

Efficient Temporal Kernels between Feature Sets for Time Series Classification

Supplementary material

Romain Tavenard¹, Simon Malinowski², Laetitia Chapel³, Adeline Bailly¹,
Heider Sanchez⁴, and Benjamin Bustos⁴

¹ Univ. Rennes 2 / LETG-Rennes COSTEL, IRISA

² Univ. Rennes 1 / IRISA

³ Univ. Bretagne Sud / IRISA

⁴ Dept. of Computer Science, Univ. Chile

1 Full results table

The performance in terms of classification error rate of the proposed method (K_{FS}) is given in Table 1, together with competing methods.

Dataset	BOSS [5]	DTD _C [3]	LS [4]	TSBF[2]	BoW[1]	K_{FS}
Adiac *	0.235	0.299	0.478	0.230	0.248	0.179
ArrowHead *	0.166	0.280	0.154	0.246	0.194	0.246
Beef *	0.200	0.333	0.133	0.433	0.267	0.200
BeetleFly	0.100	0.350	0.200	0.200	0.100	0.000
BirdChicken	0.050	0.200	0.200	0.100	0.200	0.200
CBF *	0.002	0.020	0.009	0.012	0.000	0.000
Car *	0.167	0.217	0.233	0.217	0.083	0.067
ChlorineConcentration *	0.339	0.287	0.408	0.308	0.442	0.135
CinCECGtorso	0.113	0.148	0.130	0.288	0.180	0.013
Coffee *	0.000	0.000	0.000	0.000	0.000	0.000
Computers	0.244	0.284	0.416	0.244	0.328	0.296
CricketX	0.264	0.246	0.259	0.295	0.236	0.233
CricketY	0.246	0.226	0.282	0.264	0.226	0.236
CricketZ	0.254	0.226	0.259	0.285	0.203	0.215
DiatomSizeReduction *	0.069	0.085	0.020	0.101	0.111	0.046
Distal[...]AgeGroup *	0.252	0.338	0.281	0.288	0.147	0.175
Distal[...]Correct *	0.272	0.275	0.221	0.217	0.178	0.180
DistalPhalanxTW	0.324	0.424	0.374	0.324	0.210	0.215
ECG200 *	0.130	0.160	0.120	0.160	0.100	0.100
ECG5000 *	0.059	0.076	0.068	0.060	0.051	0.060
ECGFiveDays *	0.000	0.178	0.000	0.123	0.000	0.055
Earthquakes *	0.252	0.295	0.259	0.252	0.202	0.199
ElectricDevices *	0.201	0.406	0.413	0.297	0.361	0.340
FaceAll *	0.218	0.101	0.251	0.256	0.226	0.198
FaceFour *	0.000	0.182	0.034	0.000	0.023	0.045
FacesUCR *	0.043	0.092	0.061	0.133	0.054	0.075

FiftyWords	0.295	0.246	0.270	0.242	0.266	0.198
Fish	0.011	0.074	0.040	0.166	0.023	0.023
FordA	0.070	0.235	0.043	0.150	0.085	0.070
FordB	0.289	0.347	0.083	0.401	0.111	0.082
GunPoint	0.000	0.013	0.000	0.013	0.020	0.020
Ham *	0.333	0.448	0.333	0.238	0.352	0.276
HandOutlines *	0.097	0.135	0.519	0.146	0.137	0.119
Haptics *	0.539	0.601	0.532	0.510	0.490	0.506
Herring *	0.453	0.453	0.375	0.359	0.438	0.391
InlineSkate *	0.484	0.491	0.562	0.615	0.591	0.629
InsectWingbeatSound *	0.477	0.527	0.394	0.375	0.426	0.372
ItalyPowerDemand *	0.091	0.049	0.040	0.117	0.056	0.050
LargeKitchenAppliances	0.235	0.205	0.299	0.472	0.163	0.136
Lightning2 *	0.164	0.131	0.180	0.262	0.164	0.246
Lightning7	0.315	0.342	0.205	0.274	0.288	0.260
Mallat	0.062	0.073	0.050	0.040	0.146	0.068
Meat *	0.100	0.067	0.267	0.067	0.117	0.100
MedicalImages *	0.282	0.255	0.336	0.295	0.253	0.246
Middle[...]AgeGroup *	0.455	0.500	0.429	0.422	0.200	0.215
Middle[...]Correct *	0.220	0.258	0.220	0.186	0.312	0.318
MiddlePhalanxTW *	0.455	0.500	0.494	0.403	0.386	0.358
MoteStrain	0.121	0.232	0.117	0.097	0.169	0.179
NonInvasiveFatalECGThorax1	0.162	0.159	0.741	0.158	0.076	0.045
NonInvasiveFatalECGThorax2	0.099	0.110	0.230	0.138	0.061	0.042
OSULeaf	0.045	0.116	0.223	0.240	0.087	0.145
OliveOil *	0.133	0.133	0.833	0.167	0.167	0.133
PhalangesOutlinesCorrect *	0.228	0.239	0.235	0.170	0.219	0.179
Phoneme *	0.735	0.732	0.782	0.724	0.731	0.752
Plane	0.000	0.000	0.000	0.000	0.000	0.000
Proximal[...]AgeGroup *	0.166	0.205	0.166	0.151	0.132	0.180
Proximal[...]Correct *	0.151	0.206	0.151	0.127	0.192	0.131
ProximalPhalanxTW *	0.200	0.229	0.224	0.190	0.210	0.205
RefrigerationDevices	0.501	0.555	0.485	0.528	0.523	0.485
ScreenType *	0.536	0.563	0.571	0.491	0.525	0.523
ShapeletSim *	0.000	0.400	0.050	0.039	0.011	0.161
ShapesAll *	0.092	0.162	0.232	0.815	0.092	0.142
SmallKitchenAppliances *	0.275	0.352	0.336	0.328	0.285	0.171
SonyAIBORobotSurface1	0.368	0.290	0.190	0.205	0.092	0.218
SonyAIBORobotSurface2	0.141	0.108	0.125	0.222	0.179	0.153
StarlightCurves	0.022	0.038	0.053	0.023	0.021	0.022
Strawberry *	0.024	0.043	0.089	0.046	0.047	0.044
SwedishLeaf *	0.078	0.104	0.093	0.085	0.082	0.061
Symbols *	0.033	0.037	0.068	0.054	0.017	0.109
SyntheticControl	0.033	0.003	0.003	0.007	0.003	0.007
ToeSegmentation1	0.061	0.193	0.066	0.219	0.026	0.044
ToeSegmentation2	0.038	0.285	0.085	0.200	0.077	0.092
Trace	0.000	0.010	0.000	0.020	0.000	0.000

TwoLeadECG	0.019	0.015	0.004	0.134	0.024	0.040
TwoPatterns	0.007	0.000	0.007	0.024	0.007	0.000
UWaveGestureLibraryAll *	0.061	0.062	0.047	0.074	0.173	0.029
UWaveGestureLibraryX	0.238	0.225	0.209	0.169	0.231	0.171
UWaveGestureLibraryY	0.315	0.302	0.297	0.264	0.312	0.246
UWaveGestureLibraryZ	0.305	0.321	0.253	0.228	0.274	0.236
Wafer	0.005	0.007	0.004	0.005	0.008	0.005
Wine *	0.259	0.389	0.500	0.389	0.370	0.352
WordSynonyms *	0.362	0.270	0.393	0.312	0.365	0.281
Worms	0.442	0.351	0.390	0.312	0.343	0.304
WormsTwoClass	0.169	0.377	0.273	0.247	0.271	0.265
Yoga	0.082	0.144	0.166	0.181	0.131	0.077
Average rank	3.165	4.371	3.912	3.753	3.176	2.624

Table 1: Comparison of error rates between our temporal feature set kernel ($D = 8192$) and local feature based state-of-the-art methods. Recall that BoW baseline operates on the exact same feature sets as our K_{FS} . Stars indicate datasets for which cross-validation leads to take temporal information into account ($\gamma_t > 0$).

2 Parameter ranges

Table 2 gives the search range of the different parameters used for cross-validation.

Parameter name	Kernels	Parameter range
C	All kernels	$10^0 - 10^6$
γ_K	All kernels	$10^{-1} - 10^5$
γ_f	All variants of K_{FS}	$10^0 - 10^6$
γ_t	All temporal variants of K_{FS}	$10^0 - 10^6$

Table 2: Parameter grids for all kernels used in the paper. For all parameters, 5 values are selected at regular logscale locations in the range.

3 Timing experiments on more datasets

Fig. 1 represents the mean square error approximation of the approximated kernel matrix versus the timing for two datasets: *Adiac* and *Ham*. These datasets have been chosen so as to complement conclusions drawn in the paper for *ECG200* dataset. Indeed, *Ham* corresponds to a long time series case while *Adiac* has more training time series than *ECG200*. This figure matches Figure 2 of the paper and confirms that a good trade-off is obtained by the Fourier approximation.

Fig. 2 shows the error rate versus the feature map dimension for *Adiac* and *Ham* datasets. It confirms the statements made in the paper (Figure 4) about

the benefits of taking temporal information into account and of the kernel normalization.

Fig. 3 depicts the training and testing time of BoW, SQFD- k -means and SQFD-Fourier as a function of the size of the training set for the same datasets. This figure matches Figure 3 of the paper, and identical conclusions can be drawn. The Fourier approximation leads to lower execution times than k -means approximation. Compared to BOW, training time is reduced, especially for large training sets.

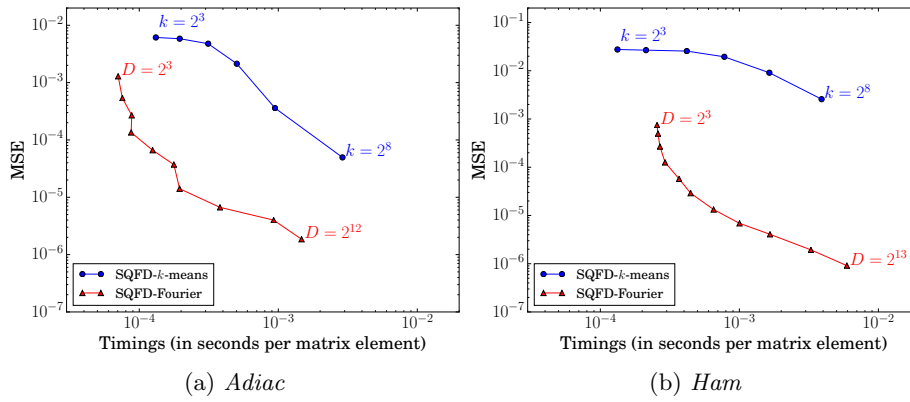


Fig. 1: Mean Squared Error (MSE) vs timings of the approximated kernel matrix (*Adiac* and *Ham*).

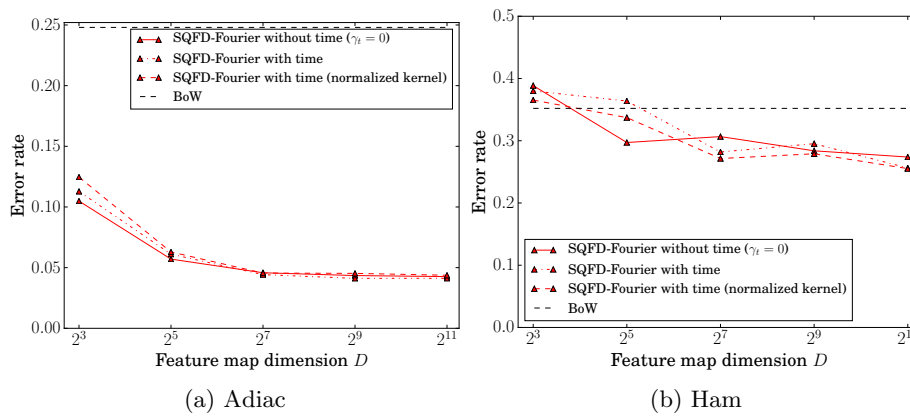


Fig. 2: Error rates as a function of the feature map dimension (*Adiac* and *Ham*).

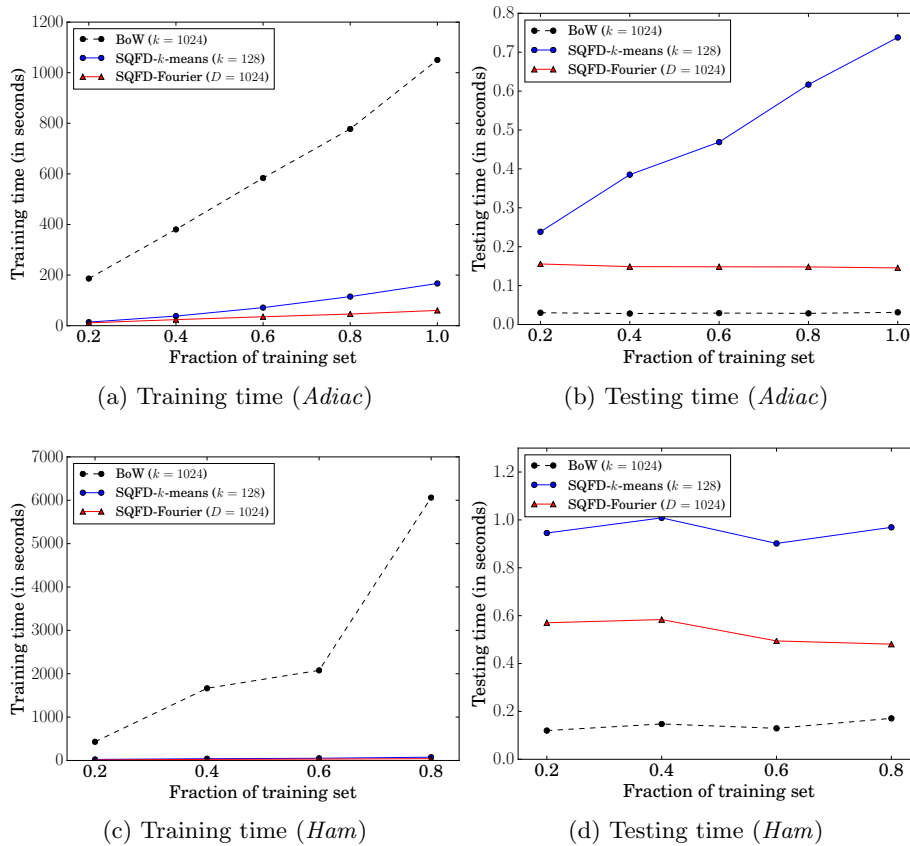


Fig. 3: Training and testing times as a function of the amount of training data (*Adiac* and *Ham*).

4 Pairwise comparisons

Figure 4 compares the performance of the proposed approach with two competing approaches : 1-nearest-neighbour (1NN) combined with DTW and BOSS-VS [6]. Classification error rates on the 85 datasets of the UCR-UEA repository are depicted, together with the Win/Tie/Lose scores and the p -values corresponding to one-sided Wilcoxon signed rank tests.

References

1. Bailly, A., Malinowski, S., Tavenard, R., Chapel, L., Guyet, T.: Dense Bag-of-Temporal-SIFT-Words for Time Series Classification. Lecture Notes in Artificial Intelligence 9785, 17–30 (2016)

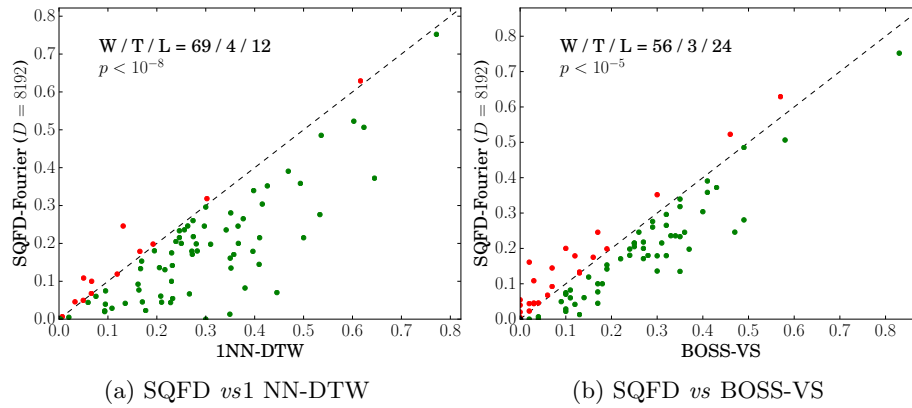


Fig. 4: Pairwise performance comparisons.

2. Baydogan, M.G., Runger, G., Tuv, E.: A Bag-of-Features Framework to Classify Time Series. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35(11), 2796–2802 (2013)
3. Górecki, T., Łuczak, M.: Non-isometric transforms in time series classification using DTW. *Knowledge-Based Systems* 61, 98–108 (2014)
4. Grabocka, J., Schilling, N., Wistuba, M., Schmidt-Thieme, L.: Learning time-series shapelets. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. pp. 392–401 (2014)
5. Schäfer, P.: The BOSS is concerned with time series classification in the presence of noise. *Data Mining and Knowledge Discovery* 29(6), 1505–1530 (2014)
6. Schäfer, P.: Scalable time series classification. *Data Mining and Knowledge Discovery* pp. 1–26 (2015)