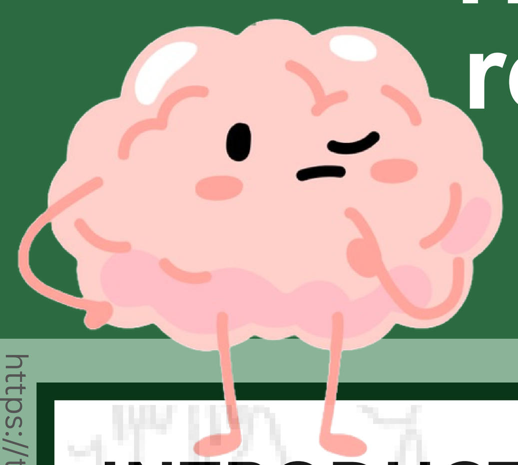


How to downsample large EEG signals, keeping clinically relevant waveforms while minimizing end-to-end latency

Erik Koprivanacz[1] Supervisors: Ricardo Marroquim[1], Arthur-Ervin Avramiea[2]

[1]Delft University of Technology, [2]Vrije University Amsterdam

INTRODUCTION

Electroencephalography (EEG) is one of the most widely used non-invasive techniques for measuring brain activity. It is relatively cheap and measures data at high frequencies; however, this also poses the challenge of visualizing large amounts of data in real time. While modern EEG analysis toolkits provide methods for resampling data, their primary objective is to preserve the signal's statistical properties, which introduces computational overhead and may not preserve the data's visual properties. Thus, **the goal of this project is to investigate whether different downsampling methods could achieve superior visual fidelity at a faster runtime.**

DOWNSAMPLING ALGORITHMS

Five **candidate algorithms** were selected for evaluation, which cover a wide range of traditional downsamplers:

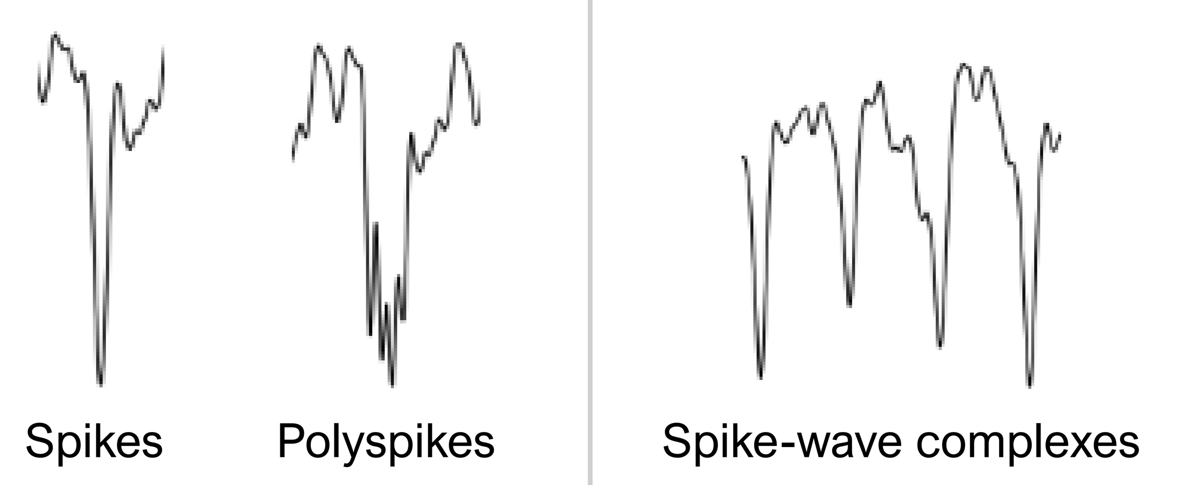
1. **Uniform Stride:** takes every nth point
2. **Min-Max:** divides the data into buckets, preserves the minimum and maximum
3. **M4:** like Min-Max but also preserves first and last points
4. **Largest Triangle Three Buckets:** divides data into buckets, selects points with the largest triangle areas across buckets
5. **Min-Max Largest Triangle Three Buckets:** Min-Max preselection with LTTB applied after

Additionally, four modern **EEG analysis toolkits** were compared with the candidate algorithms: Fieldtrip and EEGLab, MATLAB EEG analysis frameworks, while MNE is their counterpart in Python. Additionally, SciPy was added as a general-purpose signal processing framework written in Python.

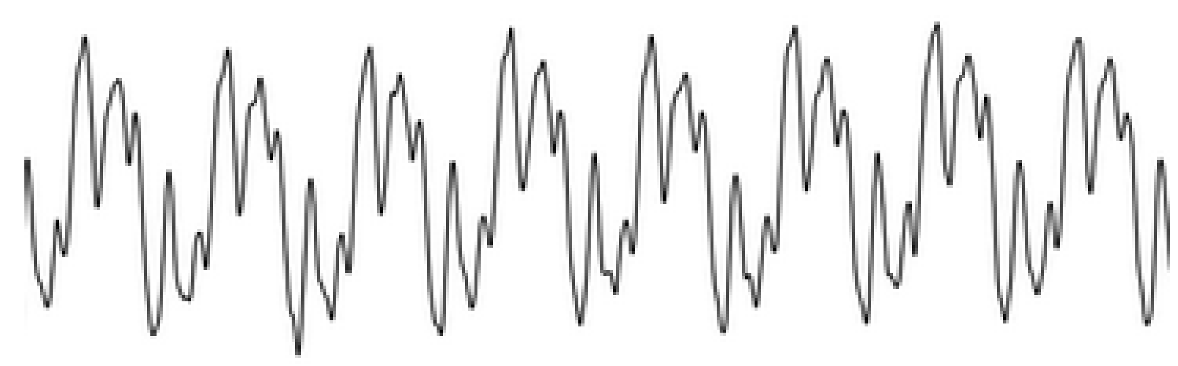
These frameworks, unlike the candidate algorithms, were **designed to preserve the signal's statistical properties.**

EEG WAVEFORMS

To assess visual fidelity, EEG data was downsampled and compared. Thus, to accurately assess performance, it is important to be familiar with EEG waveforms. This study focused on **interictal epileptiform discharges (IEDs)**, which are only a subset of clinically relevant waveforms; thus, in the future, it would be crucial to assess the results of this project on other types of EEG data. Examples of IEDs are:



An example of a EEG signal without anomalies is:



VISUAL EVALUATION

The algorithms were evaluated against the raw signal using **Intersection over Union (IoU) and a set of six metrics** designed to capture various details of EEG signals. The six metrics are: spike preservation, direction change preservation, complexity preservation, Euclidean distance, shape-based distance, and time-series structural similarity. Overall, the **candidate algorithms scored high** across the metrics and over IoU (fig. 1), **outperforming the EEG analysis toolkit methods**, meaning they capture the details of the original signal well.

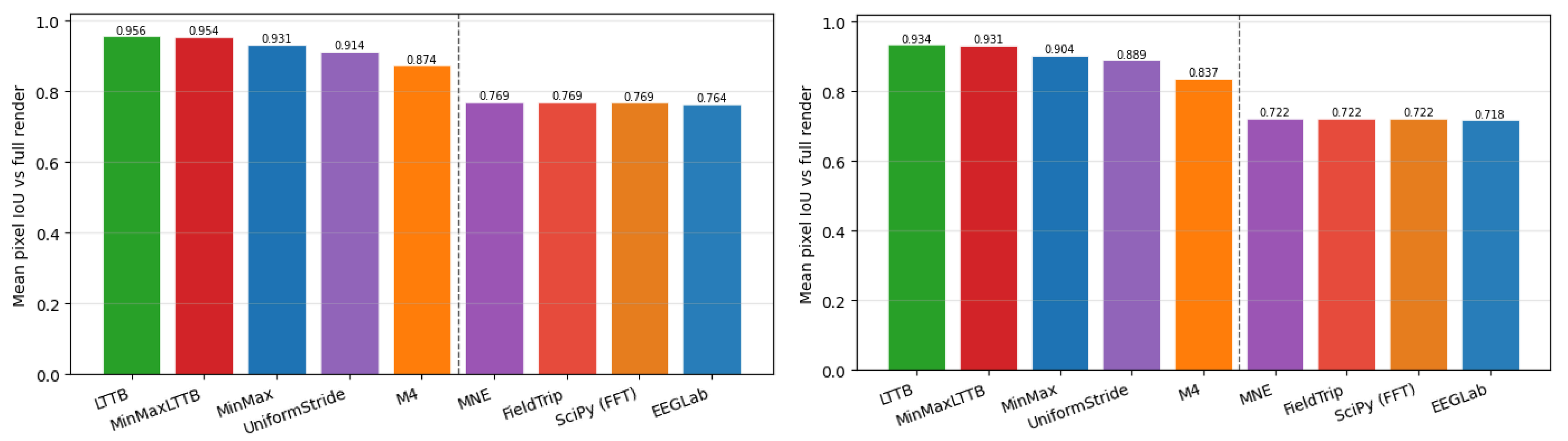


Fig. 1: IoU vs raw signal at HD (left) and 4K (right) resolution

INTER-ALGORITHM AGREEMENT

Since the algorithms scored similarly during visual evaluation, an inter-algorithm comparison was performed to determine how closely their outputs matched one another and the EEG analysis toolkit methods. **Similarity was quantified using IoU and structural similarity.**

The algorithms **scored similarly to each other, except for M4** (fig. 2). This is because M4 created less granular buckets. On the other hand, **agreement with the EEG analysis toolkit methods was substantially lower**, further reinforcing the fact that these algorithms were optimized for **different use cases**.

Fig. 2: IoU against each other (left) and against the EEG methods (right)

PERFORMANCE EVALUATION

Lastly, the algorithm's performance was quantified using **runtime analysis**. The median runtime and standard deviation were calculated over 492 ten-second windows. (tab. 1) While uniform stride and M4 perform the fastest, both introduce potential risks of missing critical features after downsampling. Thus, the **fastest algorithm with adequate visual fidelity is Min-Max**. The **EEG analysis methods showed longer runtimes** than the candidate algorithm, reflecting their focus on preserving statistical signal properties.

Algorithm	Median (ms)	Std (ms)
Uniform Stride	0.0053	0.0077
M4	0.0361	0.0568
Min-Max	0.0392	0.0259
LTTB	0.0520	0.0297
Min-Max LTTB	0.0876	0.0510
SciPy (FFT)	0.1118	0.0576
MNE	0.4299	0.2049
EEGLab	29.0227	0.7681
FieldTrip	39.1767	9.4338

Tab. 1: runtime of algorithms

Learn More!

CONCLUSIONS

In summary, when both visual fidelity and computational efficiency is taken into account, **Min-Max provides the best tradeoff.** LTTB follows closely behind providing better visualization at the cost of slightly longer runtime.

The **EEG analysis toolkit algorithms overall performed** worse in both the visual and performance evaluations, underscoring that they were optimized for different use cases.