CSE475: Machine Learning[Fall 2025]

Section : 02

Group Number : 08

Submitted by:

Student Name	Student ID
Tirtho Das	2023-1-60-052
Nura Alam	2023-1-60-081

Related Work

Ref.	Title	Dataset Description	Methods	Results	Pros	Cons	Future Work
1	**Study on the use of standard 12-lead ECG data for rhythm-type ECG classification problems** (Park et al.)	Large standard 12-lead ECG dataset (13,241 tests) from **Konkuk University Hospital** (used for analysis) and Seoul National University Bundang Hospital (used for verification),. Data was categorized into **5 rhythm classes** (Normal, AF, APC, VPC, and "other"). Data frequency was downsampled from 500 Hz to **200 Hz**	Proposed a methodology for **fusion of various single-lead ECG data as training data** for single-lead classification,. Used a **152-layer Squeeze-and-Excitation Residual Network (SE-ResNet)** as the baseline model. Included transformation of Lead aVR to **Lead –aVR**. Evaluation used Accuracy and F1 score.	The classification performances achieved by the 152-layer SE-ResNet using Leads −aVR and II were the best overall, associated with F1 scores of 97.5% and 97.4%, respectively . Specifically, the model trained on the fusion of single-lead ECGs achieved a high F1 score using **Lead** −aVR for the most fundamental	Methodology offers an **alternative to improve the ECG classification performance of wearable devices** that measure only single-lead ECG. Performance on critical arrhythmias (AF, APC, VPC) was **higher than previous studies**.	**F1 scores for APC were much lower** than for other diagnoses, suggesting potential **insufficient APC-scored data**. The **optimal combination of single-lead ECG data was not confirmed**.	Focus on **increasing the F1 scores for APC**. Develop methods to **find the best combination of ECG leads**, including **nonlinear transformation**. Implement training data that includes **mechanically generated noise** (e.g., power line noise, baseline wander).

				rhythm classes: • **Normal** (98.7%) • **AF** (98.2%) • **APC** (95.1%) • **VPC** (97.4%)			
2	**In-Distribution and Out-of-Distribution Self-supervised ECG Representation Learning for Arrhythmia Detection** (Soltanieh et al.)	**PTB-XL** (12-lead, 21,837 records, 100Hz), **Chapman** (12-lead, 10,646 records, downsampled to 100Hz), and **Ribeiro** (test set only, 827 samples, downsampled to 100Hz),,. Datasets confirmed suitable for cross-dataset OOD analysis.	Conducted systematic investigation of **Self-Supervised Learning (SSL)** in **In-Distribution (ID) and Out-of-Distribution (OOD)** settings. Evaluated three SSL methods: **SimCLR**, **BYOL**, and **SwAV**, using **xResNet1d50** as the backbone. Used diverse augmentations (Gaussian Noise, Masking, Time Warping, etc.)	**SwAV consistently achieved the best overall results** across all explored datasets in ID and OOD settings. SSL methods achieved **comparable performances for ID and OOD data**, indicating learned representations generalize well. The SSL model (SwAV) showed **greater robustness against synthetic 60 Hz power line interference** than the	SSL methods **outperform fully supervised training**. SSL effectively enhances model **generalization to OOD scenarios**. The models, particularly SwAV, are robust against common ECG noise.	Features corresponding to some diseases (HYP and STTC) proved **more challenging to learn** compared to others.	Investigate **finer step sizes and wider parameter ranges** for augmentations. Examine **more sophisticated augmentations** (e.g., in the frequency domain). Expand the number of **SSL methods and backbone encoders**. Collect a **combination of different datasets for pre-training** for improved generalizability.

				fully-supervised model			
3	**Subject-based Non-contrastive Self-Supervised Learning for ECG Signal Processing (Atienza et al.)**	Training Dataset: **The Sleep Heart Health Study (SHHS)** (2,643 subjects with two records each),. Evaluation Datasets: **MIT-BIH Arrhythmia Database (MIT Arr), MIT-BIH Atrial Fibrillation Database (MIT AFib), and MIT-BIH Polysomnographic Database (MIT PSG)** (all publicly available on PhysioNet)	Proposed **SBnCL** (Subject-based Non-contrastive Self-Supervised Learning) method that operates **without data augmentation or negative pairs**. Compared performance using a **Vision Transformer (ViT)-based model** against a **CNN-based XResNet50**,. Evaluation included gender classification and downstream tasks like age regression and sleep stage classification	The ViT-SBnCL model achieved comparable performance to existing methods in extracting **stable, subject-based static features** such as gender (**95.7%** accuracy),. The **ViT model showed a higher learning capability** than the CNN-based XResNet50	Demonstrated that **neither data augmentations nor negative pairs are necessary** for learning useful representations. The proposed method allows for the design of SSL models capable of capturing **dissimilarities** (instead of just similarities).	The model **failed to capture dynamic characteristics** such as **arrhythmias (AFib) or sleep stages**. The training dataset (SHHS) contained a **smaller number of subjects** (2,643 subjects) compared to other studies (e.g., PCLR study).	Design a **new SSL method that captures dissimilarities as well as similarities** to improve representation learning for **transient states** and different rhythm diseases
4	**ECG**	Multiclass	Proposed the	Overall	Achieved	SqueezeNet	Improve the

	Classification for Detecting ECG Arrhythmia Empowered with Deep Learning Approaches (Rahman et al.)	ECG classification dataset taken from **Kaggle** (16,879 images), which was **augmented and fused with the MIT-BIH dataset**,. The model classified **five classes**: F, N, Q, S, and V. Images were preprocessed to dimensions of 227×227.	**CAA-TL** (ECG Classification for Detecting ECG Arrhythmia Empowered with Transfer Learning) model. Compared and used transfer learning on three deep learning models: **AlexNet** (25 layers), **SqueezeNet** (68 layers), and **ResNet50** (177 layers),. The train/validation split was 80:20.	accuracy achieved was: **AlexNet 98.38%**, SqueezeNet 90.08%, and ResNet50 91%,. **AlexNet** was found to be the **most effective method** as it could train and validate **all** the classes. SqueezeNet and ResNet50 showed outstanding accuracy for classes Q, S, and V but **failed to train the N (Normal) class** effectively,,.	**remarkable accuracy** compared to previous machine learning methods. **Augmentation** and image processing were applied to increase the number of images.	and ResNet50 showed **poor performance or failure to detect specific classes** (N and F),. The computational complexity is significant.	dataset quality for classes showing less accuracy (**N and F** classes),. Improve computational speed using **GPU or AWS cloud computing service**. Apply the innovative approach of **federated deep learning** to enhance security and consistency. Utilize **K fold Cross-validation**.
5	**Deep Learning for ECG Analysis: Benchmarks and Insights from PTB-XL** (Strodthoff et al.)	Primary Dataset: **PTB-XL** (21,837 12-lead ECG records from 18,885 patients). Annotations cover three	Proposed and evaluated a new **resnet-adaptation (xresnet1d101)** along with other CNN (ResNet and	**xresnet1d101** showed the **strongest performance** across all tasks (macro AUC up to **0.96** in *rhythm* category),.	Established **comprehensive benchmarking results** for deep learning algorithms on the PTB-XL dataset. Demonstrated	Recurrent architectures were **less performant** than convolutional counterparts. The performance of	Conduct further studies on **co-occurring pathologies** to identify hidden stratification. Analyze the correlation of **human-provid**

		non-mutually exclusive categories: *diagnostic*, *form*, and *rhythm*. Secondary dataset for transfer learning validation: **ICBEB2018** (6877 12-lead ECGs).	Inception variants), Recurrent (LSTM, GRU), and feature-based (Wavelet+NN) architectures,,. Tasks included ECG statement prediction, age regression, and sex prediction. Used **transfer learning** by fine-tuning models pretrained on PTB-XL onto ICBEB2018. Primary metric: macro-averaged AUC.	**Transfer learning was highly effective in the small dataset regime**. Age regression achieved MAE = **6.86 years** (on normal subjects). Sex prediction achieved AUC of **0.96** (on normal patients).	the **prospects of transfer learning** using PTB-XL as a resource, reducing the amount of target data required. Provided evidence of **hidden stratification** and successful comparison of **model uncertainty with human-provided diagnosis likelihoods**,.	the feature-based approach (**Wavelet+NN**) was significantly lower than the deep learning models.	**ed diagnosis likelihoods with model uncertainties**. Apply **interpretability methods** (e.g., relevance propagation) for statistical analysis of attribution maps,.
6	**SELF-SUPERVISED PRE-TRAINING with Joint-Embedding Predictive Architecture Boosts ECG Classification Performance** (Weimann & Conrad, 2023)	Unsupervised pre-training dataset combining **ten public ECG databases** (including MIMIC-IV-ECG, CODE-15, PTB-XL, etc.) totaling **over one million**	Proposed the use of the **Joint-Embedding Predictive Architecture (JEPA)** for self-supervised pre-training, which predicts latent features,. The model architectures	JEPA achieved **state-of-the-art performance** on the PTB-XL *all statements* task with an AUC of **0.945** (ViT-S, two-stage fine-tuning), surpassing the previous	JEPA overcomes limitations of traditional SSL methods by **predicting latent features** and **not relying on hand-crafted data augmentations**	Increasing model size (ViT-B) or expanding dataset diversity beyond MIMIC-IV-ECG provided **negligible impact** on performance,	Study **different techniques for unsupervised or self-supervised pretraining** that do not rely on expensive human annotations.

		ECG records. Evaluation performed on various **PTB-XL benchmarks**.	were **Vision Transformers (ViT)** in three configurations: **ViT-B, ViT-S, and ViT-XS**. Implemented a **masking strategy** to create compatible pairs. Evaluation included linear evaluation, fine-tuning, and two-stage fine-tuning.	benchmark (CPC, 0.942 AUC),,. **JEPA consistently learned the highest quality representations**, confirmed by linear evaluation (ViT-S: **0.940 AUC**),. Pre-training was **always advantageous**, even when using no additional data beyond PTB-XL.	. Achieved **state-of-the-art performance** across multiple PTB-XL benchmarks,. Demonstrated strong feature quality in **linear evaluation**.	suggesting larger models may require even more data to utilize full capacity.	
7	**ECG classification using 1-D convolutional deep residual neural network** (Khan et al.)	**PhysioNet MIT-BIH Arrhythmia database** (48 two-channel recordings, 360 Hz),. Heartbeats are grouped into **five AAMI Standard classes**: N, S, V, F, and Q. The dataset is **highly**	Proposed a **1-D convolutional deep Residual Neural Network (ResNet)** for feature extraction. Employed the **Pan-Tompkins algorithm** for R-peak detection and heartbeat segmentation,.	Achieved an **average accuracy of 98.63%** and an average F1-score of **92.63%** on the test dataset,. The deeper ResNet model performed significantly well with **increased network**	The model performed feature extraction directly from heartbeats, avoiding separate, complex signal processing phases. **SMOTE effectively addressed the severe**		

		class-imbalanced (82.8% of beats belong to the N class).	Utilized the **Synthetic Minority Oversampling Technique (SMOTE) exclusively on the training dataset** to mitigate class imbalance,,. Evaluation used 10-fold cross validation.	**depth**. Accuracy was significantly higher with SMOTE (98.62%) than without (95%).	**class-imbalance problem**.		
8	**ECG signal classification based on deep CNN and BiLSTM** (Cheng et al.)	Dataset provided by the **2017 PhysioNet/CINC Challenge** (8,528 single-lead records),. Data was screened down to 7,561 records with a minimum dimension of 9,000,. Divided into four types: **AF, Normal, Other rhythm, and Noise**.	Developed a dense heart rhythm network combining a **24-layer Deep Convolutional Neural Network (DCNN)** and **Bidirectional Long Short-Term Memory (BiLSTM)**,. Preprocessing used a **combined filter of wavelet**	Obtained an overall **accuracy rate of 89.3%** and an **F1 score of 0.891** through ten-fold cross validation,. The **TMSE loss function effectively suppressed outliers** and achieved better accuracy compared to cross-entropy and MSE methods.	The model successfully mined deep-level and **time-sensitive features** using the combined DCNN and BiLSTM network. The combined filtering method (**WT and MT**) effectively reduced noise and protected signal features.	The filtering effect **declined for data collected by ECG acquisition instruments from different manufacturers**. The network layers were relatively shallow due to **hardware limitations**.	Study **more robust preprocessing methods** to handle different manufacturers' data. Explore using **deeper layers** for the DCNN and BiLSTM networks. **Transplant the classification algorithm to portable terminal devices** for dynamic ECG monitoring.

			transform (WT) and median filtering (MT) for noise reduction and baseline drift elimination,,. Proposed a new loss function: **Tan Mean Square Error (TMSE)**, to control loss fluctuation and suppress outliers.				
9	**Transfer learning for ECG classification** (Weimann & Conrad, 2021)	Upstream (Pre-training) dataset: **Icentia11K** (single lead, 250 Hz, 11,000 patients, over 630,000 hours of ECG). Downstream (Finetuning) dataset: **PhysioNet/CinC Challenge 2017 data set** (8528 single-lead episodes). Also	Used **transfer learning** to pretrain CNNs (primarily **ResNet-18v2**, **ResNet-34v2**, and **ResNet-50v2**) and fine-tune for AF classification. Pretraining tasks included supervised (Beat/Rhythm Classification) and self-supervised	Pretraining improved the macro F1 score on the target task by up to **6.57%** (Beat Classification pretraining) over random initialization. The pretraining improvement was substantial in the **small dataset regime**. **ResNet-34v2** yielded the	Transfer learning **improved performance and accelerated training**. The approach **effectively reduces the number of annotations required** to achieve comparable performance. Unsupervised pretraining	Supervised pretraining was generally the **better choice** if labels were available. The largest model (**ResNet-50v2**) suffered a performance decline, suggesting potential resource or pretraining schedule limitations.	Collect a **massive ECG database** that reflects data heterogeneity to pretrain large networks. Study **different techniques for unsupervised or self-supervised pretraining** since they avoid expensive human annotations.

		tested generalization on PTB-XL and ICBEB2018 (12-lead datasets).	(**Future Prediction**, an adaptation of **Contrastive Predictive Coding [CPC]** using a **Transformer-based attention pooling module**), The primary metric was the macro **F1 score**.	highest performance improvement (4.51% increase over baseline),.	methods are a **viable method** for improving performance.		

Reference:

1. Park, Junsang, Junho An, Jinkook Kim, Sunghoon Jung, Yeongjoon Gil, Yoojin Jang, Kwanglo Lee, and Il-young Oh. "Study on the use of standard 12-lead ECG data for rhythm-type ECG classification problems." *Computer Methods and Programs in Biomedicine* 214 (2022): 106521.
2. Soltanieh S, Hashemi J, Etemad A. In-distribution and out-of-distribution self-supervised ecg representation learning for arrhythmia detection. IEEE Journal of Biomedical and Health Informatics. 2023 Nov 10;28(2):789-800.
3. Atienza, A., Bardram, J. and Puthusserypady, S., 2023. Subject-based Non-contrastive Self-Supervised Learning for ECG Signal Processing. *arXiv preprint arXiv:2305.10347*.
4. Rahman, Atta-Ur, Rizwana Naz Asif, Kiran Sultan, Suleiman Ali Alsaif, Sagheer Abbas, Muhammad Adnan Khan, and Amir Mosavi. "ECG classification for detecting ECG arrhythmia empowered with deep learning approaches." *Computational intelligence and neuroscience* 2022, no. 1 (2022): 6852845.
5. Gitau, Antony M., Ayoade Adeyemi, Belal Tavashi, and Bjørn-Jostein Singstad. "[Re] Deep Learning for ECG Analysis: Benchmarks and Insights from PTB-XL." *medRxiv* (2025): 2025-01.
6. Weimann, K. and Conrad, T.O., 2024. Self-Supervised Pre-Training with Joint-Embedding Predictive Architecture Boosts ECG Classification Performance. *arXiv preprint arXiv:2410.13867*.

7. Weimann, K. and Conrad, T.O., 2024. Self-Supervised Pre-Training with Joint-Embedding Predictive Architecture Boosts ECG Classification Performance. *arXiv preprint arXiv:2410.13867*.
8. Cheng, J., Zou, Q. and Zhao, Y., 2021. ECG signal classification based on deep CNN and BiLSTM. *BMC medical informatics and decision making*, *21*(1), p.365.
9. Weimann, K. and Conrad, T.O., 2021. Transfer learning for ECG classification. *Scientific reports*, *11*(1), p.5251.