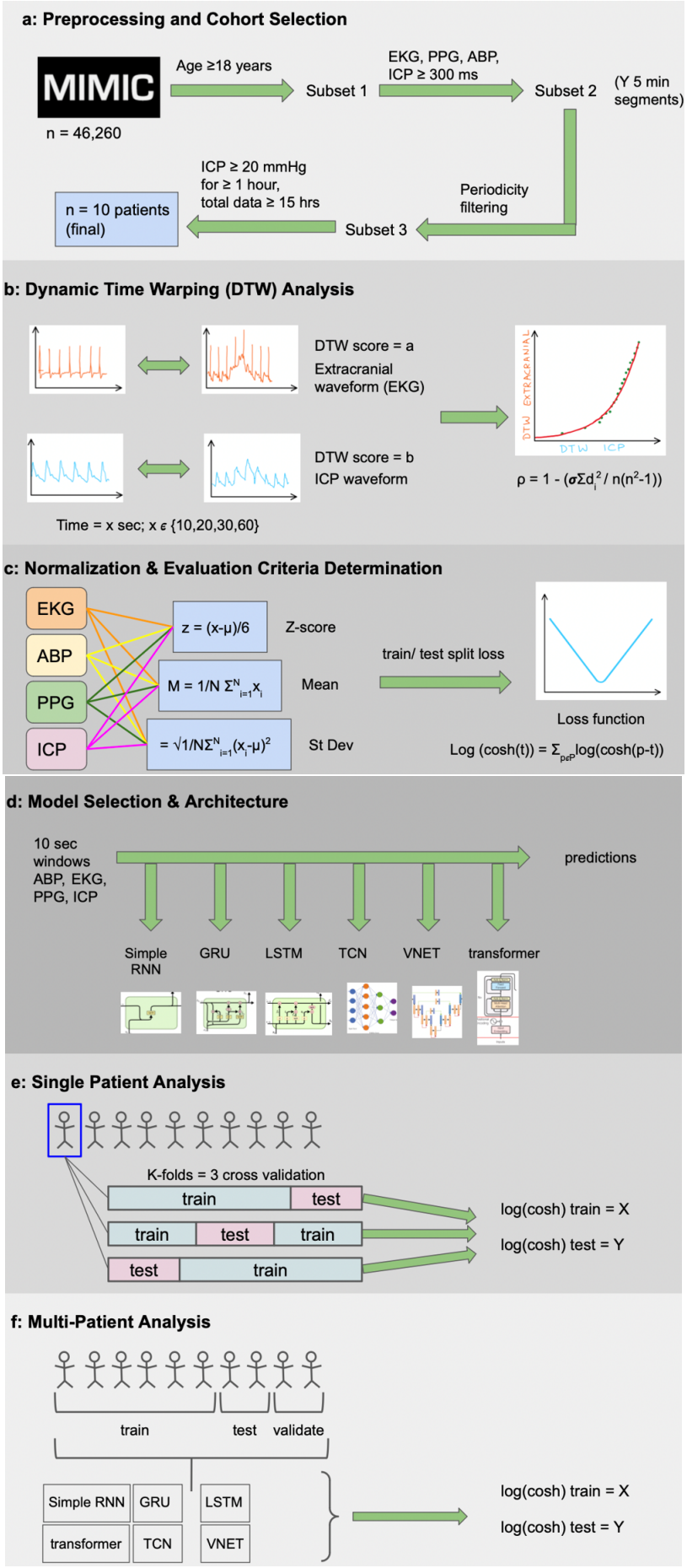


FIG. 2: Overview of Methodology and Analysis Performed in Study

V. RESULTS

Based on the single patient analysis for predicting ICP using EKG, ABP, and PPG in 10-second windows, TCN performed better than LSTM (Table 2 and Table 3). The mean testing logcosh for TCN was 1.13 ± 0.70 mmHg (median 0.95 mmHg) compared to 1.84 ± 1.78 mmHg (median 1.26 mmHg) for LSTM. The mean testing MAE for TCN was 1.77 ± 0.61 mmHg (median 1.60 mmHg) compared to 2.50 ± 1.91 (median 1.88 mmHg). The mean testing RMSE for TCN was 3.44 ± 0.56 mmHg (median 3.29 mmHg) compared to 4.11 ± 1.74 mmHg (median 3.54 mmHg) for LSTM. The LSTM algorithm had particular trouble finding the proper global optimization for Pat 3, which had a significantly higher training and testing loss for LSTM compared to TCN. TCN's improved performance in comparison to the LSTM has also been noted in a previous study assessing waveform data for anomaly detection [43].

Consequently, it was determined that TCN was the best model in this study and was therefore used for the multi-patient analysis. Figure 3 shows two examples of strong predictions by TCN. When training on 6 patients, validating on 2, and testing on 2, we found that the logcosh testing loss was 4.79 mmHg, MAE testing loss was 5.44 mmHg, and RMSE testing loss was 6.78 mmHg (Table 4). The higher loss in the multi-patient analysis is due to poor patient generalizability from the training to the test set. As we continue to increase the number of patients in this study, we expect to create a more representative sample of patients with ICP recordings to improve patient-to-patient model generalizability.

VI. DISCUSSION

A. Key Findings

We have developed a continuous, non-invasive, and real-time methodology for deriving ICP from extracranial waveforms. Our results show that deep learning algorithms are capable of modeling ICP from extracranial waveforms, as demonstrated in our single patient analysis study where LSTM and TCN had a MAE of 1.77 ± 0.61 mmHg and 2.50 ± 1.91 mmHg respectively. Furthermore, our results show that our TCN algorithm is proficient in generalizing across patients in a small training set (n = 6 patients) to predict ICP with an MAE of 5.44 mmHg. The strength of these results can be partly explained by the relationship between these extracranial waveforms and intracranial hemodynamics. In addition these results show there exists valuable information encoded in the waveform morphologies of EKG, PPG, and ABP that relate to ICP.

Given the novelty of our work, there exist limited studies that attempt to predict ICP from extracranial waveforms on a continuous and real-time basis. However, we can compare our algorithms' performances to current marketed technologies that aim to measure ICP both less-invasively and invasively. In particular, our TCN algorithm's multi-patient analysis performance is significantly better than current, less-invasive techniques for monitoring ICP. Transcranial Doppler ultrasonography has an error range of ± 12 mmHg [44] and cannot monitor ICP continuously. Tympanic membrane displacement has an error range of ± 15 mmHg and requires a probe to be inserted into the canal of the ear [8]. Optic sheath diameter, a retinal imaging based technique, reports a margin of error from 5-10 mmHg and also cannot monitor ICP continuously [8]. Furthermore, our TCN algorithm performed similarly to other state-of-the-art alternatives to the IVC such as intracranial transducers and intraparenchymal monitoring which both report a mean error of ± 7 mmHg [8,45]. While both techniques offer real-time and continuous monitoring, they are on the same level of invasiveness as an IVC, requiring surgical implantation.

B. Clinical Error Analysis Procedure

To improve the accuracy of predictions for individual patients, an error analysis was conducted to identify the factors contributing to the varying degrees of prediction error across the testing dataset. The analysis investigated potential correlations between prediction error and various factors, including autoregulation index, mean ICP values, ICP signal variance, and clustering of features in a reduced dimensionality space. This investigation aimed to identify any underlying reasons for the variation in prediction error across the testing dataset and to inform the development of improved models for predicting ICP in individual patients and across patients.

To investigate the relationship between the autoregulation index and prediction error, we computed the Pearson correlation coefficient between the ABP and ICP to obtain an autoregulation index for each prediction in the testing set. The autoregulation index represents the degree of correspondence between ABP and ICP waveforms. We hypothesized that a higher autoregulation index would be associated with lower prediction error. However, the Pearson correlation between autoregulation index and prediction error revealed no statistically significant relationship between the two variables.

Next, data imbalance was assessed to examine whether poorly represented data resulted in worse predictions. This was particularly important in cases of high mean ICP, which was a small fraction of all data yet the most clinically relevant. To assess whether a high ICP was associated with poor prediction, the mean ICP value was calculated for each window of the testing dataset, and the Pearson correlation coefficient was computed between the ICP value and prediction loss. No statistically significant correlation was observed between prediction error and high ICP.

Additionally, to investigate the correlation between

Table III. Single Patient Analysis Results for LSTM

	Pat 1	Pat 2	Pat 3	Pat 4	Pat 5	Pat 6	Pat 7	Pat 8	Pat 9	Pat 10	Mean $\pm Std$	Median
Logcosh Test Loss (Train Loss) in mmHg	2.79 (2.09)	1.42 (0.34)	6.52 (4.77)	1.14 (1.18)	1.84 (1.40)	0.77 (0.60)	0.71 (0.59)	0.55 (0.21)	0.80 (0.38)	1.83 (1.43)	1.84 $\pm$1.78 (1.30 $\pm$ 1.36)	1.28 (0.89)
MAE Test Loss (Train Loss) in mmHg	3.27 (2.61)	2.00 (1.11)	7.66 (5.41)	1.75 (1.79)	2.37 (1.98)	1.45 (1.31)	1.40 (1.31)	1.27 (1.00)	1.48 (1.14)	2.36 (2.00)	2.50 $\pm$1.91 (1.97 $\pm$ 1.31)	1.88 (1.55)
RMSE Test Loss (Train Loss) in mmHg	4.81 (3.94)	3.65 (2.84)	8.80 (6.76)	3.43 (3.46)	3.99 (3.61)	3.15 (3.03)	3.11 (3.02)	2.99 (2.75)	3.18 (2.87)	3.98 (3.66)	4.11 $\pm$1.74 (3.59 $\pm$ 1.18)	3.54 (3.25)

Table IV. Single Patient Analysis Results for TCN

	Pat 1	Pat 2	Pat 3	Pat 4	Pat 5	Pat 6	Pat 7	Pat 8	Pat 9	Pat 10	Mean $\pm Std$	Median
Logcosh Test Loss (Train Loss) in mmHg	2.63 (1.31)	1.40 (0.48)	0.99 (0.38)	1.47 (0.60)	0.90 (0.52)	0.68 (0.43)	0.47 (0.35)	0.46 (0.16)	0.52 (0.33)	1.81 (0.68)	1.13 $\pm$0.70 (0.52 $\pm$ 0.31)	0.95 (0.46)
MAE Test Loss (Train Loss) in mmHg	3.12 (1.90)	1.97 (1.21)	1.63 (1.14)	2.04 (1.31)	1.56 (1.25)	1.38 (1.18)	1.21 (1.12)	1.20 (0.97)	1.25 (1.10)	2.34 (1.37)	1.77 $\pm$0.61 (1.26 $\pm$ 0.25)	1.60 (1.20)
RMSE Test Loss (Train Loss) in mmHg	4.67 (3.56)	3.63 (2.94)	3.32 (2.87)	3.69 (3.03)	3.25 (2.97)	3.09 (2.91)	2.93 (2.85)	2.92 (2.72)	2.97 (2.84)	3.96 (3.08)	3.44 $\pm$0.56 (2.98 $\pm$ 0.23)	3.29 (2.93)

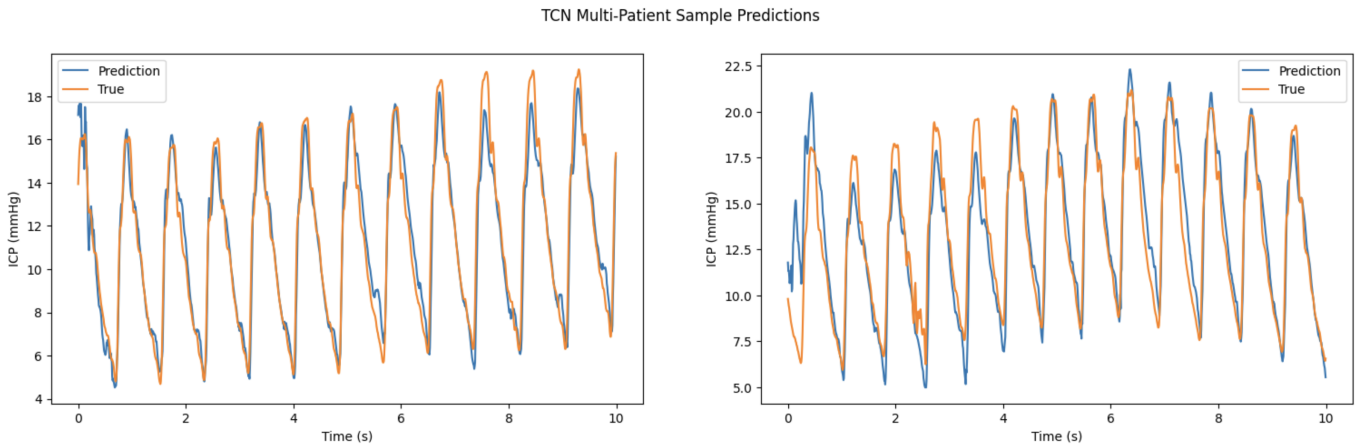


FIG. 3: TCN Multi-Patient Waveform Prediction Visualizations

Table V. Multi-Patient Analysis Results for TCN

Logcosh Test Loss (Train Loss) in mmHg	4.79 (1.78)
MAE Test Loss (Train Loss) in mmHg	5.44 (2.31)
RMSE Test Loss (Train Loss) in mmHg	6.78 (3.93)

ICP signal variance and prediction error, the variance of the ICP signal was calculated for each instance, and the Pearson correlation coefficient was computed between the prediction loss and the ICP signal variance. No statistically significant correlation was observed.

Lastly, to examine if the quality or similarity of input features had an effect on prediction error, all three input waveforms were reduced into a two-dimensional space using the Uniform Manifold Approximation and Projection (UMAP) Python library. Prediction error was overlaid onto these points and there was no clear trend of prediction error across clusters.

Overall, the results of this error analysis procedure suggest that the autoregulation index, mean ICP values, ICP signal variance, and clustering of features in a reduced dimensionality space are not significantly correlated with prediction accuracy. Ultimately, further investigation is required to determine the causal factor for poor prediction error in our algorithms.

C. Novelty

This research is the first of its kind to demonstrate a proof of concept that it is possible to accurately infer ICP from extracranial waveforms. The real-time and non-invasive nature of our algorithms makes it possible for them to be integrated into the current clinical workflow of ICP monitoring, a profound advancement towards eliminating the need for an IVC. In addition, by assessing multiple deep learning algorithms for accomplishing this task, we have also conducted a robust methodological study on the performance of deep learning algorithms on time-series data. The results of this study demonstrate that algorithmic prediction of ICP using a TCN performs better than current, less invasive alternatives and on par with alternative invasive options to the IVC.

D. Limitations

We acknowledge the limitations of this study. Firstly, the sample size of this study was small, consisting of only 10 patients, which may impede our ability to generalize model predictions to a broader population. Secondly, our application of a selective pre-processing scheme to discern high-quality waveform data could introduce subjectivity and bias into the training data, potentially compromising model performance and robustness. Finally, our study lacked external validation through the use of an independent dataset, thereby restricting our ability to evaluate the generalizability and reliability of the findings beyond the data sourced from the MIMIC III Waveform Database.

E. Future Direction

Forthcoming endeavors will concentrate on augmenting both the sample size of evaluated patients and deep learning algorithms implemented in this investigation. Although 10 patients were used in this study, their data quality proved invaluable in drawing clinically significant conclusions and correlations. Nevertheless, we plan to expand our inclusion criteria by accommodating missing values (presently, five-minute intervals that comprise any missing values) and adding additional patients, thereby enhancing patient representation and improving model generalizability. We will also test the models on an external dataset outside of the MIMIC III Waveform Database to assess the generalizability of model performance. In particular, we intend to externally validate our data on a more homogenous group of TBI patients using the CENTER-TBI database. Finally, we are currently examining the performance of four additional deep learning algorithms, namely simple RNN, GRU, VNET, and Transformer.

VII. ACKNOWLEDGEMENTS

We would like to acknowledge Johns Hopkins University for facilitating this study. Specifically all the staff who helped maintain Crunchr, Phoenix, and Rockfish. Thank you to Beth Israel Deaconess Medical Center for making MIMIC III data publicly available for analysis. We look forward to continuing our work.

References

1. C. A. Taylor, J. M. Bell, M. J. Breiding and L. Xu, "Traumatic Brain Injury-Related Emergency Department Visits, Hospitalizations, and Deaths - United States, 2007 and 2013," MMWR Surveill Summ, vol. 66, no. 9, pp. 1-16, Mar. 2017.
2. Centers for Disease Control and Prevention, "National Center for Health Statistics: Mortality Data on CDC WONDER," accessed April 2023, https://wonder.cdc.gov/mcd.html.
3. L. Shen, Z. Wang, Z. Su, S. Qiu, J. Xu, Y. Zhou, A. Yan, R. Yin, B. Lu, X. Nie, S. Zhao, and R. Yan, "Effects of Intracranial Pressure Monitoring on Mortality in Patients with Severe Traumatic Brain Injury: A Meta-Analysis," PLoS One, vol. 11, no. 12, pp. e0168901, Dec. 2016, https://doi.org/10.1371/journal.pone.0168901.
4. V. L. Pinto, P. Tadi, and A. Adeyinka, "Increased Intracranial Pressure," StatPearls, updated Aug. 4, 2021. Available: https://www.ncbi.nlm.nih.gov/books/NBK482119/
5. M. J. Albeck, S. E. Borgesen, F. Gjerris, et al., "Intracranial pressure and cerebrospinal fluid outflow conductance in healthy subjects," J. Neurosurg., vol. 74, pp. 597-600, Apr. 1991.
6. L. Rangel-Castilla, S. Gopinath, and C. S. Robertson, "Management of intracranial hypertension," Neurologic Clinics, vol. 26, no. 2, pp. 521-x, May 2008.
7. M. H. Wilson, "Monro-Kellie 2.0: The dynamic vascular and venous pathophysiological components of intracranial pressure," Journal of cerebral blood flow and metabolism: official journal of the International Society of Cerebral Blood Flow and Metabolism, vol. 36, no. 8, pp. 1338-1350, Aug. 2016. doi: 10.1177/0271678X16648711.
8. X. Zhang, J. E. Medow, B. J. Iskandar, F. Wang, M. Shokoueinejad, J. Koueik and J. G. Webster, "Invasive and noninvasive means of measuring intracranial pressure: a review," Physiological measurement, vol. 38, no. 8, pp. R143-R182, Aug. 2017, doi: 10.1088/1361-6579/aa7256.
9. M. Dey, J. Jaffe, A. Stadnik, and I. A. Awad, "External ventricular drainage for intraventricular hemorrhage," Current neurology and neuroscience reports, vol. 12, no. 1, pp. 24–33, 2012.
10. A. Silverman and N. H. Petersen, "Physiology, Cerebral Autoregulation," StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing, updated Mar 15, 2023. Available: https://www.ncbi.nlm.nih.gov/books/NBK553183/
11. J. T. Levis, "ECG Diagnosis: Deep T Wave Inversions Associated with Intracranial Hemorrhage," The Permanente journal, vol. 21, pp. 16-049, 2017. doi: 10.7812/TPP/16-049.
12. H. Liu, F. Pan, X. Lei, J. Hui, R. Gong, J. Feng, and D. Zheng, "Effect of intracranial pressure on photoplethysmographic waveform in different cerebral perfusion territories: A computational study," Frontiers in Physiology, vol. 14, article no. 1085871, 2023, doi: 10.3389/fphys.2023.1085871.
13. F. M. Kashif, G. C. Verghese, V. Novak, M. Czosnyka, and T. Heldt, "Model-based noninvasive estimation of intracranial pressure from cerebral blood flow velocity and arterial pressure," Sci. Transl. Med., vol. 4, no. 129, p. 129ra44, Apr. 2012, doi: 10.1126/scitranslmed.3003249.
14. K. B. Evensen, M. O'Rourke, F. Prieur, et al., "Non-invasive Estimation of the Intracranial Pressure Waveform from the Central Arterial Blood Pressure Waveform in Idiopathic Normal Pressure Hydrocephalus Patients," Sci. Rep., vol. 8, no. 1, p. 4714, Mar. 2018, doi: 10.1038/s41598-018-23142-7.
15. P. Wijayatunga, L. D. Koskinen, and N. Sundström, "Probabilistic prediction of increased intracranial pressure in patients with severe traumatic brain injury," Scientific Reports, vol. 12, no. 1, p. 9600, 2022. https://doi.org/10.1038/s41598-022-13732-x
16. H. -J. Lee, H. Kim, Y. -T. Kim, K. Won, M. Czosnyka and D. -J. Kim, "Prediction of Life-Threatening Intracranial Hypertension During the Acute Phase of Traumatic Brain Injury Using Machine Learning," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 10, pp. 3967-3976, Oct. 2021, doi: 10.1109/JBHI.2021.3085881
17. A. Uryga, A. Ziółkowski, A. Kazimierska, A. Pudełko, C. Mataczyński, E. W. Lang, M. Czosnyka, M. Kasprowicz, CENTER-TBI High-Resolution ICU (HR ICU) Sub-Study Participants and Investigators, and CENTER-TBI High-Resolution Sub-Study Participants and Investigators, "Analysis of intracranial pressure pulse waveform in traumatic brain injury patients: a CENTER-TBI study," J. Neurosurg., vol. Online, pp. 1-11, Oct. 2022, Art no. 2022.10.JNS221523, doi: 10.3171/2022.10.JNS221523.
18. D. Petrov, S. P. Miranda, R. Balu, C. Wathen, A. Vaz, V. Mohan, C. Colon, and R. Diaz-Arrastia, "Prediction of intracranial pressure crises after severe traumatic brain injury using machine learning algorithms," J. Neurosurg., vol. Online, pp. 1-8, Dec. 2022, Art no. 2022.12.JNS221860, doi: 10.3171/2022.12.JNS221860.
19. G. Ye, V. Balasubramanian, J. K. Li, and M. Kaya, "Machine Learning-Based Continuous Intracranial Pressure Prediction for Traumatic Injury Patients," IEEE J. Transl. Eng. Health Med., vol. 10, p. 4901008, 2022, doi: 10.1109/JTEHM.2022.3179874.
20. T. Goto, K. Furihata, and K. Hongo, "Natural resonance frequency of the brain depends on only intracranial pressure: clinical research," Sci. Rep., vol. 10, no. 1, p. 2526, Feb. 2020, doi: 10.1038/s41598-020-59376-7.
21. A. E. Johnson, T. J. Pollard, L. Shen, L. W. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark, "MIMIC-III, a freely accessible critical care database," Sci. Data, vol. 3, p. 160035, May 2016, doi: 10.1038/sdata.2016.35.
22. S. Meek and F. Morris, "ABC of clinical electrocardiography. Introduction. I-Leads, rate, rhythm, and cardiac axis," BMJ (Clin. Res. Ed.), vol. 324, no. 7334, pp. 415-418, Feb. 2002, doi: 10.1136/bmj.324.7334.415.
23. R. D. Stevens, M. Shoykhet, and R. Cadena, "Emergency Neurological Life Support: Intracranial Hypertension and Herniation," Neurocrit. Care, vol. 23 Suppl 2, no. Suppl 2, pp. S76-S82, Dec. 2015, doi: 10.1007/s12028-015-0168-z.
24. *Dynamic Time Warping* Information Retrieval for Music and Motion, Berlin, Heidelberg, Springer, 2007, pp. 69-84. doi: 10.1007/978-3-540-74048-3_4.
25. A. Nayebi, S. Tipirneni, B. Foreman, J. Ratcliff, C. K. Reddy, and V. Subbian, "Recurrent Neural Network based Time-Series Modeling for Long-term Prognosis Following Acute Traumatic Brain Injury," in AMIA Symposium, 2021, pp. 900-909.
26. D. Alis, C. Alis, M. Yergin, et al., "A joint convolutional-recurrent neural network with an attention mechanism for detecting intracranial hemorrhage on noncontrast head CT," Sci. Rep., vol. 12, no. 1, p. 2084, 2022, doi: 10.1038/s41598-022-05872-x.
27. S. Wang, C. Shao, J. Zhang, et al., "Traffic flow prediction using bi-directional gated recurrent unit method," Urban Info, vol. 1, no. 1, p. 16, 2022, doi: 10.1007/s44212-022-00015-z.
28. C. Li and G. Qian, "Stock Price Prediction Using a Frequency Decomposition Based GRU Transformer Neural Network," Appl. Sci., vol. 13, no. 1, p. 222, 2023, doi: 10.3390/app13010222.
29. S. Gao, J. Yang, T. Shen, and W. Jiang, "A Parallel Feature Fusion Network Combining GRU and CNN for Motor Imagery EEG Decoding," Brain Sci., vol. 12, no. 9, p. 1233, 2022, doi: 10.3390/brainsci12091233.
30. S. Rudregowda, S. Tippannavar, N. Sharma, M. Rashid, K. Bhat, and A. Rana, "EEG-based Brain Wave Recognition using GRU and LSTM," in 2022 International Conference on Innovative Computing and Communication (ICICC), 2022, pp. 1-5, doi: 10.1109/IC3I56241.2022.10072419.
31. A. Paul, "Prediction of missing EEG channel waveform using LSTM," in 2020 4th International Conference on Computational Intelligence and Networks (CINE), Kolkata, India, 2020, pp. 1-6, doi: 10.1109/CINE48825.2020.234393.
32. X. Ma, S. Qiu, Y. Zhang, X. Lian, and H. He, "Predicting Epileptic Seizures from Intracranial EEG Using LSTM-Based Multi-task Learning," in Pattern Recognition and Computer Vision, PRCV 2018, Lecture Notes in Computer Science, vol. 11257, Springer, Cham, 2018, doi: 10.1007/978-3-030-03335-4_14.
33. W. Liu, S. Du, N. Pang, L. Zhang, G. Sun, H. Xiao, Q. Zhao, L. Xu, Y. Yao, J. Alastruey, and A. Avolio, "Central Aortic Blood Pressure Waveform Estimation Based on Temporal Convolutional Network," in IEEE Journal of Biomedical and Health Informatics, vol. PP, 2023, doi: 10.1109/JBHI.2023.3268886.
34. K. Zhou, Z. Yin, Y. Peng, and Z. Zeng, "Methods for Continuous Blood Pressure Estimation Using Temporal Convolutional Neural Networks and Ensemble Empirical Mode Decomposition," in Electronics, vol. 11, no. 9, p. 1378, 2022, doi: 10.3390/electronics11091378.
35. Y. K. Musallam, N. I. AlFassam, G. Muhammad, S. U. Amin, M. Alsulaiman, W. Abdul, H. Altaheri, M. A. Bencherif, and M. Algabri, "Electroencephalography-based motor imagery classification using temporal convolutional network fusion," Biomedical Signal Processing and Control, vol. 69, p. 102826, 2021, doi: 10.1016/j.bspc.2021.102826.
36. A. Salami, J. Andreu-Perez, and H. Gillmeister, "EEG-ITNet: An Explainable Inception Temporal Convolutional Network for Motor Imagery Classification," in IEEE Access, vol. 10, pp. 36672-36685, 2022. doi: 10.1109/ACCESS.2022.3161489.
37. B.L. Hill, N. Rakocz, Á. Rudas, et al., "Imputation of the continuous arterial line blood pressure waveform from non-invasive measurements using deep learning," Scientific Reports, vol. 11, no. 1, p. 15755, 2021, doi: 10.1038/s41598-021-94913-y.
38. S. M. Mousavi, W. L. Ellsworth, W. Zhu, Y. Shi, A. M. Zarei and B. A. Romanowicz, "Earthquake Transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking," in Nature Communications, vol. 11, no. 1, p. 3952, 2020, doi: 10.1038/s41467-020-17591-w.
39. P. Verma and C. Chafe, "A Generative Model for Raw Audio Using Transformer Architectures," in 2021 24th International Conference on Digital Audio Effects (DAFx), Vienna, Austria, 2021, pp. 230-237, doi: 10.23919/DAFx51585.2021.9768298.
40. R. A. Saleh and A. K. M. E. Saleh, "Statistical Properties of the log-cosh Loss Function Used in Machine Learning," arXiv preprint arXiv:2208.04564, 2022. [Online]. Available: https://arxiv.org/abs/2208.04564.
41. P. Chen, G. Chen and S. Zhang, "Log Hyperbolic Cosine Loss Improves Variational Auto-Encoder," 2019. [Online]. Available: https://openreview.net/forum?id=rkglvsC9Ym.
42. X. Xu, J. Li, Y. Yang, and F. Shen, "Toward Effective Intrusion Detection Using Log-Cosh Conditional Variational Autoencoder," in IEEE Internet of Things Journal, vol. 8, no. 8, pp. 6187-6196, Apr. 2021. doi: 10.1109/JIOT.2020.3034621.
43. S. Gopali, F. Abri, S. Siami-Namini and A. S. Namin, "A Comparison of TCN and LSTM Models in Detecting Anomalies in Time Series Data," in 2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, 2021, pp. 2415-2420, doi: 10.1109/BigData52589.2021.9671488.
44. D. Cardim, C. Robba, M. Bohdanowicz, J. Donnelly, B. Cabella, X. Liu, M. Cabeleira, P. Smielewski, B. Schmidt, and M. Czosnyka, "Non-invasive Monitoring of Intracranial Pressure Using Transcranial Doppler Ultrasonography: Is It Possible?," Neurocrit Care, vol. 25, no. 3, pp. 473-491, Dec. 2016, doi: 10.1007/s12028-016-0258-6.
45. P. H. Raboel, J. Bartek Jr, M. Andresen, B. M. Bellander and B. Romner, "Intracranial Pressure Monitoring: Invasive versus Non-Invasive Methods-A Review," Crit. Care Res. Pract., vol. 2012, pp. 950393, 2012, doi: 10.1155/2012/950393.