Supplementary Material for Time Series Averaging Using Multi-Tasking Autoencoder

Tsegamlak Terefe, Maxime Devanne, Jonathan Weber, Dereje Hailemariam, Germain Forestier
Emails:{tsegamlak-terefe.debella,maxime.devanne,jonathan.weber,germain.forestier}@uha.fr
{tsegamlak.terefe,dereje.hailemariam}@aait.edu.et

I. Experimental Results

This supplementary material is provided to present the experimental results for the paper *"Time Series Averaging Using Multi-Tasking Autoencoder"*. In the paper, emphasis was given to the multitasking autoencoder shown in Figure 1. The network was trained to obtain a compact per class latent space features. The features were then utilized to mimic multiple alignment which is a challenging and computationally intensive task in time series averaging [1]–[4]. To this end, we have first trained the proposed network and utilized the encoder portion of the network to project the time series into a latent space. The latent features were then utilized to make the estimation of class averages. In the experimental setups we have utilized data sets obtained from the University of California Univariate Time Series Repository (UCR) [5]. The data sets are provided in two separate ".tsv" files; i.e., one for training and one for testing. In our experimental setups, we have sliced the training file into a 20/80 format, i.e., 80% for training and 20% for validation. Moreover, to make the estimation of the per class averages, we have taken the per class arithmetic means of the latent features of the train data sets; which were also latter projected to the time domain using the decoder. In practice, the time domain arithmetic mean of time series data sets is a sub optimal estimation of the mean in most cases [3]. This is better demonstrated in Figure 2.

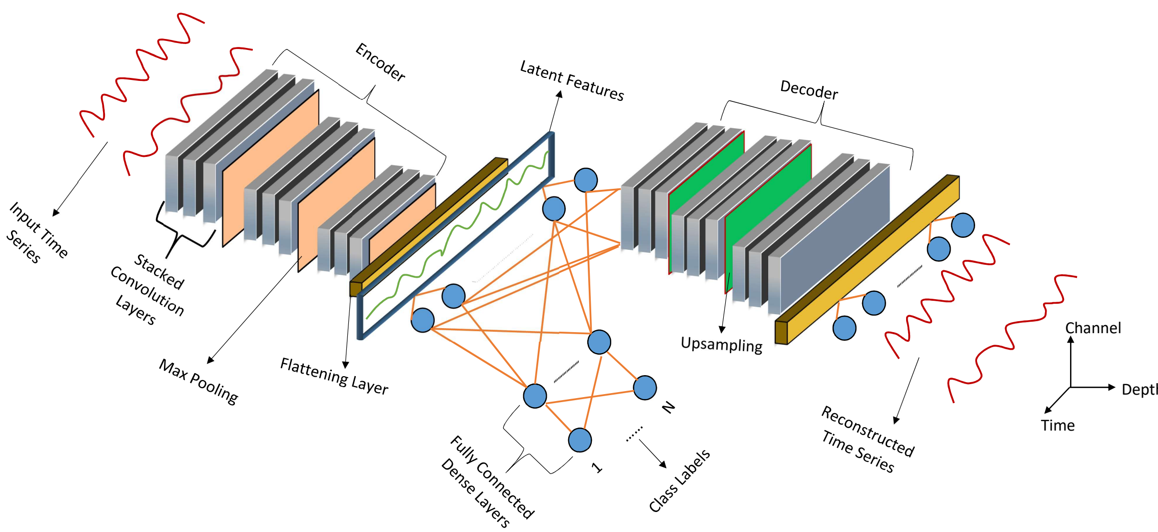

Fig. 1. Proposed multi-tasking autoencoder.

In general, in the paper, two major experiments were conducted. In the first experimental setup, a non multitasking auto encoder was utilized to obtain the latent space representation of the uni variate time series [5]. These experiments were conducted to evaluate the performance of a basic convolutional autoencoder on obtaining a separable and compact per class latent features. To this end, we first removed the classifier portion of the multitasking autoencoder and trained the basic convolutional autoencoder with reconstruction loss, (1), as an objective function.

$$L_{reconstruction} = \frac{1}{N}\sum_{i=0}^{N}(r_i - x_i)^2. \tag{1}$$

where, R is the reconstructed time series and X is the input time series.

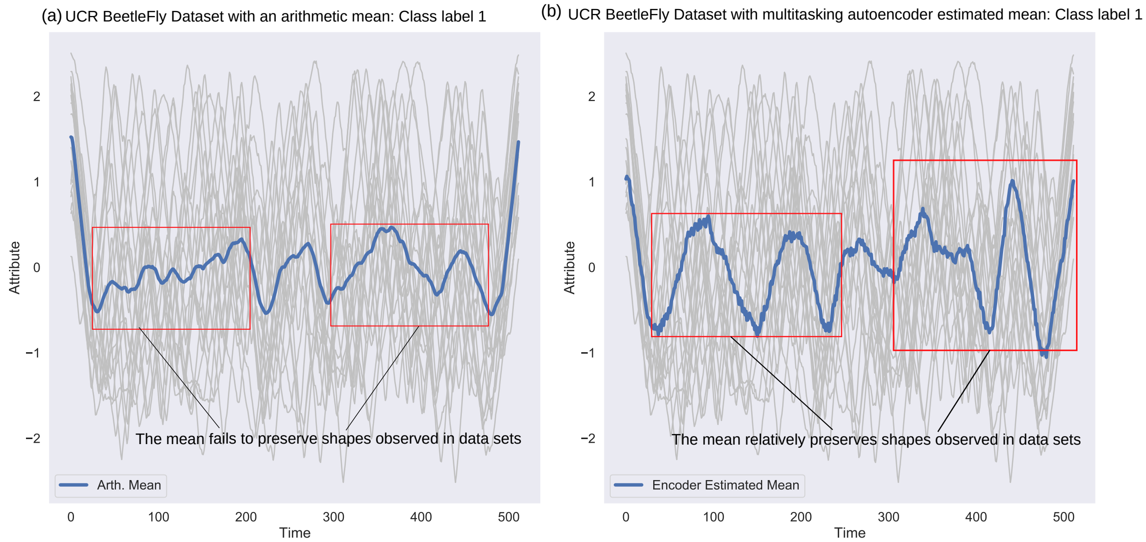

Fig. 2. Comparison of estimated means.(a) Arithmetic mean, (b) Multitasking autoencoder estimated mean

We have then utilized the encoder portion of the trained network to project the train data sets into a latent space. Following this, the per class arithmetic mean of the latent features were taken as the estimate of the class means. Following this, to measure the quality of the estimated latent means, a one nearest centroid classification was conducted using the per class latent means and the latent features of the test data set. To conduct the classification, we have utilized Euclidean distance as a means of similarity measure. The outcomes of these classifications are reported as ”Enc. Lat” in Table II & III. Moreover, the latent means were also projected to the time domain using the decoder portion of the autoencoder to conduct a one nearest centroid classification in the time domain. However, we have utilized Dynamic Time Warping (DTW) distance as a similarity measure to account for temporal distortion in the time domain. The outcomes of these experiments are reported as ”Enc. Time” in Table II & III. On the other hand, the multitasking autoencoder was trained with categorical cross entropy and reconstruction loss as an objective function. i.e.,

$$L_{multi}(x, r, h, P) = \frac{1}{N}\sum_{i=0}^{N}(r_i - x_i)^2 - \sum_{c=0}^{C} h_{o,c}\log p_{o,c}, \tag{2}$$

where r and x are the reconstructed and input time series, and h and p are the categorical representation and the softmax activation values for each C catagories, respectively.

After training, the encoder portion of the network was utilized to project the data sets to the latent space. The projected latent features were then utilized to preform one nearest centroid classification as conducted in the basic autoencoder. The latent space and the time domain results are respectively reported as ”MT. Enc. Lat” and ”Mt. Enc. Time” in Table II, III & IV.
In the experimental setups, the basic autoencoder, i.e., an autoencoder performing encoding and decoding, was trained for 2500 epochs and zero L2 regularization. However, we have trained the multitasking autoencoder for five different L2 regularization setups as shown in Table I. For the first L2 regularization setup, the network was trained for 600 epochs. On the contrary, for the remaining four setups the network was trained for 2500 epochs. Moreover, for the multitasking autoencoder, we have reported the best latent space classification result and its time domain counterpart. Additionally, we have extracted results for other averaging techniques from [6]. Even if [6] provided experimental results for 84 data sets, we have conducted the two major tests on 85 data sets. Moreover, we have provided experimental results for additional 14 data sets using the multitasking autoencoder.

TABLE I
L2 REGULARIZATION SCHEMES USED WHILE TRAINING THE MULTI TASKING AUTOENCODER.

No.	Layer	L2 Regularizations
1.	Encoder	$0, 10^{-4}, 10^{-3}, 10^{-3}, 10^{-2}$
2.	Decoder	$0, 10^{-4}, 10^{-3}, 10^{-3}, 10^{-2}$
3.	Classifier	$0, 10^{-3}, 10^{-3}, 10^{-2}, 10^{-2}$

TABLE II

COMPARISON OF ONE NEAREST CENTROID CLASSIFICATION WITH DIFFERENT AVERAGING TECHNIQUES. THE RESULTS FOR DTAN, DBA AND SOFT DBA ARE EXTRACTED FROM [6]. ACCORDING TO [6] THE DTAN NETWORK WAS TRAINED USING A 80/20 SPLIT; WHERE, THE NETWORK WAS TRAINED FOR 2500 EPOCHS WITH DIFFERENT LSTM ITERATION AND REPORTED THE BEST ACCURACY.

No.	Data	Enc. Time	Enc. Lat	MT. Enc.Time	MT. Enc.Lat	DTAN	DBA	SDBA	Arth. Mean
1.	Adiac	47.06	31.71	39.90	55.75	**69.57**	46.29	50.13	47.06
2.	ArrowHead	48.57	53.71	64.00	**79.43**	74.86	52.00	52.00	46.86
3.	Beef	40.00	63.33	43.33	**66.67**	63.33	40.00	56.67	46.67
4.	BeetleFly	50.00	60.00	50.00	80.00	80.00	**90.00**	85.00	70.00
5.	BirdChicken	45.00	45.00	65.00	**85.00**	80.00	60.00	70.00	55.00
6.	Car	40.00	55.00	51.67	76.67	**81.67**	63.33	68.33	31.67
7.	CBF	72.22	87.33	89.67	**97.44**	91.44	96.56	97.11	61.56
8.	ChlorineConcentration	32.66	32.76	30.86	**34.90**	33.31	32.37	34.82	31.69
9.	CinCECGTorso	31.28	75.90	24.71	**97.25**	61.59	44.57	39.86	15.80
10.	Coffee	96.43	92.86	**100**	**100.00**	**100.00**	96.43	96.43	96.43
11.	Computers	50.00	52.00	54.80	58.80	59.20	61.60	**64.00**	51.20
12.	CricketX	21.54	32.05	30.77	43.08	42.31	57.44	**60.26**	19.23
13.	CricketY	19.49	42.05	22.31	48.97	54.10	54.10	**57.18**	20.77
14.	CricketZ	17.18	31.28	26.15	45.38	42.05	60.51	**61.54**	19.74
15.	DiatomSizeReduction	76.14	92.81	90.86	95.10	**97.06**	95.10	95.10	84.64
16.	DistalPhalanxOutlineAgeGroup	72.66	70.50	74.10	74.82	84.75	84.00	**85.00**	73.38
17.	DistalPhalanxOutlineCorrect	41.67	**78.26**	43.12	76.09	47.17	48.83	49.00	64.49
18.	DistalPhalanxTW	62.59	53.24	63.31	63.31	**78.00**	76.00	75.50	61.87
19.	Earthquakes	25.90	56.83	74.82	64.03	77.33	57.45	**82.30**	25.18
20.	ECG200	61.00	74.00	72.00	**84.00**	79.00	72.00	73.00	67.00
21.	ECG5000	80.96	89.64	80.29	**90.51**	89.13	83.47	85.38	83.82
22.	ECGFiveDays	49.94	60.86	58.65	96.17	**97.79**	65.85	67.02	52.96
23.	ElectricDevices	24.67	44.99	16.88	**63.70**	53.48	53.90	53.97	17.27
24.	FaceAll	36.21	32.96	41.54	70.59	**82.78**	79.65	80.47	41.30
25.	FaceFour	65.91	78.41	69.32	**95.45**	82.95	85.23	85.23	76.14
26.	FacesUCR	46.68	58.68	46.78	82.49	**85.71**	77.46	81.27	40.78
27.	FiftyWords	24.62	51.65	29.45	60.44	**65.27**	61.54	61.54	16.70
28.	Fish	41.71	56.00	49.14	76.00	**90.29**	65.14	69.71	40.00
29.	FordA	51.59	58.64	51.59	**91.97**	60.48	54.96	55.29	51.59
30.	FordB	49.51	51.73	49.51	**63.83**	57.98	56.85	59.13	53.21
31.	GunPoint	53.33	67.33	50.00	**97.33**	88.00	70.00	73.33	54.67
32.	Ham	60.95	67.62	60.00	75.24	**79.05**	72.38	73.33	62.86
33.	HandOutlines	70.00	82.70	71.89	**90.81**	85.0	80.40	81.20	67.57
34.	Haptics	24.35	33.44	27.60	42.53	**45.78**	35.06	37.34	27.27
35.	Herring	57.81	56.25	57.81	64.06	**70.31**	54.69	60.94	50.00
36.	InlineSkate	18.98	19.66	20.18	22.91	**26.00**	23.27	25.27	16.73
37.	InsectWingbeatSound	20.25	55.20	18.03	53.74	**58.74**	28.94	32.83	17.53
38.	ItalyPowerDemand	80.47	93.59	84.26	**96.60**	96.21	73.08	75.02	76.77
39.	LargeKitchenAppliances	34.13	41.60	37.87	47.47	48.27	72.80	**73.33**	34.40
40.	Lightning2	59.02	67.21	60.66	**75.41**	72.13	63.93	62.30	57.38
41.	Lightning7	47.95	43.84	52.05	61.64	71.23	**72.60**	69.86	31.51
42.	Mallat	30.06	96.29	93.86	96.67	**96.89**	95.27	95.39	93.94

TABLE III
CONTINUED COMPARISON OF ONE NEAREST CENTROID CLASSIFICATION RESULTS

No.	Data	Enc. Time	Enc. Lat	MT. Enc.Time	MT. Enc.Lat	DTAN	DBA	SDBA	Arth. Mean
43.	Meat	**95.00**	88.33	91.67	93.33	93.33	91.67	93.33	91.67
44.	MedicalImages	31.71	35.79	29.74	46.58	**46.84**	43.68	46.18	29.47
45.	MiddlePhalanxOutlineAgeGroup	59.74	57.14	58.44	57.79	73.75	71.25	**79.50**	59.74
46.	MiddlePhalanxOutlineCorrect	59.45	**79.38**	61.86	65.29	54.33	49.50	48.33	57.73
47.	MiddlePhalanxTW	53.90	42.86	42.86	42.86	**59.65**	55.64	58.15	48.05
48.	MoteStrain	84.42	82.43	82.59	86.42	**90.42**	82.67	84.35	80.99
49.	NonInvasiveFetalECGThorax1	53.28	57.51	65.24	79.85	**85.34**	71.30	71.09	64.73
50.	NonInvasiveFetalECGThorax2	62.54	71.50	69.82	82.24	**90.53**	76.39	77.30	75.67
51.	OliveOil	46.67	70.00	76.67	76.67	**86.67**	76.67	80.00	76.67
52.	OSULeaf	25.62	35.54	23.14	40.08	46.28	43.80	**47.52**	20.25
53.	PhalangesOutlinesCorrect	61.19	62.00	60.96	**80.65**	64.22	63.29	63.75	61.77
54.	Phoneme	1.05	8.28	2.43	14.87	10.18	18.25	**20.46**	3.27
55.	Plane	99.05	96.19	99.05	96.19	**100.00**	**100.00**	99.05	93.33
56.	ProximalPhalanxOutlineAgeGroup	80.98	79.02	83.46	81.43	**85.37**	84.39	**85.37**	81.95
57.	ProximalPhalanxOutlineCorrect	63.92	63.57	88.66	**90.03**	64.26	64.95	72.51	63.92
58.	ProximalPhalanxTW	65.85	67.80	67.32	73.66	**81.75**	73.50	74.75	64.88
59.	RefrigerationDevices	26.40	43.47	36.00	52.00	46.67	58.40	**58.67**	34.13
60.	ScreenType	33.87	40.00	**49.07**	45.87	44.53	37.87	38.93	33.07
61.	ShapeletSim	50.00	67.22	50.00	**67.78**	53.89	52.22	58.89	56.11
62.	ShapesAll	41.67	52.83	42.83	56.33	**62.83**	60.33	**62.83**	40.33
63.	SmallKitchenAppliances	33.07	60.53	32.53	**73.07**	62.13	66.13	65.87	7.47
64.	SonyAIBORobotSurface1	79.20	84.69	71.05	88.69	**89.35**	83.53	89.35	**89.52**
65.	SonyAIBORobotSurface2	70.51	77.65	73.56	**87.51**	81.11	76.60	77.23	75.34
66.	StarlightCurves	79.92	76.13	80.99	89.87	NA	14.29	14.29	75.42
67.	Strawberry	57.84	56.49	72.43	74.05	**84.34**	61.66	64.93	58.38
68.	SwedishLeaf	51.20	71.36	52.00	**83.04**	80.64	68.16	72.32	41.12
69.	Symbols	83.22	86.33	84.52	78.59	89.75	85.73	**95.48**	91.46
70.	SyntheticControl	72.67	88.00	83.67	**99.00**	95.00	98.00	98.00	73.67
71.	ToeSegmentation1	48.68	61.84	50.44	**69.30**	64.04	67.11	61.40	48.25
72.	ToeSegmentation2	82.31	55.38	63.08	76.15	75.38	83.85	**85.38**	68.46
73.	Trace	78.00	64.00	76.00	**99.00**	78.00	97.00	97.00	57.00
74.	TwoLeadECG	58.65	54.26	81.04	93.06	**95.61**	81.12	80.16	70.32
75.	TwoPatterns	14.22	50.15	47.00	91.80	55.58	97.50	**98.98**	32.60
76.	UWaveGestureLibraryAll	54.86	85.17	59.99	90.34	**92.07**	83.19	83.36	48.91
77.	UWaveGestureLibraryX	56.98	65.21	55.42	70.07	68.12	67.64	**70.69**	56.53
78.	UWaveGestureLibraryY	50.64	58.99	49.53	59.58	**61.17**	52.54	56.48	49.94
79.	UWaveGestureLibraryZ	47.88	55.05	62.65	48.91	**64.21**	59.24	60.41	47.15
80.	Wafer	40.06	77.22	62.64	98.72	**98.90**	51.10	64.94	41.79
81.	Wine	50.00	48.15	**59.26**	50.00	57.41	51.85	57.41	51.85
82.	WordSynonyms	10.50	31.66	10.03	37.46	**47.49**	34.48	41.22	8.31
83.	Worms	18.18	50.65	27.27	**50.65**	25.97	41.44	40.88	20.78
84.	WormsTwoClass	50.65	53.25	48.05	63.64	61.88	59.12	**65.19**	49.35
85.	Yoga	46.40	51.33	47.03	**70.57**	63.17	55.70	57.40	46.43

TABLE IV
EVALUATION OF REMAINING DATA SETS

No.	Name	MT. Enc. Time	MT. Enc. Lat
86.	UMD	55.56	78.47
87.	InsectEPGRegularTrain	100.00	100.00
88.	FreezerSmallTrain	73.02	78.32
89.	FreezerRegularTrain	76.28	95.72
90.	GunPointOldVersusYoung	97.14	88.89
91.	InsectEPGSmallTrain	100.00	100.00
92.	Adiac	43.73	56.78
93.	BME	62.00	78.67
94.	PowerCons	87.78	98.33
95.	GunPointAgeSpan	82.28	87.03
96.	GunPointMaleVersusFemale	72.47	67.41

References

[1] D. Schultz and B. Jain, "Nonsmooth analysis and subgradient methods for averaging in dynamic time warping spaces," *Pattern Recognition*, vol. 74, pp. 340–358, 2018.

[2] V. Niennattrakul and C. A. Ratanamahatana, "Shape averaging under time warping," in *2009 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, vol. 2, pp. 626–629, IEEE, 2009.

[3] F. Petitjean and P. Gançarski, "Summarizing a set of time series by averaging: From steiner sequence to compact multiple alignment," *Theoretical Computer Science*, vol. 414, no. 1, pp. 76–91, 2012.

[4] R. A. Shapira Weber, M. Eyal, N. Skafte, O. Shriki, and O. Freifeld, "Diffeomorphic temporal alignment nets," in *Advances in Neural Information Processing Systems 32* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, eds.), pp. 6574–6585, Curran Associates, Inc., 2019.

[5] Y. Chen, E. Keogh, B. Hu, N. Begum, A. Bagnall, A. Mueen, and G. Batista, "The ucr time series classification archive," July 2015. www.cs.ucr.edu/~eamonn/time_series_data/.

[6] R. A. Shapira Weber, M. Eyal, N. Skafte, O. Shriki, and O. Freifeld, "Diffeomorphic temporal alignment nets: Supplemantary material," in *Advances in Neural Information Processing Systems 32* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, eds.), pp. 6574–6585, Curran Associates, Inc., 2019.