



HAL
open science

A Deep Learning-Based Automated Framework for Subpeak Designation on Intracranial Pressure Signals

Donatien Legé, Laurent Gergelé, Marion Homme, Jean-Christophe Lapayre,
Yoann Launey, Julien Henriet

► **To cite this version:**

Donatien Legé, Laurent Gergelé, Marion Homme, Jean-Christophe Lapayre, Yoann Launey, et al..
A Deep Learning-Based Automated Framework for Subpeak Designation on Intracranial Pressure
Signals. *Sensors*, 2023, 23 (18), pp.7834 (17). hal-04224624

HAL Id: hal-04224624

<https://hal.science/hal-04224624>

Submitted on 2 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Deep Learning-Based Automated Framework for Subpeaks Designation on Intracranial Pressure Signal

Donatien LEGÉ^{1,3,‡,‡,*}, Laurent GERGELÉ^{2,‡}, Marion PRUD'HOMME^{1,†}, Jean-Christophe LAPAYRE³, Yoann LAUNEY⁴ and Julien HENRIET^{4,†}

¹*Sophysa*

²*UHSE*

³*uFC FEMTO-ST*

⁴*UHR*

*email@email.org

Abstract: \LaTeX The intracranial pressure (ICP) signal, as monitored on patients in intensive care units, contains pulses of cardiac origin on which P1 and P2 sub-peaks can often be observed. When calculable, the ratio of their relative amplitudes is an indicator of the patient's cerebral compliance. This information about the overall state of the cerebrospinal system is especially useful when it comes to adjusting sedation to the patient's needs.

We developed a recurrent neural network-based framework for P2/P1 ratio computation that only takes a raw PCI signal as an input. Two tasks are performed, namely pulses classification and subpeaks designation. Performances are evaluated on the basis of 10 labeled ICP recordings of one hour duration.

Pulses classification was achieved with an area under the curve of 0.90 on a 4344-pulse testing dataset, whereas peaks designation identified pulses with a P2/P1 ratio > 1 with a 97.92% accuracy.

ICP monitoring bedside devices can be improved with our real-time P2/P1 ratio calculation algorithm.

1. Introduction

Intracranial pressure (ICP) is classically monitored invasively in intensive care units (ICU) in the event of brain damage. One of the main objectives for the clinician is to limit the time spent by the patient above a threshold of cerebral hypertension, described by international guidelines [?]. However, the ICP signal is a combination of different periodic components, both affected by cardiac and respiratory frequencies. Thus, the only mean ICP cannot capture all the information provided by such a complex signal [?]. Especially, this single number does not describe the ability of the cerebrospinal system to compensate the changes in volume caused by blood and cerebrospinal fluid (CSF) displacements, so that the ICP is maintained into an acceptable range. This pressure-volume relationship, generally called "cerebral compliance", require the clinician specific manipulations to be measured punctually with CSF infusion tests [?] [?] [?]. That is why different characterizations of cerebral compliance, based on a mathematical analysis of ICP waveform, have been proposed in the literature [?] [?]. Notably, cardiac pulses morphology varies according to the cerebral compliance [?]. When the latter is at a normal state, three subpeaks of decreasing amplitudes are generally visible (see figure 1). Those peaks are called P1, P2 and P3, in accordance with their apparition order. While it is broadly admitted that P1 is due to the systolic pressure wave, the origin of P2 and P3 remain unclear [?]. MRI measurements tend to associate P2 with a maximum volume in the cerebral arteries [?] [?], whereas P3 could be linked to venous outflow [?]. In any case, as cerebral compliance is degraded, P2 and P3 become increasingly higher compared to P1 [?]. At the same time, their appearance times get closer [?], until the pulse takes a triangular shape centered on P2. Therefore, the ratio of the relative amplitudes of P2 and P1 (designated as the P2/P1 ratio) has been used as an indicator

47 of cerebral compliance [?]. This ratio is all the more relevant given that Kazimierska *et al.* [?] 48 demonstrated its good correlation with cerebral compliance assessed by classical infusion tests.

49 However, P1 and P2 automated detection on ICP signal faces different issues due to the 50 highly variable pulses morphology. Only a few automated frameworks allowing for P2 and P1 51 designation have been proposed in the literature [?] [?] [?]. Most of them rely on clustering 52 algorithms to only analyze one characteristic pulse over a predefined period, as proposed by the 53 authors of Morphological Clustering and Analysis of Continuous Intracranial Pressure (MOCAIP) 54 algorithm [?]. MOCAIP-based automated frameworks are designed to compute a large amount 55 of morphological features of the ICP pulses, including P2/P1 ratio. However, in addition to 56 the raw ICP signal, their data processing workflows require both an eletrocardiogram (ECG) 57 monitoring and an extensive reference library of non-artifact pulses, which can be difficult to 58 implement into an on-board bedside device. To perform real-time P2/P1 ratio calculation, neural 59 network-based algorithms seem to be the tools of choice to circumvent these prerequisites, due 60 to their ability to directly integrate the information provided by previous examples into trained 61 models. Especially, convolutional Neural Networks (CNN) and Long Short-Term Memory 62 (LSTM) recurrent networks have been successfully used for similar tasks, such as ECG beats 63 detection and classification (respectively [?] [?] [?] and [?] [?]).

64 Under the constraint of only using ICP signal as an input, we developed a deep learning-based 65 framework to detect the subpeaks P2 and P1, and compute the ratio of their relative amplitudes 66 when possible. Its conception was performed by achieving a comparative study of proposed 67 deep learning network architectures, enhanced with pre- and post-treatments and applied on 68 our dataset provided by the ICU of the University Hospital of Saint-Etienne. Our framework 69 is designed to perform two tasks sequentially. The first one is a classification task, aiming to 70 eliminate all the pulses without the P1 and P2 subpeaks. The second one, only performed on the 71 remaining pulses, aims to identify the subpeaks P1 and P2 to calculate the ratio of their relative 72 amplitudes. As an output, our framework provides a discontinuous signal of P2/P1 ratio values, 73 post-processed to make it as readable as possible for the clinician. In this article, we provide a 74 description of the neural network (NN) architectures we compared for pulse selection (3.2) and 75 for subpeaks designation (3.3). The performances obtained for each of the task are respectively 76 reported in sections (4.1) and (4.2). We finally tested our completed automated framework on a 77 dedicated testing dataset (section 4.3).

78 2. Dataset overview

79 The studied ICP signals came from 10 adult patients suffering from traumatic brain injury, 80 admitted to the ICU of the University Hospital of Saint-Etienne (France), between March 2022 81 and March 2023. For each of them, ICP was invasively monitored with an intraparenchymal 82 sensor (Pressio, Sophysa, Orsay, France) for a duration of 8.3 ± 5 days (min = 3.8, max = 15) at a 83 sampling frequency of 100Hz.

84 The dataset used in this study was constituted by randomly sampling 5 one-hour sections for 85 each record. 4 of them were affected to the training dataset, whereas the last one was affected 86 to the testing dataset. After the pulses were preprocessed and individualized as described in 87 section 3.1, 1 of out 15 was selected to be part of the final datasets. Those pulses were labellized 88 with the positions of P1 and P2 if both of them were visible, [0, 0] otherwise. The training 89 dataset was finally composed of 13,127 pulses, including 12,308 with a calculable P2/P1 ratio. 90 Its testing counterpart was composed of 4,344 pulses, including with 3847 a calculable P2/P1 91 ratio. This proportions are in accordance with Rashidinejad *et al.* ([?]) who estimated a missing 92 subpeak probability at less than 10% based on their 700-hour dataset.

93 To assess the performances of the final dataset, an additional 10-minute segment was randomly 94 sampled from each of the 10 patients. This second testing dataset, hence divided into 10 95 contiguous segments, was composed of 7,399 pulses, among which 6,815 had a calculable P2/P1

96 ratio.

97 **3. Materials and Methods**

98 Our data processing pipeline is divided into four parts. After preprocessing and a cardiac pulses
99 detection step, a selection is performed to eliminate all the pulses without a calculable P2/P1
100 ratio. The subpeaks are then designated on the remaining pulses. Finally, a postprocessing step
101 is performed to remove outliers and deal with missing values.

102 *3.1. Data preprocessing*

103 A fourth-order Butterworth bandpass filter between 0.3 Hz and 20 Hz is first applied to the raw
104 signal. It is meant to isolate cardiac pulses from rapid oscillations of electronic origin, respiratory
105 waves and baseline variations. The modified Scholkmann algorithm is then applied to the filtered
106 signal in order to detect the pulses onsets [?]. The characteristic duration L is set to 500 ms,
107 which offers a security margin compared to the quarter of a mean pulse duration recommended
108 as a minimum by the authors. The amplitude of each single pulse is normalized between 0
109 and 1, whereas the length is set to 180 points by a third degree polynomial interpolation. This
110 preprocessing step is close to the one performed by Mataczynski *et al.* ([?]) for pulse shape index
111 calculation. As an output, a $N \times 180$ matrix of N pulses is provided to the selection algorithm.

112 *3.2. Pulses selection*

113 A major difficulty in monitoring the P2/P1 ratio is that not all subpeaks are systematically
114 visible on all pulses. Therefore, a selection step is needed so that the detection algorithm is
115 only provided with pulses on which P1 and P2 are visible. This selection is performed by a
116 neural network. Three architectures are compared for this task, namely a 1-dimensional CNN, a
117 LSTM-based recurrent network and a Long Short-Term Memory Fully Convolutional Network
118 (LSTM-FCN), which is a combination of both. All the models are trained to perform the same
119 binary classification task, by minimizing a Binary Cross-Entropy (BCE) loss. Before calculating
120 it, a sigmoid function is applied to the neural networks outputs to obtain values between 0 and 1.

121 *3.2.1. 1-dimensional CNN architecture*

122 These architectures extract relevant features by applying convolutional filters on the input tensor.
123 CNN have been successfully used for medical images segmentation, but it is also possible
124 to adapt the layers dimensions to process 1-dimensional vectors the same way. Our CNN is
125 constituted of 3 encoding blocks, each one composed of the sequence Convolutional Layer- Batch
126 Normalization - ReLU activation, followed by a max pooling layer. The output is post-processed
127 by two dense layers separated with a ReLU activation layer. To reduce overfitting, a dropout
128 with a probability of 0.2 is applied at the end of the encoder and to the first dense layer. The
129 dimensions of each layer appear on figure 2.

130 *3.2.2. LSTM-based recurrent network*

131 Recurrent networks are designed to capture the underlying time dependencies of sequential data.
132 They are generally composed of one or more cells whose outputs are computed based on the
133 current input state and on the outputs of previous states. Past predictions can be taken into account
134 by different ways ; LSTM cells are specifically designed to track long-term dependencies [?]. The
135 proposed recurrent network is a single bi-directional LSTM cell, followed by two dense layers
136 separated by a ReLU activation. Hence, the input vector is processed in both reading directions
137 by the LSTM cell, which produces two outputs that are concatenated and post-processed by the
138 two dense layers. A dropout with a probability of 0.2 was applied at the end of the LSTM cell and
139 to the first dense layer. The dimensions of each layer appear on figure 2.

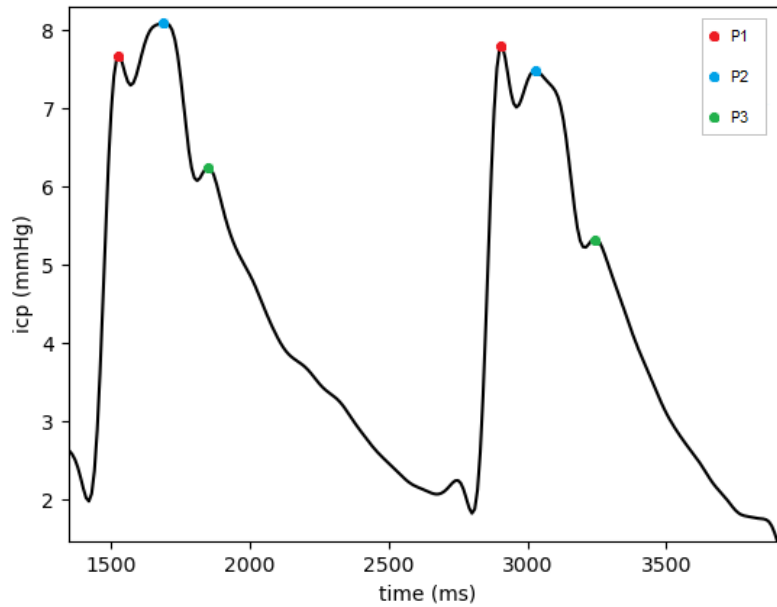


Fig. 1. Two pulses of cardiac origin on an ICP signal. The left one has a P2/P1 ratio > 1, whereas the right one has a P2/P1 ratio < 1 .

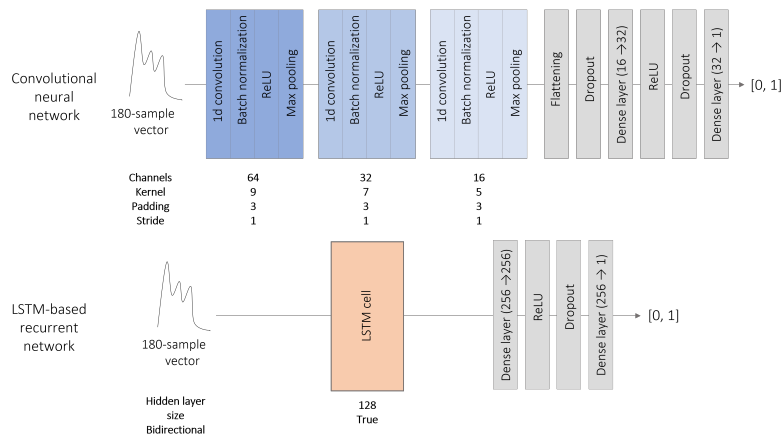


Fig. 2. CNN and LSTM-based recurrent network architectures used for pulses selection. In both cases, dropout was applied with a probability of 0.2. A sigmoid function was used to map the NN output into the interval [0, 1].

140 3.2.3. LSTM-FCN network

141 The two above-mentioned architectures process the input data with different objectives. Whereas
142 CNN focus on the neighborhood of each point, recurrent neural networks are meant to exploit the
143 causalities inherent to sequential data. LSTM-FCN networks attempt to combine both strategies,
144 and were specifically designed for time series classification [?]. Moreover, Mataczynski *et al.*
145 ([?]) obtained good results with such an architecture for pulse shape index calculation. The
146 LSTM-FCN network we implemented contains a three-block encoder, put in parallel with an
147 LSTM cell. Their respective dimensions are identical to those used for the CNN and for the
148 LSTM-based recurrent network. Both the computations are performed in parallel. The outputs
149 are then concatenated and processed by two dense layers. As above, a dropout with a probability
150 of 0.2 was applied to the first dense layer.

151 3.3. Subpeaks designation

152 Once the pulses with a calculable P2/P1 ratio are selected, subpeaks P1 and P2 can be designated.
153 To do so, we studied different ways of combining the output of a neural network with the pulse
154 curvature, as used by the MOCAIP-based automated frameworks. The latter is defined as:

$$\kappa(x) = \frac{x''}{(1+x'^2)^{3/2}}$$

155 On the other side, neural networks learn a classification task. For a pulse x , the objective is a
156 180-point vector y_x , such that

$$\forall t \in [[1, 180]], y_x(t) = e^{\frac{1}{2}} \left(e^{-\frac{(x(t)-p_1(x))^2}{2}} + e^{-\frac{(x(t)-p_2(x))^2}{2}} \right)$$

157 where $p_1(x)$ and $p_2(x)$ are the respective positions of P1 and P2. More formally, during the
158 learning process, the neural networks seeks a function f^* such that

$$f^* = \operatorname{argmin}_f \sum_{x \in D} \operatorname{MSE}(f(x), y_x)$$

159 Where *MSE* denotes the *Mean Square Error* loss function, and D the training set.

160 The detection strategy consists in designating P1 and P2 from among a candidate subpeaks set.
161 To do this, two methods were compared. In both cases, the candidate subpeaks are identified by a
162 search for local maxima, either on κ (method 1) or on f (method 2). Having thus obtained a list c
163 of candidates, p_1 and p_2 are then designated as the two points of c corresponding to the highest
164 value of f . Both strategies are summarized on figure 3. To perform the peaks designation task,
165 two networks architectures were compared, namely a 1-dimensional U-Net and a LSTM-based
166 recurrent network.

167 3.3.1. 1-dimensional U-Net

168 U-Net is a particular architecture of CNN. Its three-level bottleneck structure is composed of two
169 symmetric blocks. In addition to the linear information propagations, pairwise connections are
170 set between components of same shapes. As it was originally conceived for images segmentation,
171 layers have been here modified to perform 1-dimensional convolutions. Layers dimension appear
172 on figure 4. A dropout with a probability 0.2 was applied at each convolution block.

173 3.3.2. LSTM-based recurrent network

174 We used a bidirectional LSTM-based recurrent similar to the one trained for peaks selection
175 (see section 3.2.2). Hence, the input 180-sample pulse was processed by a single LSTM cell
176 followed by two consecutive dense layers. As hidden layer size of the LSTM cell was set to
177 180, the respective input and output dimensions of the latter were (360, 360) and (360, 180). A
178 dropout with a probability 0.2 was applied to the first dense layer.

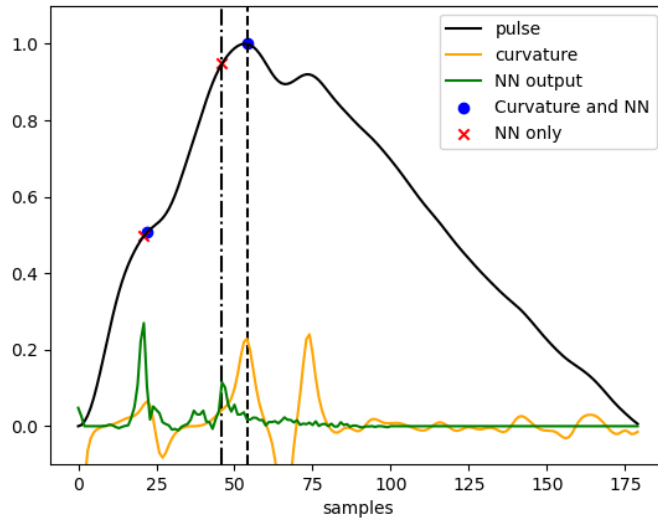


Fig. 3. Comparison of two methods of peaks designation algorithm. P1 and P2 are designated from among a set of candidates either based on the curvature analysis (method 1) or directly on the NN output (method 2).

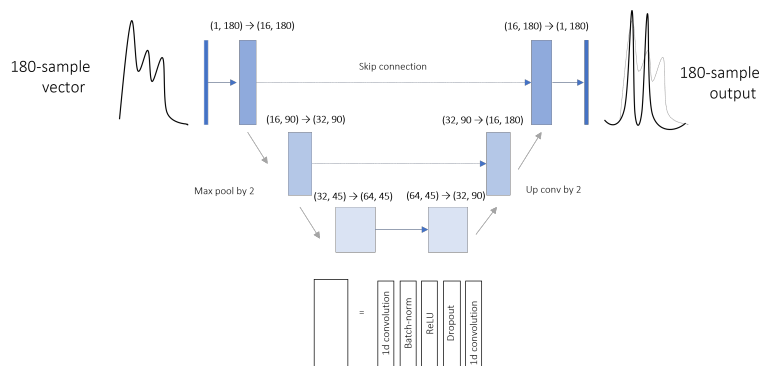


Fig. 4. U-Net architecture proposed for subpeaks detection. The NN learns to reconstitue the sum of two gaussian curves respectively centered on p_1 and p_2 .

179 3.4. Postprocessing

180 Postprocessing the P2/P1 ratio signal has to address three main issues:

- 181 • Spurious oscillations, mostly due to the intrinsic variability of the ICP signal. Even if they
182 are not a result of the data processing pipeline itself, they tend to make the record less
183 readable for the clinician.
- 184 • Missing values, since all the pulses that do not pass the selection cut are recorded as
185 missing.
- 186 • Punctal outliers. If they are not caused by the ICP signal itself, they can be due to errors in
187 the data processing pipeline. The latter can either occur at the classification step, when
188 false positive pulses are provided to the detection algorithm, or at the detection step, when
189 P1 and P2 are designated at wrong positions.

190 These different problems are alleviated at the post-processing phase, by retrospectively smoothing
191 the ratio monitoring. To do so, a 95% normal confidence interval is estimated on a 100-pulse
192 sliding window. A mean ratio is then calculated over the window if at least 50 values are
193 non-missing ; otherwise, the value corresponding to this window is reported as missing. Finally,
194 the output P2/P1 ratio signal can be displayed with a 100-pulse delay, which corresponds to about
195 one minute.

196 4. Results

197 Experiments were performed separately on the pulse selection and on the peaks detection tasks,
198 in order to select a single neural network for each of them. The same training and testing datasets
199 of labelled pre-processed pulses were used for both tasks, with 10% of the training set used
200 for validation. After having our framework completed with two trained neural networks, we
201 entirely processed 10-minute labelled segments randomly sampled from each of the recordings.
202 To ensure the reproductibility of our experiments, each of the three steps were performed using
203 a dedicated processing pipeline designed with Snakemake 7.25 [?]. All the associated scripts
204 were coded in Python 3.11. Neural networks were implemented with Pytorch 2.0 [?]. All the
205 experiments described below were performed on a Windows 10 machine powered by WSL2
206 Ubuntu 20.04.5, equipped with a 12th Gen Intel(R) Core(TM) i7-12850HX 2.10 GHz 16 CPU,
207 a Nvidia RTX A3000 12GB Laptop GPU, and 16 GB of RAM. Pipelines used for comparing
208 neural networks performances are available at the following address: _

209 4.1. Pulse selection

210 The three models (i.e. CNN, LSTM recurrent network and LSTM-FCN) were trained on 150
211 epochs with the Adam optimizer, an initial learning rate of 0.001 and a batch size of 256. For each
212 of them, the area under the receiver operating characteristic curve (ROC) curve was calculated by
213 plotting the True Positive Rate (TPR) against the False Positive Rate (FPR), defined as:

$$TPR = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}, FPR = \frac{\text{False Positive}}{\text{False Positive} + \text{True Negative}}$$

214 The three ROC curves are displayed on figure 5. For the final framework, the optimal decision
215 threshold was chosen to maximize the difference $TPR - FPR$.

216 Our LSTM-based recurrent network architecture overperformed the convolution-based ones,
217 with an area under the curve of 0.903. The confusion matrices corresponding to the respective
218 optimal decision thresholds of each NN architecture are presented in table 1.

219 The amounts of false-positive pulses and false-negative pulses correspond to respectively 1.8%
220 and 9.7% of the total testing data set when using the LSTM-based architecture for classification.

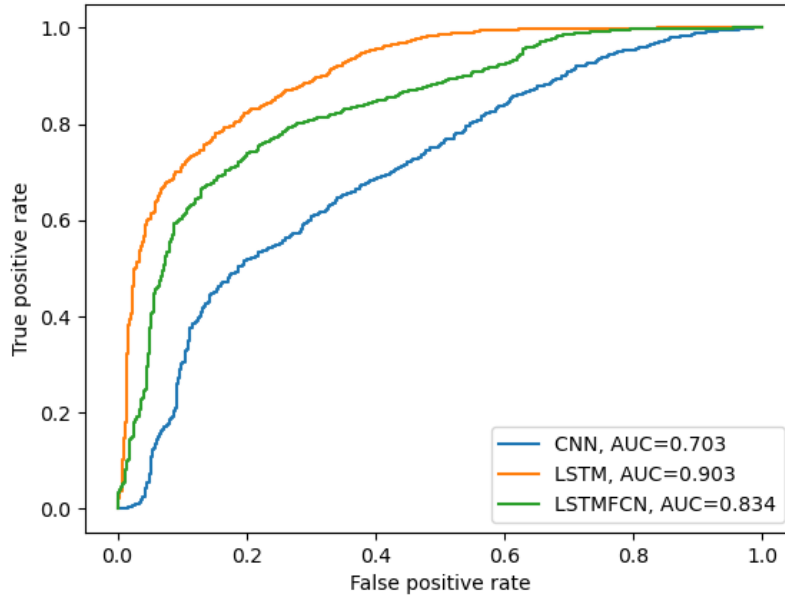


Fig. 5. Areas under the ROC curve (AUC) of the three neural network architectures used for pulses selection. Positive class corresponds to pulses with a calculable P2/P1 ratio.

Table 1. Confusion matrices of the 3 NN architectures compared for pulses selection. Positive class corresponds to pulses with a calculable P2/P1 ratio.

NN architecture	CNN		LSTM		LSTM-FCN	
	-	+	-	+	-	+
Prediction						
True -	399	98	421	76	397	100
True +	1865	1982	847	3000	1005	2842
True Positive Rate (%)	51.5		78.0		11.2	
False Positive Rate (%)	19.7		8.52		26.1	

221 In contrast, this percentages amount to respectively 2.3% and 42.9% when using the convolutive
222 network.

223 4.2. Peaks designation

224 The experimental pipeline was designed to compare the four possible combinations between peak
225 designation method (i.e., by using or not the curvature function) and neural network architecture
226 (i.e., 1-d convolutional U-Net or LSTM-based recurrent network). In addition, a designation only
227 using the first two local maxima of curvature was performed as a baseline. Both models were
228 trained on 150 epochs with the Adam optimizer, an initial learning rate of 0.001 and a batch size
229 of 256. Mean average time appearance error, exprimed in percentage of the whole pulse duration,
230 and mean average ratio error were calculated. The results are reported in table 2. In addition, as
231 it is the most interpretable information for the clinician, we assessed the ability of our models to
232 detect pulses where P2 is higher than P1. To do so, we calculated a confusion matrix for classes
233 "+": "P2/P1 ratio > 1" and "-": "P2/P1 ratio < 1" and the associated accuracy, defined as the
234 proportion of correct predictions over the whole testing dataset.

235 As for the pulse selection task, the recurrent architecture overformed the convolutional one.
236 Without the curvature-based candidate peaks selection step, the LSTM-RE architecture performed
237 the classification task with an accuracy 3% higher than our 1d-Unet. Moreover, it achieved the
238 most accurate estimation of the P2/P1 ratio, with a mean average error of 0.03. Achieving the
239 candidate peaks selection step with the means of the curvature function tends to improve the
240 algorithm's ability to discriminate pulses with a P2/P1 ratio > 1, at the cost of a slightly less
241 accurate ratio estimation.

242 4.3. Final automated framework

243 On the basis of previous experiments, we finally chose a LSTM-based recurrent both for pulses
244 selection and for subpeaks designation. For the latter step, candidate subpeaks selection was
245 performed using the pulse curvature. For each of the ten patients, the complete workflow was
246 used to process a randomly chosen labelled 10-minute section. An example of such an output is
247 presented figure 6.

248 The performances were assessed for each individual 10-minute segment. We used the same
249 respective metrics as above for pulses selection and subpeaks designation. In addition, we
250 calculated the percentage of pulses that have been assigned a ratio value, and the percentage of
251 non-missing values in the final post-processed ratio signal. Table 3 contains value calculated over
252 the total 110-min dataset, but 10-min segment individualized metrics are available table A??.

253 False positive rate and true positive rate are both about 7 points higher than their respective
254 equivalents calculated when selecting the NN architecture. However, subpeaks designation
255 performances are consistent with previous experiments. Table 4 corresponds to the overall
256 confusion matrix calculated for pulses selection. As above, individualized confusion matrices
257 are available table ??.

258 It is noticeable that the only 2nd segment sample contains 91% of the negatively labeled pulses.
259 In this segment, pulse selection algorithm performed with a 13.5% false positive rate (table ??).

260 False-positive pulses and false-negative pulses amount to respectively 1.14% and 7.49% of
261 the total testing dataset. This proportions are consistent with those previously calculated on the
262 4344-pulse testing dataset.

263 5. Discussion

264 Our deep-learning based framework is designed to perform P1 and P2 detection and P2/P1 ratio
265 computation directly on a bedside device. For convenience concerns, we designed it under
266 the constraint of only using the ICP signal, which was made possible by a well-established
267 efficient preprocessing step. Hence, we were able to focus our deep-learning based analysis

Table 2. Performances of five methods for P1 and P2 detection. P1 and P2 are designated as the two candidate subpeaks corresponding to the two highest NN output value. Local maxima of either curvature or NN output are selected as candidate subpeaks. As a baseline, the algorithm “Curvature” corresponds to the designation of the two first local maxima of pulse curvature as P1 and P2. Mean absolute errors (MAE) on the appearance time of P1 and P2 are expressed in percentage of the total pulse duration..

-0cm CCCCC Algorithm	Candidate peaks selection	P1 MAE (%)	P2 MAE (%)	Ratio MAE Accuracy(%)
[m]2*1d-Unet NN output	1.2±0.1	2.1±0.2	0.08±0.03	93.2
Curvature	0.6±0.05	2.2±0.2	0.05 ±0.02	96.6
[m]2*LSTM NN output	0.70 ±0.05	1.3±0.07	0.03±0.003	96.9
Curvature	0.70±0.06	1.3±0.2	0.05±0.02	97.9
[m]1*Curvature	-2.4±0.2	4.0±0.2	0.1±0.01	89.3

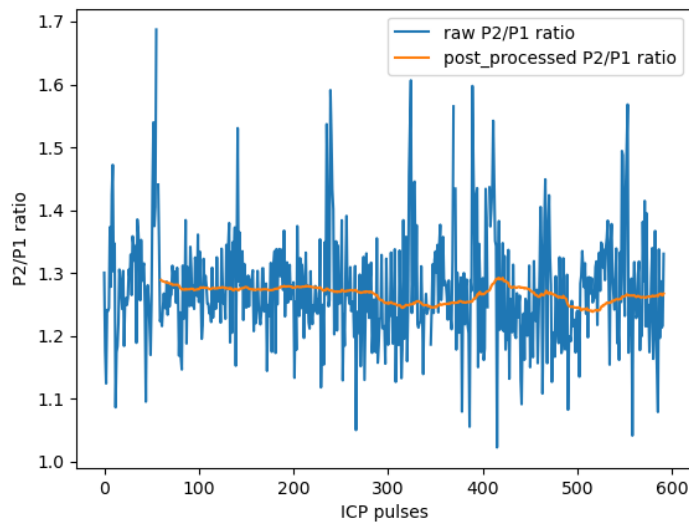


Fig. 6. Example output for a 10-minute ICP signal segment processed with the final automated framework.

Table 3. Performances of the final automated P2/P1 ratio computation framework. Metrics associated with P2/P1 ratio values (*i.e.*, P2/P1 ratio MAE and Accuracy on ratio > 1 detection) are calculated pulses with a labeled P2/P1 ratio value that passed the selection step.

-0cm CCCCC	True positive rate (%)	False positive rate (%)	P2/P1 Ratio MAE	Accuracy on ratio > 1 detection(%)	Ratio-associated pulses (%)	Displayed-ratio time(%)
87.3*	14.6*	0.044 ± 0.002	99.7*	85.8	88.3	

* Significantly higher than the same metric calculated on the testing set during NN selection (*p-value* < 0.05)

268 on short time series corresponding to single pulses of cardiac origin, which benefited to
269 not excessively deep networks architectures. Moreover, working at the cardiac cycle scale
270 allowed us to alleviate another real-life difficulty: At bedside monitoring, ICP signals are very
271 often contaminated with artifacts either due to patient movements (coughing, reactions to drug
272 administration, nursing manipulations...), or to electronic perturbations. Therefore, it can be
273 complicated, at a macroscopic scale, to determine whether an acute rise in ICP corresponds
274 to a real physiological measurement or to an artifacted zone. By only focusing on modified
275 Scholkmann algorithm-extracted candidates pulses, we were able to perform this artifact removal
276 step on the only basis of the local waveform, at the pulse selection step. In addition, as changes
277 in cerebral compliance generally occur in a progressive way, a continuous pulse-wise compliance
278 score is a tool of choice to describe as faithfully as possible the current patient state.

279 When labeling the pulses, only using the ICP signal could sometimes cause difficulties for
280 interpreting isolated single pulse waveform: Without other elements of context, pulses with only
281 two visible subpeaks systematically fell into the “non-calculable P2/P1 ratio“ category, since it
282 was not possible to know which of P1, P2 or P3 was missing. In some of these cases, ABP or
283 ECG signals may have helped to distinguish subpeaks, and thus to compute a P2/P1 ratio. In that
284 sense, the training dataset was labelled in a quite restrictive way, to limit as much as possible the
285 amount of pulses without a calculable P2/P1 ratio provided to the peak designation step. However,
286 this decision has necessarily consequences on the amount of time during which a P2/P1 ratio can
287 be displayed. In any case, recurrent architectures clearly overperformed the convolutional-based
288 ones for pulse selection, even it is probably possible to reduce the observed gap by fine-tuning
289 the proposed convolutional architecture. As the full succession of subpeaks is necessary to
290 understand the pulse waveform, recurrent networks seem to be more appropriate than CNNs to
291 perform such a classification task. In that sense, these results may contrast with similar studies
292 performed on ECG signal, on which events such as QRS complexes have more recognizable
293 shapes and thus make CNN more relevant for classification or detection tasks. Concerning the
294 consequences of misclassified pulses, it is noticeable that false-negative pulses only cause
295 spurious missing values at the end of the data processing workflow. In contrast, false-positive
296 pulses are provided to a peak designation algorithm that systematically outputs the two positions
297 of estimated P1 and P2. Therefore, the latter can do much more damage to the output P2/P1 ratio
298 signal. While we simply chose an optimal threshold that minimizes the difference $TPR - FPR$,
299 it could be relevant to optimize the decision threshold to make the algorithm more restrictive.

300 Peak detection was performed by computing a density function by the means of neural networks,
301 as it is often the case for image segmentation tasks. We chose to stick to the underlying philosophy
302 of MOCAIP-based automated framework, which include a candidates selection step before
303 subpeaks designation. It would have been possible to turn our algorithm into a regression task
304 to directly output the estimated positions, as it is sometimes done for ECG peaks detection [?].
305 This simpler strategy lead to lighter computations. However, our method offers two advantages.
306 Firstly, it is more robust and explainable in itself, as a score is affected to each point of the
307 input tensor. Secondly, it is easier to combine the output tensor with another function such as
308 the pulse curvature. Designating two peaks from among a set of candidates selected with this
309 simple and explainable criterion offers guarantees for the generalization abilities of the algorithm.
310 This is all the more relevant given that we could only train our deep learning-based models on a
311 relatively small set of patients, whereas there is a large inter-patient morphological variability
312 in ICP waveform. In the case of our testing dataset, a preselection of candidate peaks with a
313 search for local maxima of the curvature function improved the algorithm’s ability to discriminate
314 pulses with a P2/P1 ratio superior to 1. The observed improvements in accuracy amounted to
315 respectively 1% for the recurrent network and to 3% for our U-Net.

316 The biggest limitation of our study is that only 10 patients recordings contributed to the pulses
317 database. Because of this small number, we chose to include samples from each of the ten

318 patients both in the training and in a testing datasets, in order to train our neural networks with as
319 much diversity as possible. By doing this, we made the assumption that a single patient CP signal
320 variability over 8 days (that is to say, the average monitoring duration) was enough to neglect the
321 effects of a commune underlying distribution. However, generalization abilities of our automated
322 framework still have to be improved by expanding our datasets with further inclusions. This is all
323 the more important that we obtained quite different false positive rates during the model selection
324 (8.52%) and during the final automated framework evaluation (14.6%).

325 While designing the data processing pipeline, we considered better taking into account the
326 neighborhood of each single pulse. For instance, the pulse selection process could have integrated
327 all the pulses occurred over the last minute before the one to be classified, thus helping the
328 interpretation of pulse waveform. However, it would have require a much more computation-
329 intensive training step, since the recurrent networks would have had to capture more long-term
330 dependencies. In addition, the database would have had to be composed of contiguous labelled
331 samples, which would have had drawbacks on the diversity covered this way. We faced the exact
332 same issue when sampling the final testing dataset, which was particularly disbalanced with 90%
333 of its false-negative pulses occurring in the same segment.

334 The latter observation leads us to discuss the main drawbacks of monitoring the P2/P1 ratio. As
335 mentioned earlier, this information is not always available, and depends on biological mechanisms
336 still not fully understood [?]. A more complete picture of cerebral compliance could be obtained
337 by combining P2/P1 ratio with other indicators such as mean ICP, pulse amplitude [?] or pulse
338 shape index [?]. More generally, cerebral compliance has to be considered as part of a bundle of
339 information available on patients. Characterizing it is especially helpful when ICP is close to the
340 hypertension threshold, as a simple mean calculation is not informative enough on the current
341 state of the cerebrospinal system. Cerebral compliance may also inform specific decisions, for
342 instance when it comes to adjusting or putting sedation to an end.

343 References

- 344 1. N. Carney, A. M. Totten, C. O'Reilly, *et al.*, "Guidelines for the management of severe traumatic brain injury,"
345 *Neurosurgery* **80**, 6–15 (2017).
- 346 2. M. Czosnyka, P. Smielewski, I. Timofeev, *et al.*, "Intracranial pressure: more than a number," *Neurosurg. focus* **22**,
347 1–7 (2007).
- 348 3. A. L. Maset, A. Marmarou, J. D. Ward, *et al.*, "Pressure-volume index in head injury," *J. neurosurgery* **67**, 832–840
349 (1987).
- 350 4. E. R. Cardoso, J. O. Rowan, and S. Galbraith, "Analysis of the cerebrospinal fluid pulse wave in intracranial pressure,"
351 *J. Neurosurg.* **59**, 817–821 (1983). Publisher: Journal of Neurosurgery Publishing Group Section: Journal of
352 *Neurosurgery*.
- 353 5. A. Vallet, L. Gergel , E. Jouanneau, *et al.*, "Assessment of pressure-volume index during lumbar infusion study:
354 What is the optimal method?" *Intracranial Press. Neuromonitoring XVII* pp. 335–338 (2021).
- 355 6. M. Czosnyka, E. Guazzo, M. Whitehouse, *et al.*, "Significance of intracranial pressure waveform analysis after head
356 injury," *Acta neurochirurgica* **138**, 531–542 (1996).
- 357 7. A. Spiegelberg, M. Krause, J. Meixensberger, and V. Kurtcuoglu, "Raq: a novel surrogate for the craniospinal
358 pressure–volume relationship," *Physiol. Meas.* **41**, 094002 (2020).
- 359 8. S. Brasil, D. J. F. Solla, R. d. C. Nogueira, *et al.*, "Intracranial compliance assessed by intracranial pressure pulse
360 waveform," *Brain Sci.* **11**, 971 (2021).
- 361 9. A. A. Domogo, P. Reinstrup, and J. T. Ottesen, "Mechanistic-mathematical modeling of intracranial pressure (icp)
362 profiles over a single heart cycle. the fundament of the icp curve form," *J. Theor. Biol.* **564**, 111451 (2023).
- 363 10. E. Carrera, D.-J. Kim, G. Castellani, *et al.*, "What shapes pulse amplitude of intracranial pressure?" *J. neurotrauma*
364 **27**, 317–324 (2010).
- 365 11. M. Unnerb ck, J. T. Ottesen, and P. Reinstrup, "Icp curve morphology and intracranial flow-volume changes: a
366 simultaneous icp and cine phase contrast mri study in humans," *Acta Neurochir.* **160**, 219–224 (2018).
- 367 12. M. Czosnyka and Z. Czosnyka, "Origin of intracranial pressure pulse waveform," *Acta Neurochir.* **162**, 1815–1817
368 (2020).
- 369 13. K. Germon, "Interpretation of icp pulse waves to determine intracerebral compliance." *The J. neuroscience nursing:*
370 *journal Am. Assoc. Neurosci. Nurses* **20**, 344–351 (1988).
- 371 14. A. Zi lkowski, A. Pudelko, A. Kazimierska, *et al.*, "Peak appearance time in pulse waveforms of intracranial pressure
372 and cerebral blood flow velocity," *Front. Physiol.* **13**, 2670 (2023).

- 373 15. D. A. Godoy, S. Brasil, C. Iaccarino, *et al.*, "The intracranial compartmental syndrome: a proposed model for acute
374 brain injury monitoring and management," *Crit. Care* **27**, 1–9 (2023).
- 375 16. A. Kazimierska, M. Kasprowicz, M. Czosnyka, *et al.*, "Compliance of the cerebrospinal space: Comparison of three
376 methods," *Acta neurochirurgica* **163**, 1979–1989 (2021).
- 377 17. H.-J. Lee, E.-J. Jeong, H. Kim, *et al.*, "Morphological feature extraction from a continuous intracranial pressure
378 pulse via a peak clustering algorithm," *IEEE Trans. on Biomed. Eng.* **63**, 2169–2176 (2015).
- 379 18. X. Hu, T. Glenn, F. Scalzo, *et al.*, "Intracranial pressure pulse morphological features improved detection of decreased
380 cerebral blood flow," *Physiol. measurement* **31**, 679 (2010).
- 381 19. P. Rashidinejad, X. Hu, and S. Russell, "Patient-adaptable intracranial pressure morphology analysis using a
382 probabilistic model-based approach," *Physiol. measurement* **41**, 104003 (2020).
- 383 20. X. Hu, P. Xu, F. Scalzo, *et al.*, "Morphological clustering and analysis of continuous intracranial pressure," *IEEE*
384 *Trans. on Biomed. Eng.* **56**, 696–705 (2008).
- 385 21. L. Xiaolin, B. Cardiff, and D. John, "A 1d convolutional neural network for heartbeat classification from single lead
386 ecg," in *2020 27th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, (IEEE, 2020), pp.
387 1–2.
- 388 22. J. Jang, S. Park, J.-K. Kim, *et al.*, "Cnn-based two step r peak detection method: Combining segmentation and
389 regression," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society*
390 *(EMBC)*, (IEEE, 2022), pp. 1910–1914.
- 391 23. M. Chourasia, A. Thakur, S. Gupta, and A. Singh, "Ecg heartbeat classification using cnn," in *2020 IEEE 7th Uttar*
392 *Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, (IEEE,
393 2020), pp. 1–6.
- 394 24. S. Boda, M. Mahadevappa, and P. K. Dutta, "An automated patient-specific ecg beat classification using lstm-based
395 recurrent neural networks," *Biomed. Signal Process. Control.* **84**, 104756 (2023).
- 396 25. H. M. Rai and K. Chatterjee, "Hybrid cnn-lstm deep learning model and ensemble technique for automatic detection
397 of myocardial infarction using big ecg data," *Appl. Intell.* **52**, 5366–5384 (2022).
- 398 26. S. M. Bishop and A. Ercole, "Multi-scale peak and trough detection optimised for periodic and quasi-periodic
399 neuroscience data," in *Intracranial Pressure & Neuromonitoring XVI*, (Springer, 2018), pp. 189–195.
- 400 27. C. Mataczyński, A. Kazimierska, A. Uryga, *et al.*, "End-to-end automatic morphological classification of intracranial
401 pressure pulse waveforms using deep learning," *IEEE J. Biomed. Health Informatics* **26**, 494–504 (2021).
- 402 28. H. Sak, A. Senior, and F. Beaufays, "Long short-term memory based recurrent neural network architectures for large
403 vocabulary speech recognition," *arXiv preprint arXiv:1402.1128* (2014).
- 404 29. F. Karim, S. Majumdar, H. Darabi, and S. Chen, "Lstm fully convolutional networks for time series classification,"
405 *IEEE access* **6**, 1662–1669 (2017).
- 406 30. F. Mölder, K. P. Jablonski, B. Letcher, *et al.*, "Sustainable data analysis with snakemake," *F1000Research* **10** (2021).
- 407 31. A. Paszke, S. Gross, F. Massa, *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *CoRR*
408 **abs/1912.01703** (2019).
- 409 32. L. D'Antona, C. L. Craven, F. Bremner, *et al.*, "Effect of position on intracranial pressure and compliance: a
410 cross-sectional study including 101 patients," *J. Neurosurg.* **136**, 1781–1789 (2021).
- 411 33. C. G. Nucci, P. De Bonis, A. Mangiola, *et al.*, "Intracranial pressure wave morphological classification: automated
412 analysis and clinical validation," *Acta neurochirurgica* **158**, 581–588 (2016).

Table 4. Confusion matrix obtained for the final pulses selection step. Positive class corresponds to pulses with a calculable P2/P1 ratio.

CCC	Predicted -	Predicted +
True -	499	85
True +	554	6261