Deep Leaning For Time Series Classification

“Practical Time Series Classification: an Overview of Current Research and Solutions for Imperfect Data“

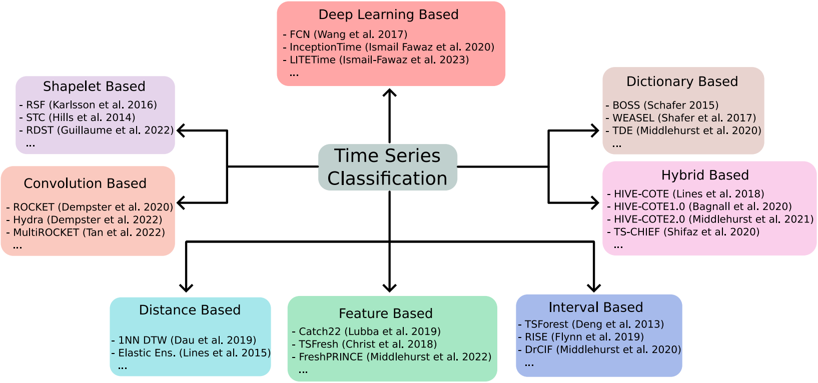
Germain Forestier & Ali Ismail-Fawaz

MSD, IRIMAS, Université de Haute-Alsace, Mulhouse France

IEEE International Conference on Data Science and Advanced Analytics (DSAA)

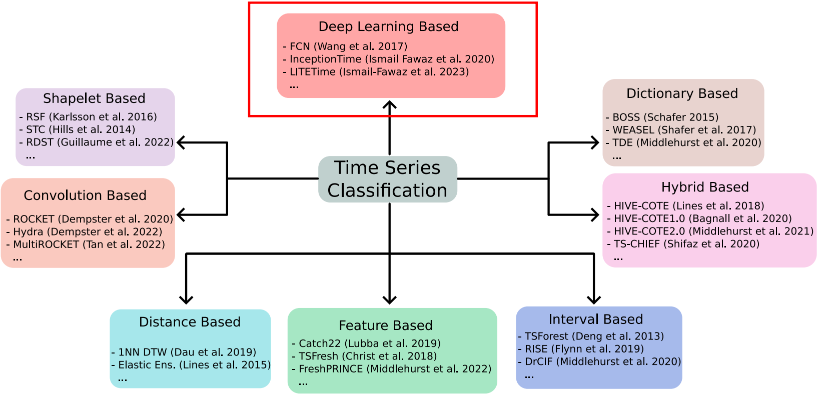
Contents

Taxonomy of methods



Middlehurst, M., Schäfer, P., & Bagnall, A. (2024). Bake off redux: a review and experimental evaluation of recent time series classification algorithms. Data Mining and Knowledge Discovery, 1-74.

Taxonomy of methods



Middlehurst, M., Schäfer, P., & Bagnall, A. (2024). Bake off redux: a review and experimental evaluation of recent time series classification algorithms. Data Mining and Knowledge Discovery, 1-74.

Why Deep Learning ?

- Around 2017, reviewers began to question the potential performance of deep learning for TSC while assessing papers on non-deep learning TSC methods.
- Deep learning has achieved great success in other data types, such as computer vision and natural language processing (NLP), so why not with time series?

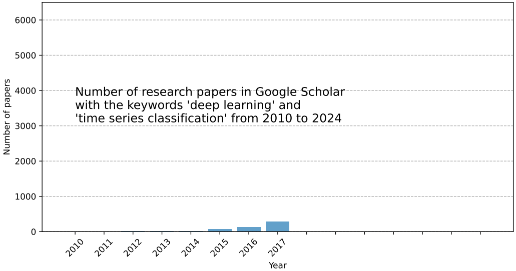


Figure: The number of research papers mentioning "deep learning" and "time series classification" increased rapidly in the last years:

Why Deep Learning ?

- Around 2017, reviewers began to question the potential performance of deep learning for TSC while assessing papers on non-deep learning TSC methods.
- Deep learning has achieved great success in other data types, such as computer vision and natural language processing (NLP), so why not with time series?

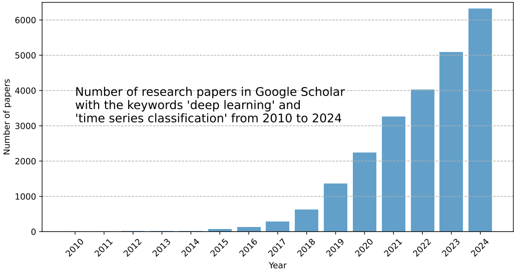
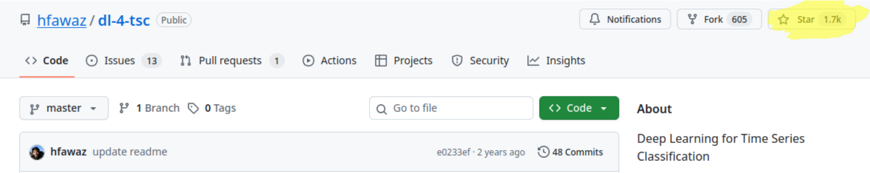


Figure: The number of research papers mentioning "deep learning" and "time series classification" increased rapidly in the last years.

DL4TSC - 2019

In **2019**, we presented a study of **Deep Learning for Time Series Classification** (cited more than 4.2K times (GoogleScholar)) [1].

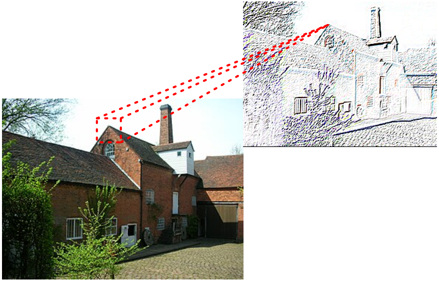
- We selected models with enough details (or available code) to reproduce the model's architecture and benchmarked them on the UCR archive [2]
- We published the code on Github for reproducibility and got very positive feedback (>1.7K stars) `https://github.com/hfawaz/dl-4-tsc`



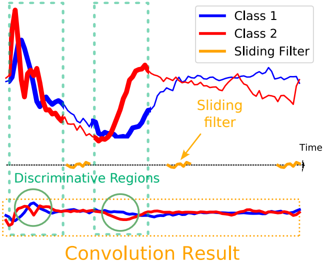
[1] Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., & Muller, P. A. (2019). Deep learning for time series classification: a review. Data mining and knowledge discovery, 33(4), 917-963.

[2] Dau, H. A., Bagnall, A., Kamgar, K., Yeh, C. C. M., Zhu, Y., Gharghabi, S. & Keogh, E. (2019). The UCR time series archive. IEEE/CAA Journal of Automatica Sinica

Convolutions on Images vs Time Series



The result of a applying an edge detection convolution on an image

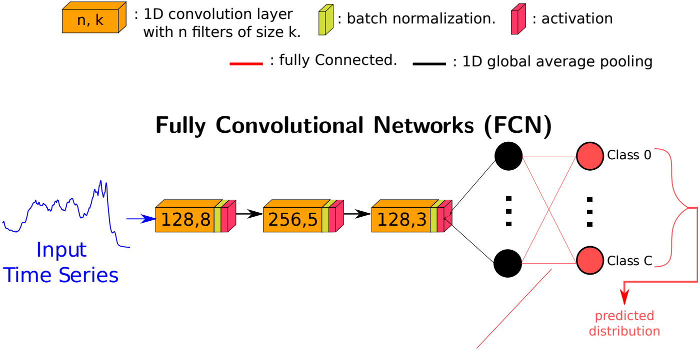


The result of applying a learned discriminative convolution on the ECG200 dataset

Image source: Sarehole Mill, Hall Green, Birmingham, Wikipedia.

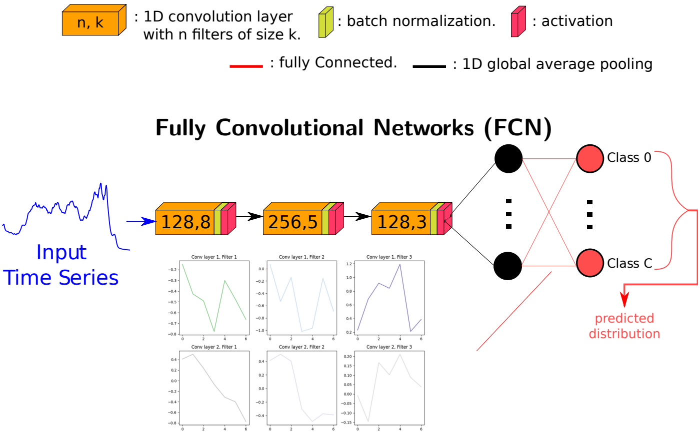
Contents

DL4TSC - Some Architectures



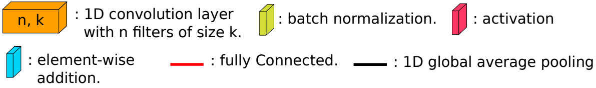
Wang, Z., Yan, W., & Oates, T. (2017, May). Time series classification from scratch with deep neural networks: A strong baseline. In 2017 International joint conference on neural networks (IJCNN) (pp. 1578-1585). IEEE.

DL4TSC - Some Architectures

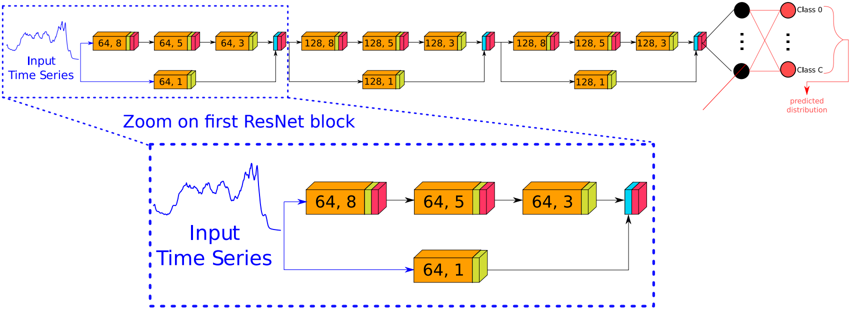


Wang, Z., Yan, W., & Oates, T. (2017, May). Time series classification from scratch with deep neural networks: A strong baseline. In 2017 International joint conference on neural networks (IJCNN) (pp. 1578-1585). IEEE.

DL4TSC - Some Architectures

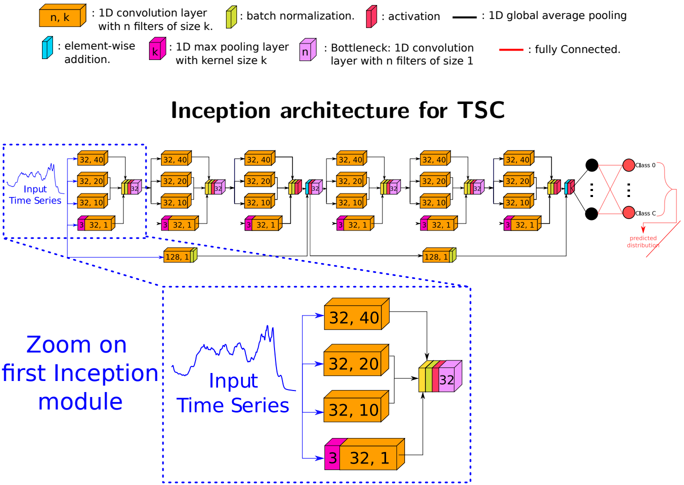


Residual Network (ResNet):



Wang, Z., Yan, W., & Oates, T. (2017, May). Time series classification from scratch with deep neural networks: A strong baseline. In 2017 International joint conference on neural networks (IJCNN) (pp. 1578-1585). IEEE.

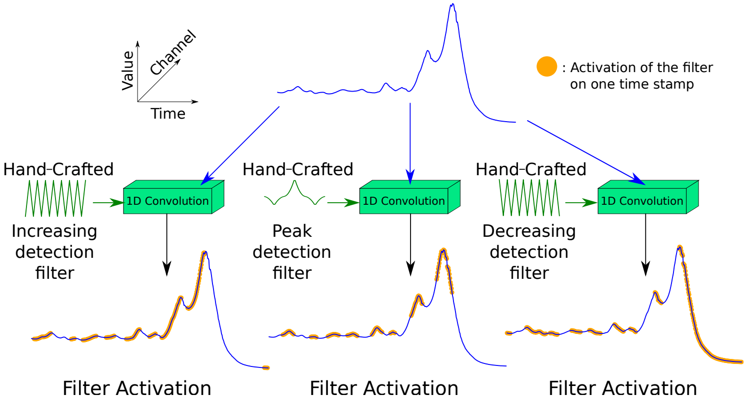
InceptionTime: Ensemble of Inception Models



Ismail Fawaz, Hassan, et al. "Inceptiontime: Finding alexnet for time series classification." Data Mining and Knowledge Discovery 34.6 (2020): 1936-1962.

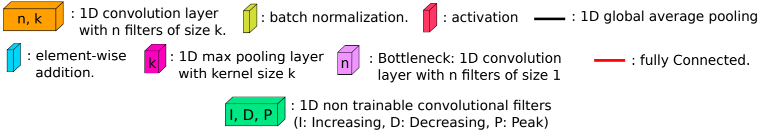
Hand-Crafted Filters

Bridging the Gap Between Random and Fully Learned Filters

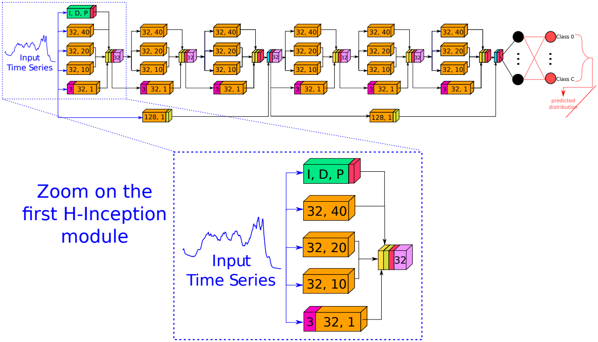


Ismail-Fawaz, A., Devanne, M., Weber, J., & Forestier, G. (2022). Deep learning for time series classification using new hand-crafted convolution filters. In 2022 IEEE International Conference on Big Data (Big Data) (pp. 972-981)

H-InceptionTime: Ensemble of Hybrid Inception Models



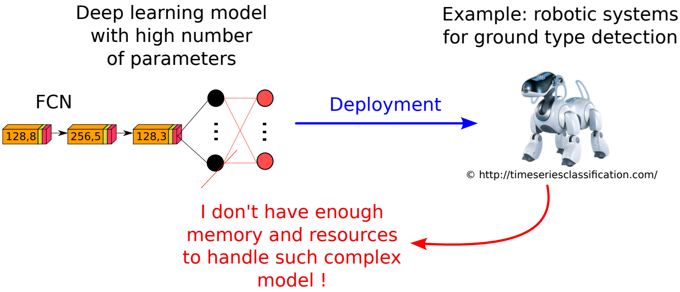
Hybrid Inception (H-Inception) Architecture



Ismail-Fawaz, A., Devanne, M., Weber, J., & Forestier, G. (2022). Deep learning for time series classification using new hand-crafted convolution filters. In 2022 IEEE International Conference on Big Data (Big Data) (pp. 972-981)

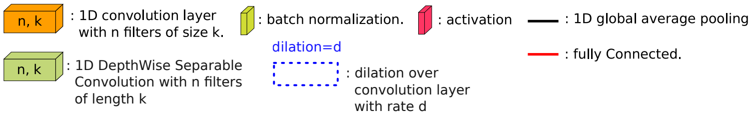
Reducing Models Size

Deep learning models can be difficult to deploy:

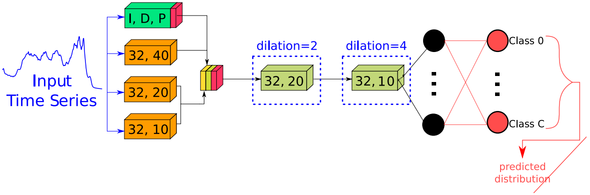


- Hand-crafted convolutional filters in H-FCN demonstrate that complexity is not always the key to better performance.
- Is it possible to reduce complexity without compromising performance?

LITETime: Ensemble of LITE Models



LITE Architecture

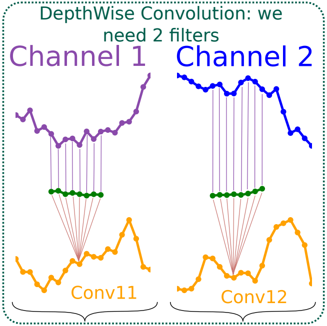


Ismail-Fawaz, A. et al. (2023). LITE: Light Inception with boosTing tEchniques for Time Series Classification. IEEE International Conference on Data Science and Advanced Analytics (DSAA)

Convolutions In LITETime

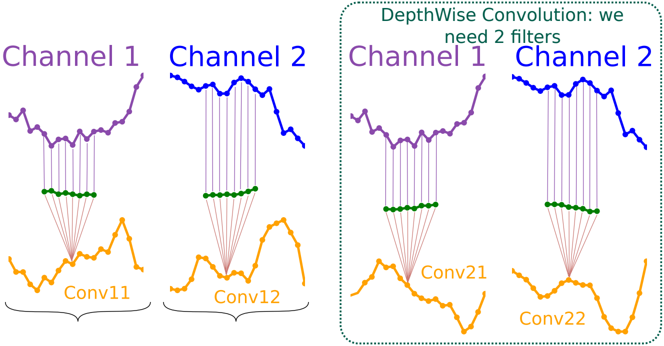
Standard vs Depthwise Separable Convolutions

Standard Convolution = DepthWise Convolution + Summation



Standard vs Depthwise Separable Convolutions

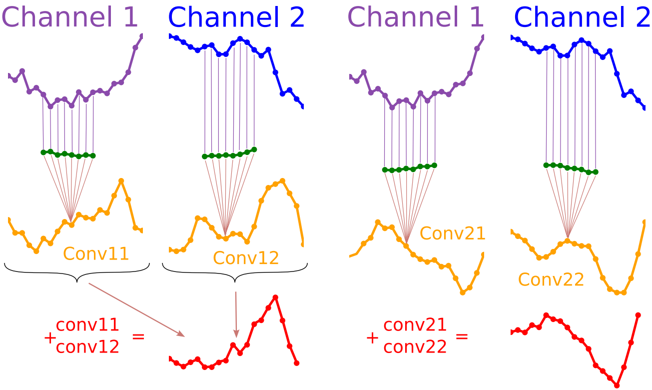
Standard Convolution = DepthWise Convolution + Summation



Convolutions In LITETime

Standard vs Depthwise Separable Convolutions

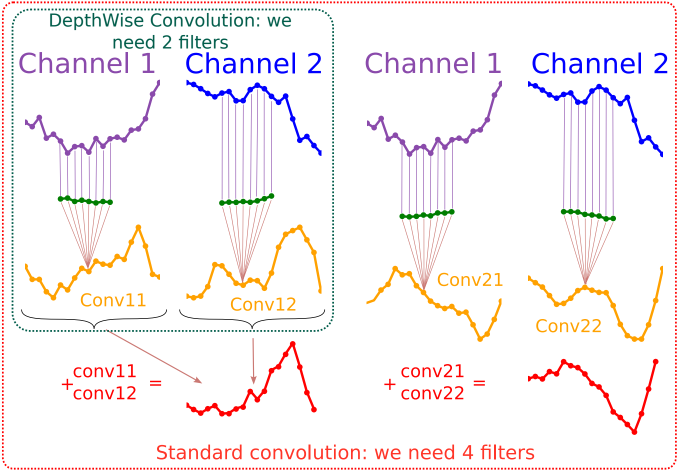
Standard Convolution = DepthWise Convolution + Summation



Convolutions In LITETime

Standard vs Depthwise Separable Convolutions

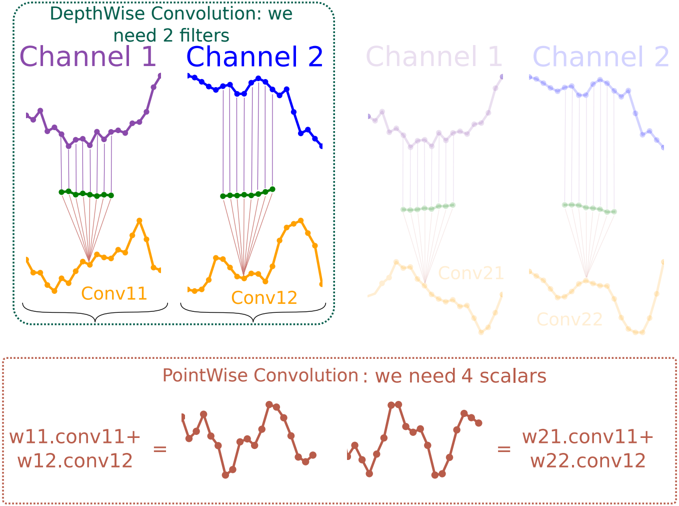
Standard Convolution = DepthWise Convolution + Summation



Convolutions In LITETime

Standard vs Depthwise Separable Convolutions

DepthWise Separable Convolution = DepthWise + PointWise Convolution



Reducing Models Size: LITE - Efficiency Comparison

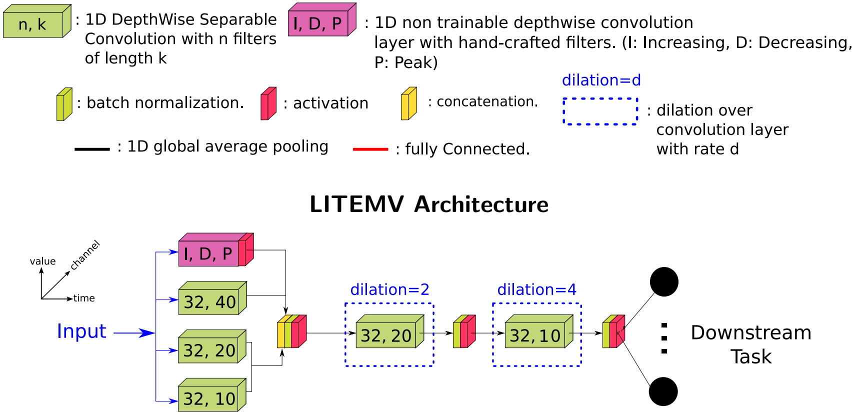
Table: Comparison between the proposed method with FCN, ResNet and Inception without ensemble.

Models	Number of parameters	Training Time (mins)	Testing Time (mins)	CO_2 (g)
Inception	420,192	2,419.2	1.296	0.2928
ResNet	504,000	2,750.4	1.008	0.3101
FCN	264,704	2,491.2	0.4464	0.2623
LITE	**9,814**	**829.8**	**0.72**	**0.1048**

CO_2 calculation: https://codecarbon.io/

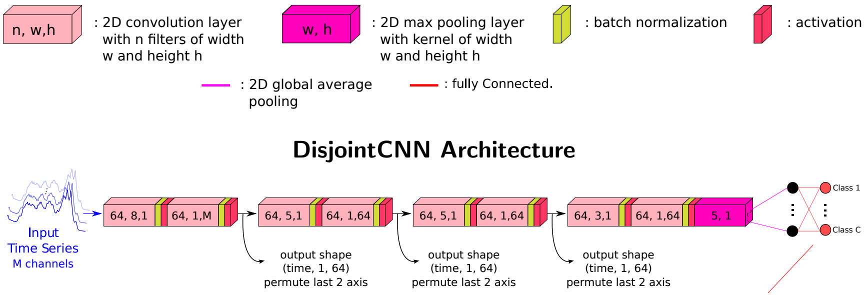
Ismail-Fawaz, A. et al. (2023). LITE: Light Inception with boosTing tEchniques for Time Series Classification. IEEE International Conference on Data Science and Advanced Analytics (DSAA)

Deep Learning for Multivariate Time Series



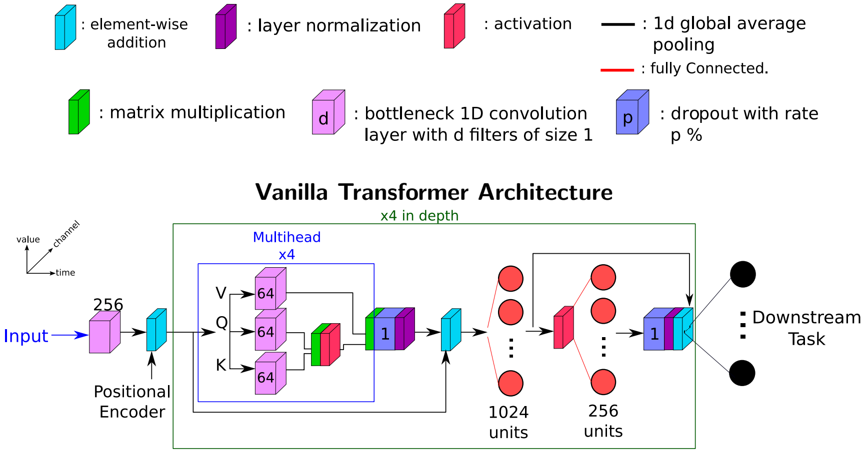
Ismail-Fawaz, A., Devanne, M., Berretti, S., Weber, J. and Forestier, G., 2025. Look into the lite in deep learning for time series classification. International Journal of Data Science and Analytics, pp.1-21.

Deep Learning for Multivariate Time Series



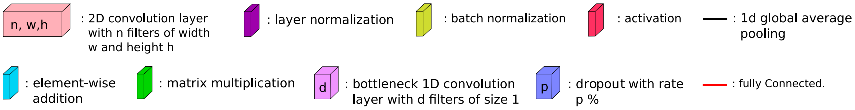
Foumani, S.N.M., Tan, C.W. and Salehi, M., 2021, December. Disjoint-cnn for multivariate time series classification. In 2021 International Conference on Data Mining Workshops (ICDMW) (pp. 760-769). IEEE.

Deep Learning for Multivariate Time Series

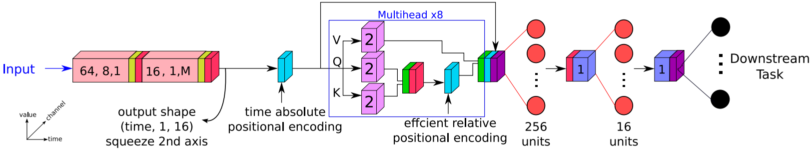


Petrovich, M., Black, M.J. and Varol, G., 2021. Action-conditioned 3d human motion synthesis with transformer vae. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10985-10995).

Deep Learning for Multivariate Time Series



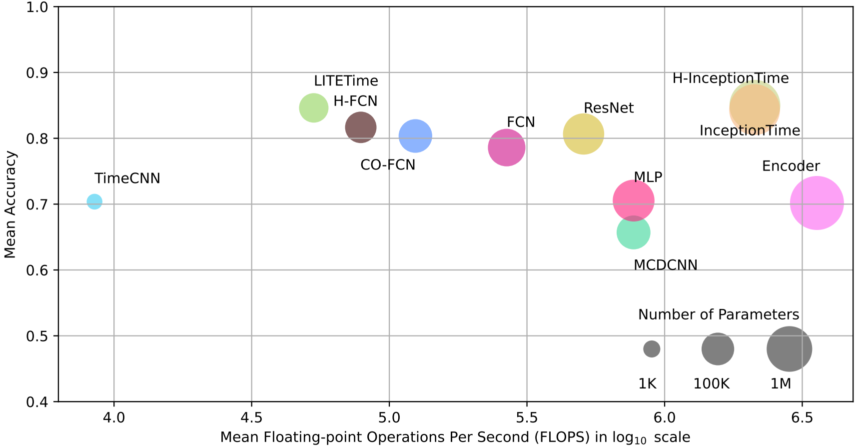
Convolution Transformer Architecture (ConvTran)



📘 Foumani, N.M., Tan, C.W., Webb, G.I. and Salehi, M., 2024. Improving position encoding of transformers for multivariate time series classification. Data mining and knowledge discovery, 38(1), pp.22-48.

Comparing all the architectures

We created a dynamic website (updated regularly) to compare all these architectures in terms of performance and complexity on univariate time series classification data:



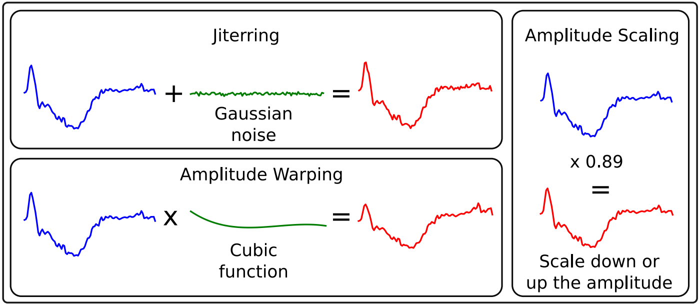
128 datasets from the UCR archive Try it out on : `https://msd-irimas.github.io/pages/dl4tsc/`

Contents

Regularization Techniques: Data Augmentation

How to create synthetic time series ?

- We add noise to the series
- We scale the time series



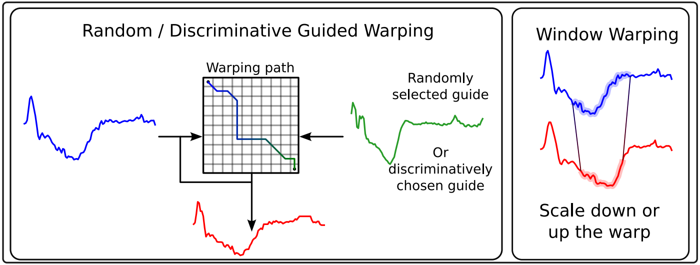
Ismail-Fawaz, A., Devanne, M., Berretti, S., Weber, J. and Forestier, G., Re-framing Time Series Augmentation Through the Lens of Generative Models. ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data, 2025

Um, T.T., Pfister, F.M., Pichler, D., Endo, S., Lang, M., Hirche, S., Fietzek, U. and Kulić, D., 2017, November. Data augmentation of wearable sensor data for parkinson's disease monitoring using convolutional neural networks. In Proceedings of the 19th ACM international conference on multimodal interaction (pp. 216-220).

Regularization Techniques: Data Augmentation

How to create synthetic time series ?

- We warp a time series using the warping path with another series
- We add warping distortions to a window in the time series



Ismail-Fawaz, A., Devanne, M., Berretti, S., Weber, J. and Forestier, G., Re-framing Time Series Augmentation Through the Lens of Generative Models. ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data, 2025

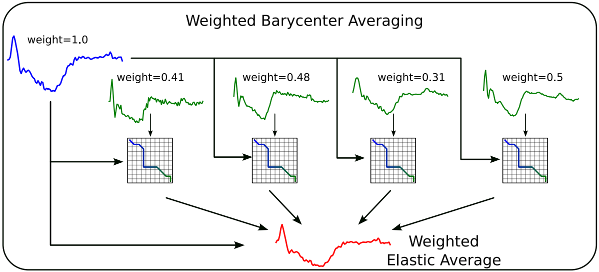
Iwana, B.K. and Uchida, S., 2021, January. Time series data augmentation for neural networks by time warping with a discriminative teacher. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 3558-3565). IEEE.

Le Guennec, A., Malinowski, S. and Tavenard, R., 2016, September. Data augmentation for time series classification using convolutional neural networks. In ECML/PKDD workshop on advanced analytics and learning on temporal data.

Regularization Techniques: Data Augmentation

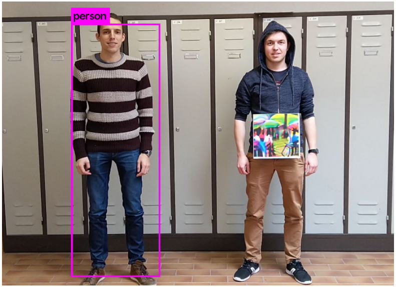
How to create synthetic time series ?

- We averaged a set of time series and took the average as a new synthetic object
- We used weighted averages to generate multiple synthetic objects

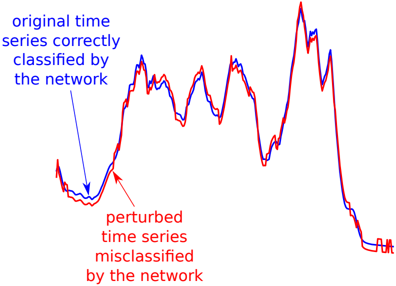


- Ismail-Fawaz, A., Devanne, M., Berretti, S., Weber, J. and Forestier, G., Re-framing Time Series Augmentation Through the Lens of Generative Models. ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data, 2025
- Petitjean, F., Ketterlin, A., & Gançarski, P. (2011). A global averaging method for dynamic time warping, with applications to clustering. Pattern Recognition, 44(3), 678-693.
- Forestier, G., et al. "Generating synthetic time series to augment sparse datasets." 2017 IEEE International Conference on Data Mining (ICDM). IEEE, 2017.
- Ismail-Fawaz, A. et al. " ShapeDBA: Generating Effective Time Series Prototypes using ShapeDTW Barycenter Averaging." ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data, 2023

Regularization Techniques: Adversarial Attacks



Object recognition [1]



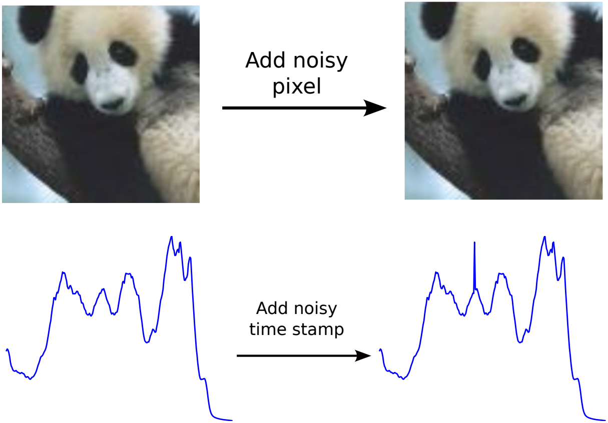
Time series classification [2]

Figure: Adversarial attacks on deep learning systems.

- [1] Van Ranst, W., Thys, S., & Goedemé, T. (2019). Fooling automated surveillance cameras: adversarial patches to attack person detection. In CVPR Workshop on The Bright and Dark Sides of Computer Vision: Challenges and Opportunities for Privacy and Security.
- [2] Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., & Muller, P. (2019). Adversarial Attacks on Deep Neural Networks for Time Series Classification. International Joint Conference on Neural Networks (IJCNN).

Regularization Techniques: Adversarial Attacks

It is easier to see attacks on time series compared to images:



Goodfellow, I.; Shlens, J.; Szegedy, C. Explaining and Harnessing Adversarial Examples. In Proceedings of the International Conference on Learning Representations

Regularization Techniques: Adversarial Attacks

Adversarial attacks on deep neural networks:

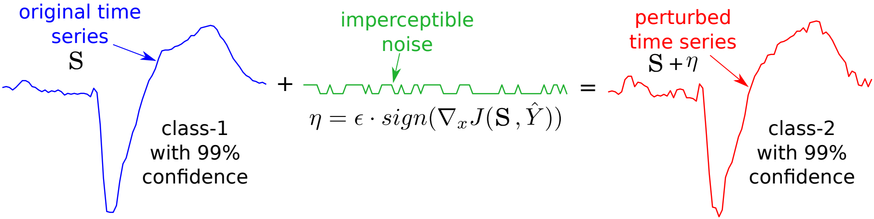


Figure: Example of perturbing the classification of an input time series (from the TwoLeadECG dataset) by adding an imperceptible noise computed using the Fast Gradient Sign Method (FGSM). Figure inspired from [1].

[1] Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and harnessing adversarial examples. International Conference on Learning Representations.

[2] Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L., Muller, P. A. (2019). Adversarial attacks on deep neural networks for time series classification. In International Joint Conference on Neural Networks (IJCNN)

Regularization Techniques: Adversarial Attacks

Smooth Perturbations for Time Series Adversarial Attacks

- A novel adversarial attack that produces smooth perturbations

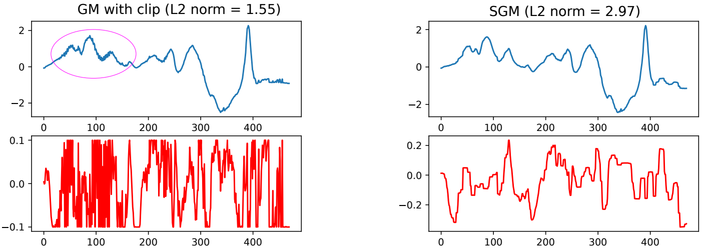


Figure: Time series from the Beef dataset. All methods perturbed time series (blue) and generated noise (red). The purple circles show the presence of saw- tooth on the perturbed time series. [1, 2].

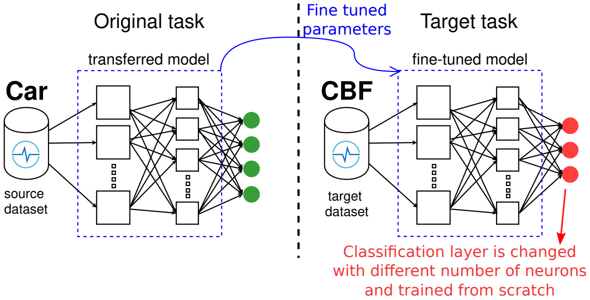
[1] Pialla et al. (2022). Smooth Perturbations for Time Series Adversarial Attacks, Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)

[2] Pialla et al. (2023). Time series adversarial attacks: an investigation of smooth perturbations and defense approaches."International Journal of Data Science and Analytics

Regularization Techniques: Transfer Learning

Transfer Learning:

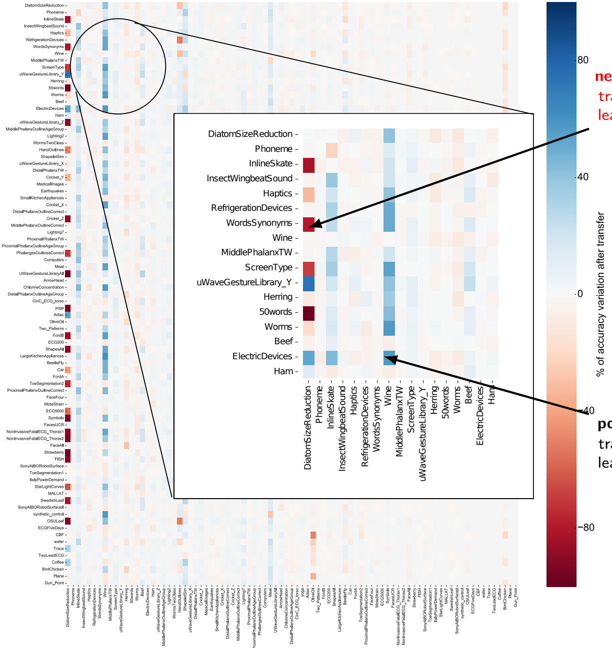
1. Train a base network on a source dataset
2. Transfer the learned features (the network's weights) to a second network and adapt the last layer (class-dependent)
3. Re-train or fine-tune the transferred network on a target dataset



Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., & Muller, P. A. (2018). Transfer learning for time series classification. IEEE International Conference on Big Data.

Ismail-Fawaz, A. et al. " Finding foundation models for time series classification with a pretext task." PAKDD International Workshop on Temporal Analytics, 2024

Regularization Techniques: Transfer Learning



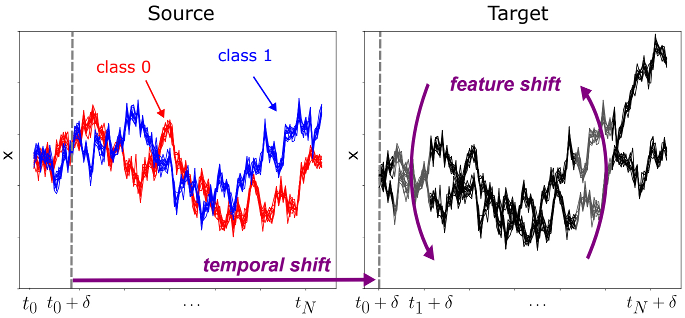
negative transfer learning

positive transfer learning

Regularization Techniques: Domain Adaptation

Domain Adaptation:

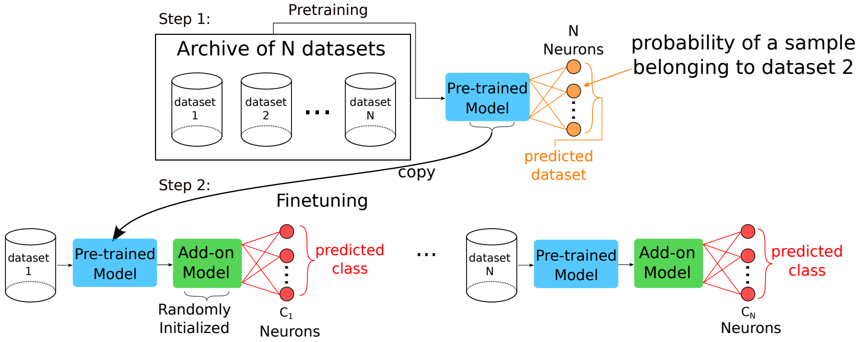
- Train a base network on a source domain
- During training on source, adapt the model's weights to align and generalize to the target domain of the same dataset
- Example: multiple subjects of the same human activity recognition datasets



Ismail Fawaz, H., Del Grosso, G., Kerdoncuff, T., Boisbunon, A. and Saffar, I., 2025. Deep unsupervised domain adaptation for time series classification: a benchmark. Data Mining and Knowledge Discovery, 39(4), p.39.

Regularization Techniques: Foundation Models

1. Instead of training on a single large or small dataset, train on multiple datasets at the same time (unsupervised pretext task)
2. Fine-tune the pre-trained model on the classification task of each dataset alone
3. Pretext task: predicting the original dataset of each time series

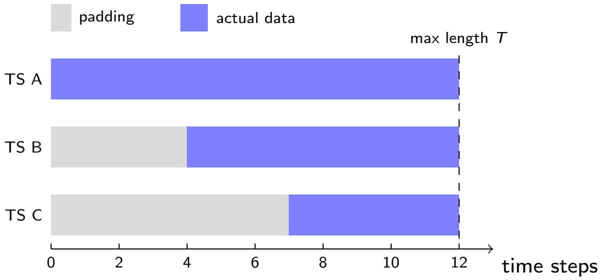


Ismail-Fawaz, A., Devanne, M., Berretti, S., Weber, J. and Forestier, G. "Finding foundation models for time series classification with a pretext task." PAKDD International Workshop on Temporal Analytics, 2024

Contents

Unequal Length:
Deep models require all time series in a batch to have the same length.
Two main strategies:

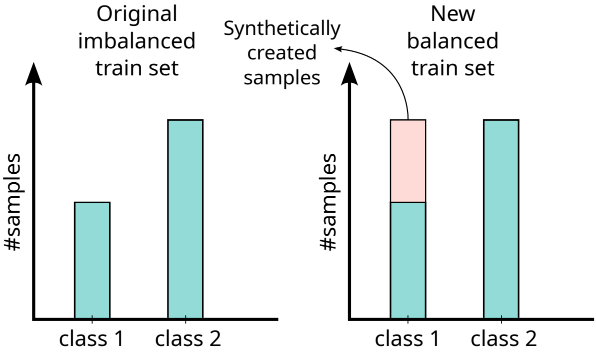
1. Global padding: pad all series (from the beginning) with zeros up to the maximum length.
2. Batch-wise padding: pad only to the longest time series within each batch.



Imperfect Data in Deep Learning for Time Series Classification

Missing Values:
Generic issue, not specific to deep learning. Usual strategies: imputation, interpolation, masking.

Class Imbalance:
Handled via data augmentation or re-sampling.

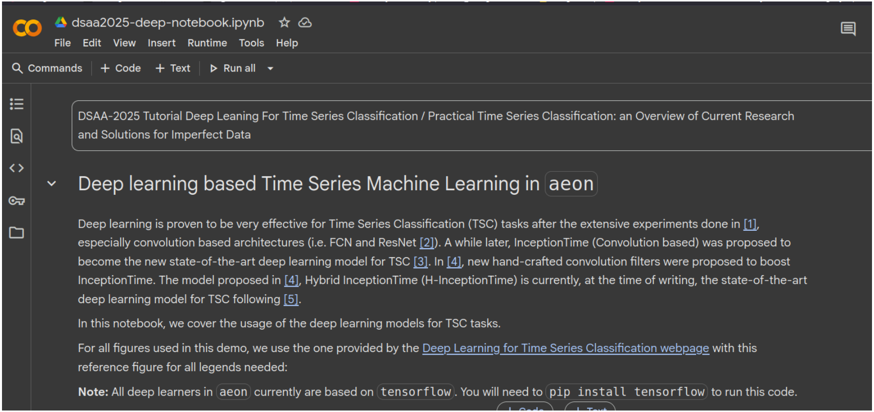


The supervised loss can be weighted during training based on the number of samples per class in each batch.

Contents

Lets code!

https://colab.research.google.com/drive/1vF0edHTBRevLkgLeHmYpRGyIOBWrWC-1?usp=sharing

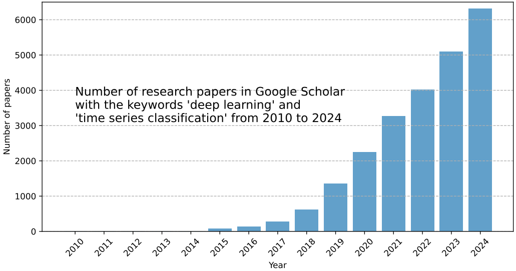


https://msd-irimas.github.io/pages/dl4tsc/

Conclusion

Hot Topics:

- Data Augmentation for Time Series Data
- Self-supervised Learning / Representation Learning
- Foundation Models for Time Series Data [1]
- Transformers for Time Series Classification [2]
- Multivariate Time Series Classification



[1] Ismail-Fawaz, A. (2023). Finding Foundation Models for Time Series Classification with a PreText Task. arXiv preprint arXiv:2311.14534.

[2] Foumani, N. M., Tan, C. W., Webb, G. I., & Salehi, M. (2023). Improving Position Encoding of Transformers for Multivariate Time Series Classification. arXiv preprint arXiv:2305.16642.