Work-let Name: LLM For Physiological Data

CLHi-MTS: Hierarchical Modeling for Medical Time-Series

Focus of the Study

• The research introduces CLHi-MTS. This is a novel, self-supervised pretraining framework for analyzing medical time-series data (like ECGs, MEGs, and EEGs).

• The goal is to automatically extract meaningful and robust feature representations from raw data without needing extensive manual labeling.

• The extracted representations are designed to be general (universal) and are then fine-tuned for classification tasks.

Problem Addressed

• **Data Complexity and Labeling:** Modern medical signals are complex and high-dimensional, making manual data annotation challenging and time-consuming for experts. Self-supervised learning is needed to address this scarcity of labeled data.

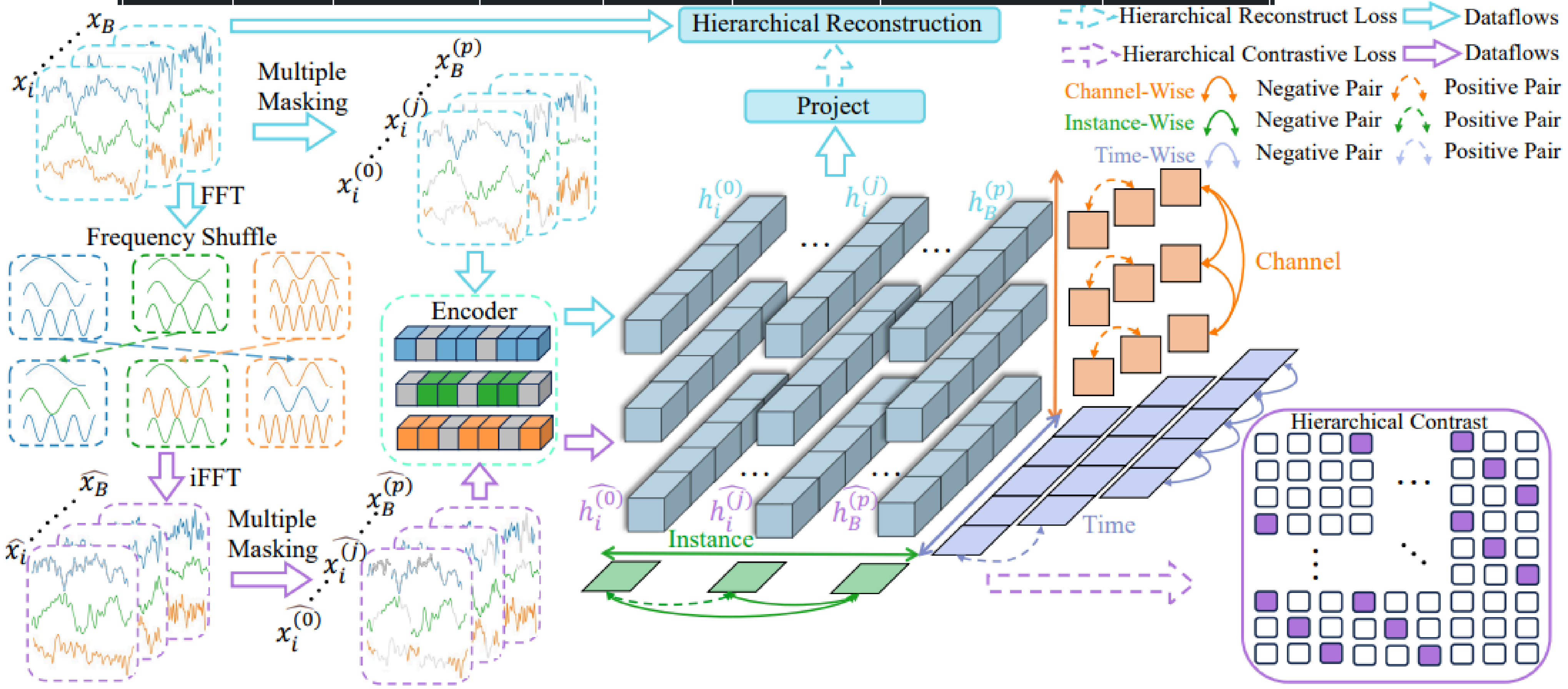
• **Oversight of Channel Relations:** Existing self-supervised methods often focus only on temporal variations or individual data points (instance level).

• **Crucial Missing Information:** These prior approaches overlook the intricate relationships and interdependencies among different data channels. These channel relationships are vital for understanding complex physiological interactions and achieving precise representation in multi-channel medical time-series data

Dataset(s) Used

Dataset	TP (Data Type)	SR (Sampling Rate)	PA (Patients)	SZ (Dataset Size)	Len (Instance Length)	CH (Channel Count/Columns)	Task (Target)
TUAB	EEG	256	2383	409.4k	10s	16	Abnormal or not
TUEV	EEG	256	390	112.4k	5s	16	IEDs or not
PTB-XL	ECG	500	18885	47.6k	5s	12	Arrhythmias or not
CHB-MIT	EEG	256	23	4.2k	10s	16	Seizure or not
Sanbo-CMR	MEG	1000	119	31.4k	5s	39	Epileptic spike or not

(Private)-> Sanbo-CMR



Reference link:-https://ieeexplore.ieee.org/document/10890624

Proposed Architecture / Methodology (CLHi-MTS)

The framework is a novel hierarchical method that combines two main approaches: masked modeling and contrastive learning.

• Hierarchical Information Learning: The model leverages feature proximity at three distinct levels:
  - **Channel-wise (interdependencies).**
  - **Time-wise (temporal fluctuations).**
  - **Instance-wise (instance-level correlations).**

• **Masked Modeling (Reconstruction)**: The model reconstructs masked segments of the data, guided by the similarities calculated across the three hierarchical levels (channel, time, instance).

• **Contrastive Learning:** Robust features are learned by ensuring representations of the same instance/channel/timestamp are pulled closer, while those from different instances/channels/timestamps are pushed apart.

• **Novel Augmentation Strategy:** A new frequency-based augmentation (Frequency Shuffling) is used to generate positive examples for contrastive learning. This technique exchanges frequency components between channels within the same sample, providing reliable augmentation and avoiding biases common in traditional techniques (like cropping or scaling).

Performance Highlights

• **Superior Performance:** CLHi-MTS consistently outperforms six state-of-the-art self-supervised methods.

• **Effectiveness Demonstrated:** The model shows superior results in both in-domain and cross-domain medical time-series classification tasks.

• **Robustness:** The high AUROC (Area Under the ROC Curve) improvement demonstrates the model's robustness on imbalanced datasets.

Limitations of the Study

• The provided sources do not explicitly state or discuss any specific weaknesses, gaps, or limitations of the proposed CLHi-MTS framework itself .

Reference link:-https://ieeexplore.ieee.org/document/10890624