2025. 01. 03.

Gyuyeon Lim (lky473736)

Testing 2 : result

MEAN AND (STANDARD DEVIATION) OF ACCURACY (%) OVER TEN EXPERIMENTS

	UCI-HAR	mHealth	PAMAP2	MobiAct
CNN [27]	92.34 (0.44)	72.53 (1.10)	64.25 (2.47)	84.40 (1.60)
DeepConvLSTM [27]	89.46 (1.95)	72.22 (1.99)	61.31 (2.85)	85.31 (0.85)
LSTM-CNN [28]	92.64 (0.28)	70.43 (1.63)	62.03 (1.78)	80.19 (1.56)
CAE [33]	87.66 (1.47)	**76.09 (0.90)**	68.22 (1.26)	82.85 (1.26)
Masked recon. [34]	69.34 (2.71)	72.15 (1.10)	62.77 (1.81)	79.24 (1.07)
CPC [38]	85.09 (1.27)	75.66 (1.25)	64.38 (1.43)	83.89 (0.51)
One-stream model	**93.32 (0.45)**	71.84 (1.16)	65.31 (2.04)	84.47 (0.91)
Two-stream model	91.37 (0.43)	73.74 (0.73)	66.68 (1.35)	89.21 (0.76)
CAGE (Ours)	92.50 (0.44)	75.48 (2.11)	**70.47 (2.85)**	**89.70 (0.63)**

(Above) the figure 3 at the paper “<Contrastive Accelerometer–Gyroscope Embedding Model for Human Activity Recognition (https://ieeexplore.ieee.org/document/9961198)>"

best validation accuracy(F1)	UCI-HAR	mHealth	PAMAP2
CAGE	86.93 (0.8680)	84.08 (0.8301)	76.69 (0.7662)
BaselineCNN	82.67 (0.8282)	83.7 (0.8367)	75.19 (0.7478)
CAE (Convolution AutoEncoder)	46.17 (0.4610)	.	.
DeepConvLSTM	77.73 (0.7765)	.	.
LSTM-CNN	89.81 (0.8971)	86.35 (0.8569)	74.71 (0.7442)

best test accuracy(F1)	UCI-HAR	mHealth	PAMAP2
CAGE	17.61 (0.0926)	0.91 (0.0002)	0.07 (0.0000)
BaselineCNN	84.36 (0.8434)	74.58 (0.6891)	62.49 (0.6117)
CAE (Convolution AutoEncoder)	17.10 (0.0878)	.	.
DeepConvLSTM	79.50 (0.7956)	.	.
LSTM-CNN	89.38 (0.8929)	74.05 (0.6742)	66.24 (0.6421)

final test accuracy(F1)	UCI-HAR	mHealth	PAMAP2
CAGE	85.00 (0.8499)	77.05 (0.7325)	59.66 (0.5968)
BaselineCNN	84.53 (0.8450)	71.84 (0.6492)	66.17 (0.6524)
CAE (Convolution AutoEncoder)	46.93 (0.4656)	.	.
DeepConvLSTM	79.13 (0.7928)	.	.
LSTM-CNN	79.13 (0.7928)	74.97 (0.6851)	64.40 (0.6199)

final validation accuracy(F1)	UCI-HAR	mHealth	PAMAP2
CAGE	84.85 (0.8473)	82.27 (0.8139)	74.30 (0.7439)
BaselineCNN	82.42 (0.8235)	84.08 (0.8220)	69.51 (0.6947)
CAE (Convolution AutoEncoder)	45.36 (0.4510)	.	.
DeepConvLSTM	75.92 (0.7594)	.	.
LSTM-CNN	88.73 (0.8858)	83.60 (0.8119)	70.68 (0.7050)

- I have contacted BMILab (Biomedical Informatics Laboratory) via email to request access to the MobiAct dataset for review and download. However, I have not yet received a response, which has prevented me from conducting experiments.
- Due to hardware limitations, I was unable to complete some experiments with CAE and DeepConvLSTM (as training all epochs requires over 12 hours to complete).
- The results of CAGE were not universally superior across all scenarios.

Overall, the scores of CAGE were not satisfactory. Could there be an issue with the experimental methodology? The experiments were conducted using the following commands.

CAGE UCI-HAR	`python train_CAGE.py --dataset UCI_HAR --batch_size 64 --epochs 200 --proj_dim 64 --lambda_ssl 1.0 --lambda_cls 1.0`
CAGE mHealth	`python train_CAGE.py --dataset mHealth --batch_size 64 --epochs 200 --proj_dim 64 --lambda_ssl 1.0 --lambda_cls 1.0`
CAGE PAMAP2	`python train_CAGE.py --dataset PAMAP2 --batch_size 64 --epochs 200 --proj_dim 128 --lambda_ssl 1.0 --lambda_cls 1.0`

DeepConv LSTM	`python train_baseline.py --dataset UCI_HAR --model DeepConvLSTM`
	`python train_baseline.py --dataset mHealth --model DeepConvLSTM`
	`python train_baseline.py --dataset PAMAP2 --model DeepConvLSTM`

LSTM ConvNet	`python train_baseline.py --dataset UCI_HAR --model LSTMConvNet`
	`python train_baseline.py --dataset mHealth --model LSTMConvNet`
	`python train_baseline.py --dataset PAMAP2 --model LSTMConvNet`

CAE	`python train_CAE.py --dataset UCI_HAR`
	`python train_CAE.py --dataset mHealth`
	`python train_CAE.py --dataset PAMAP2`

Baseline CNN	`python3 train_baseline.py --dataset UCI_HAR --model BaselineCNN`
	`python3 train_baseline.py --dataset mHealth --model BaselineCNN`
	`python3 train_baseline.py --dataset PAMAP2 --model BaselineCNN`

period