A Deep Dive into Alternatives to the **Global Average Pooling** for Time Series Classification

Cyril Meyer, Ali Ismail-Fawaz, Maxime Devanne, Jonathan Weber, Germain Forestier

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

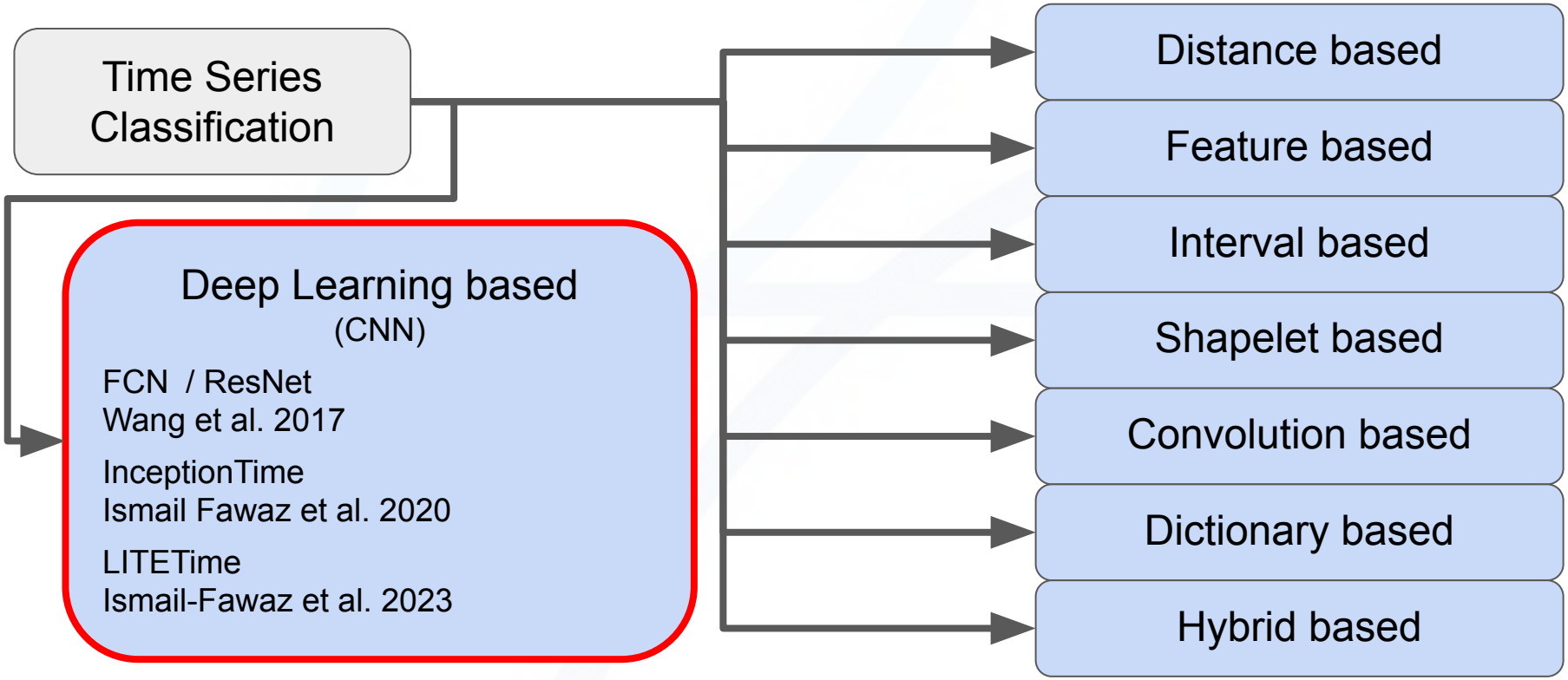
10th Workshop on Advanced Analytics and Learning on Temporal Data (AALTD 2025)

UNIVERSITÉ HAUTE-ALSACE

IRIMAS UNIVERSITÉ HAUTE-ALSACE

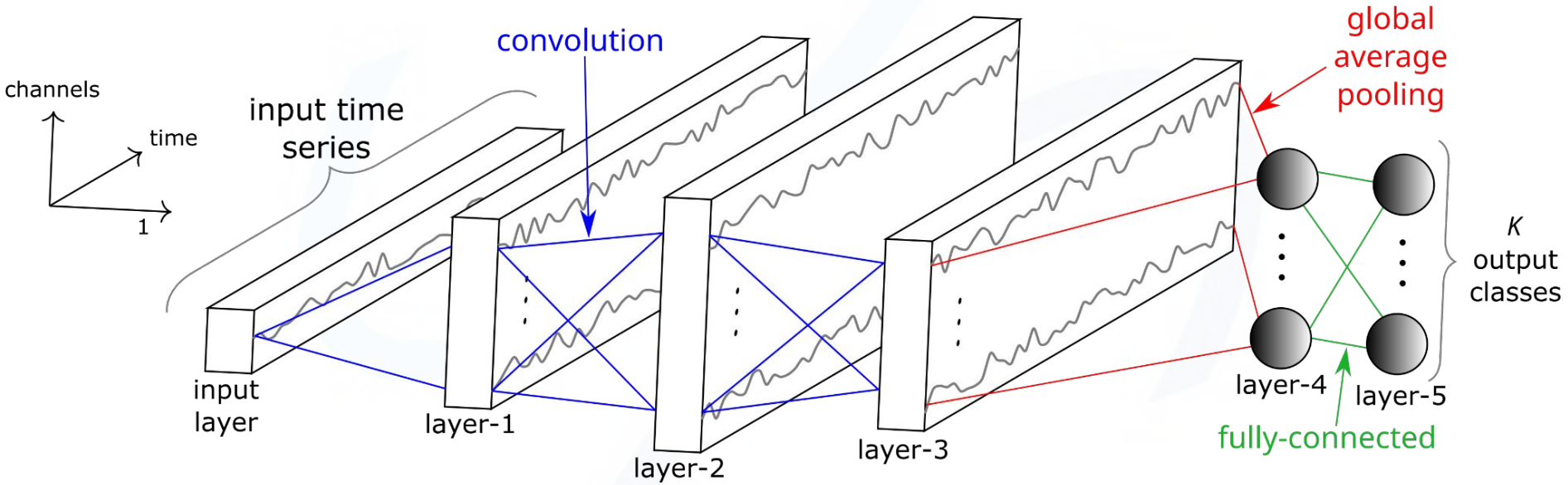
Institut de Recherche en Informatique, Mathématiques, Automatique et Signal

Deep Learning for Time Series Classification



Middlehurst, M., Schäfer, P., Bagnall, A.: Bake off redux: a review and experimental evaluation of recent time series classification algorithms. Data Min. Knowl. Discov. 38, 1958–2031 (2024).
Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., Muller, P.-A.: Deep learning for time series classification: a review. Data Min. Knowl. Discov. 33, 917–963 (2019).

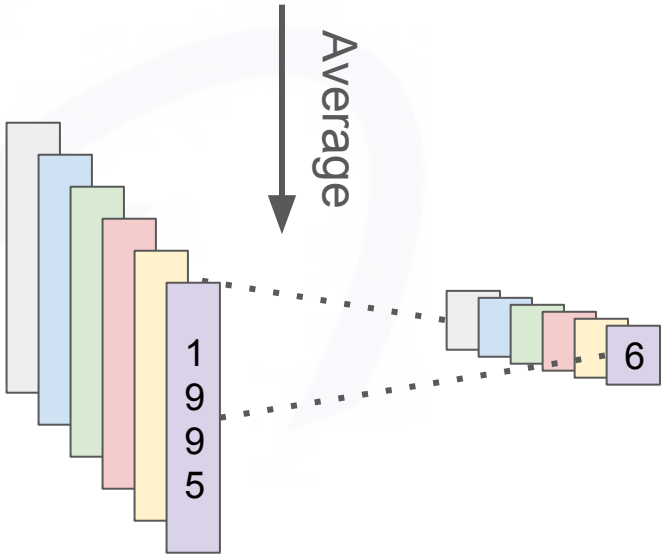
Global Average Pooling (GAP)



📝 Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., Muller, P.-A.: Deep learning for time series classification: a review. Data Min. Knowl. Discov. 33, 917–963 (2019).

Global Average Pooling (GAP)

- **Average** each **feature** map over the **time** dimension
  - **Dimensionality reduction**
    - Less parameters in FC classifier
- Assumes all temporal features contribute equally to the final decision



Alternatives to the Global Average Pooling

Alternatives to Global Average Pooling

- Categories
  - **Learning based**
  - **Feature based**
  - **Pooling based**

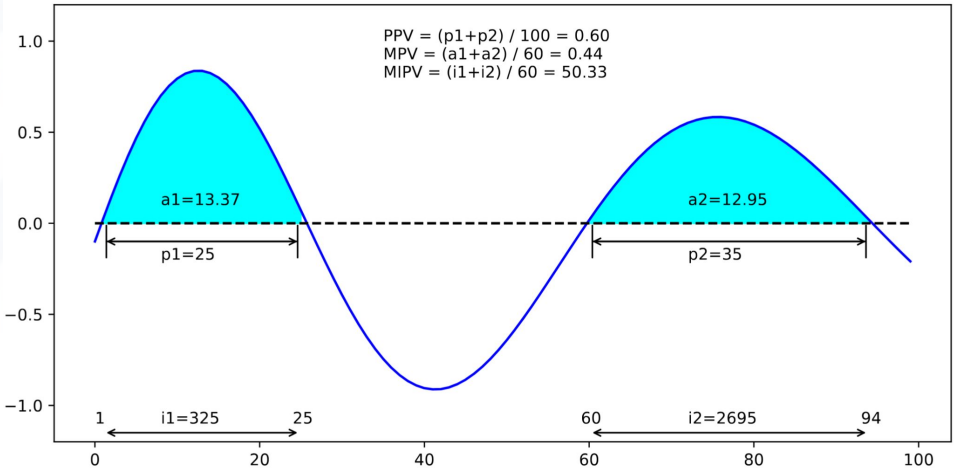
- Learning based / RNN based
  - Gated Recurrent Unit (GRU)
    - Learn to summarize feature

📝 Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y.: Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). pp. 1724–1734. Association for Computational Linguistics, Doha, Qatar (2014).

Alternatives to Global Average Pooling

- Feature based (ROCKET inspired)
  - Proportion of Positive Values (PPV)
  - Mean of Positive Values (MPV)
  - Mean Indices of Positive Values (MIPV)

- ↑PPV variation
  - ↑ Straight through estimator

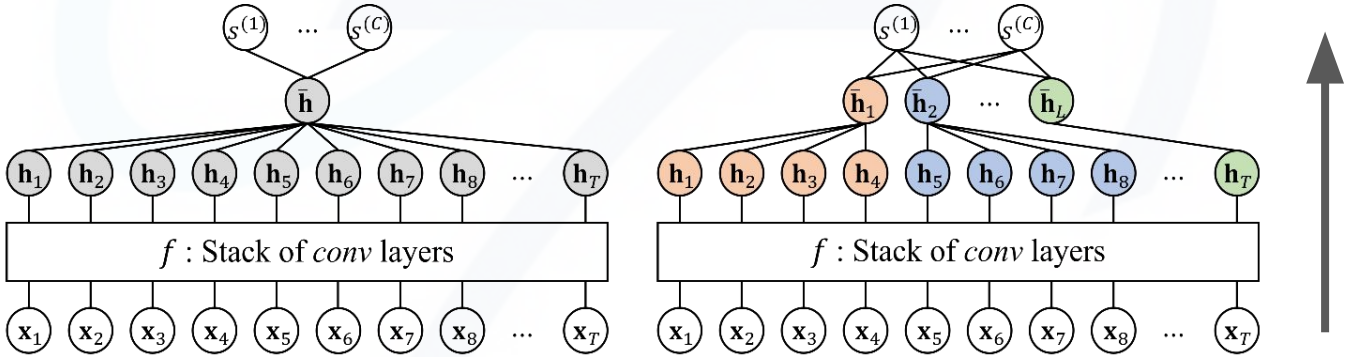


📝 Bengio, Y., Léonard, N., Courville, A.: Estimating or Propagating Gradients Through Stochastic Neurons for Conditional Computation, http://arxiv.org/abs/1308.3432, (2013).
📝 Tan, C.W., Dempster, A., Bergmeir, C., Webb, G.I.: MultiRocket: multiple pooling operators and transformations for fast and effective time series classification. Data Min. Knowl. Discov. 36 (2022).

Alternatives to Global Average Pooling

- Pooling based
  - Global **Max** Pooling (GMP)
  - **Static** Temporal **Average** Pooling (STAP)
  - **Static** Temporal **Max** Pooling (STMP)

] ℹ️ classic pooling in CNN



(a) Global Temporal Pooling (b) Static Temporal Pooling

📝 Lee, D., Lee, S., Yu, H.: Learnable Dynamic Temporal Pooling for Time Series Classification. Proc. AAAI Conf. Artif. Intell. 35, 8288–8296 (2021).

Experimentation & Results

The experiment

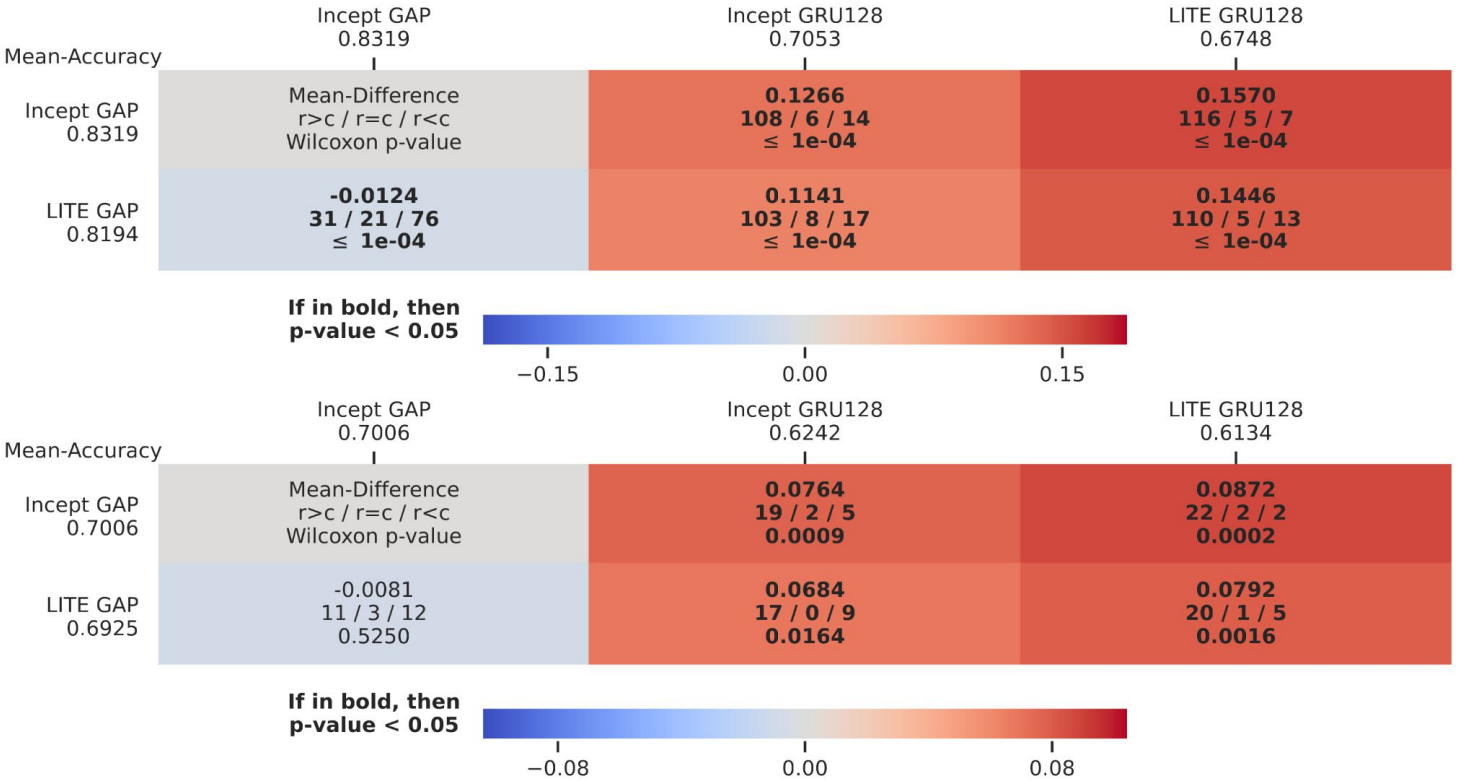
- 4 architectures : FCN / ResNet / Inception / LITE (not everything in the presentation, check the paper !)
- 20 dimensionality reduction alternatives

- Dataset
  - UCR 128 univariate series
  - UEA 26 multivariate series

- 11 repetition (more than 130k trains in total)
  - Median accuracy

📝 Dau, H.A., Bagnall, A., Kamgar, K., Yeh, C.-C.M., Zhu, Y., Gharghabi, S., Ratanamahatana, C.A., Keogh, E.: The UCR time series archive. IEEECAA J. Autom. Sin. 6, 1293–1305 (2019).
📝 Bagnall, A., Dau, H.A., Lines, J., Flynn, M., Large, J., Bostrom, A., Southam, P., Keogh, E.: The UEA multivariate time series classification archive, (2018).

GAP vs RNN



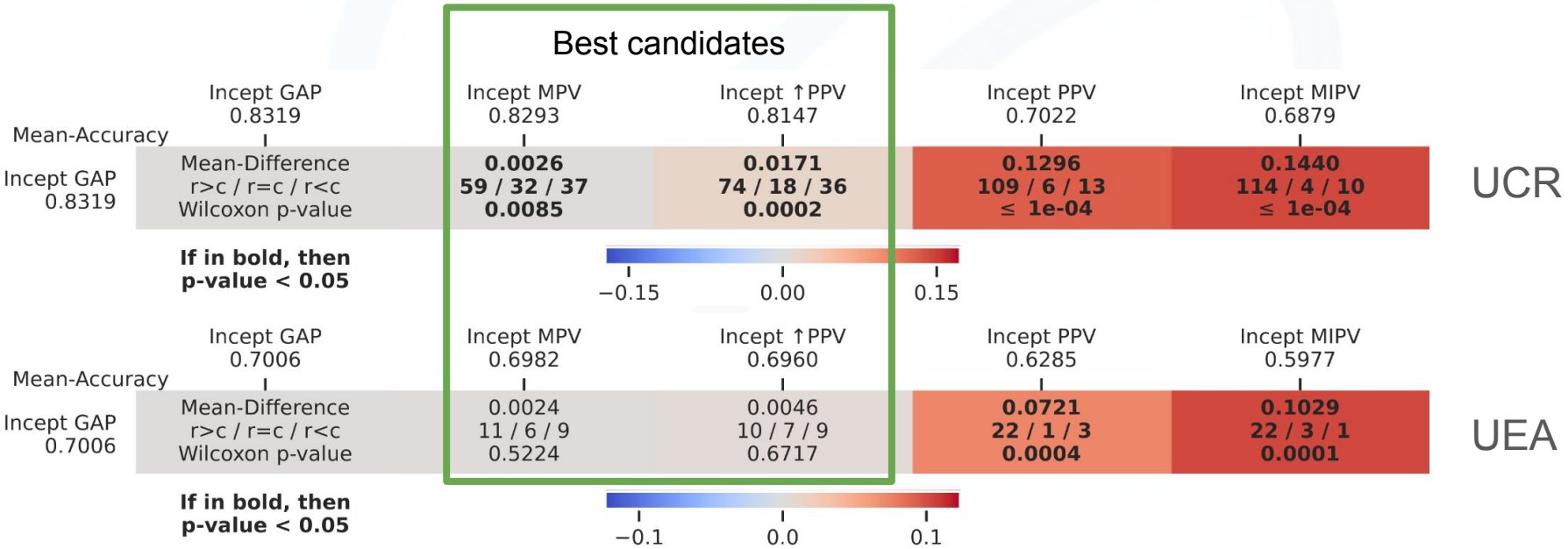
UCR

UEA

📝 Ismail-Fawaz, A., Dempster, A., Tan, C.W., Herrmann, M., Miller, L., Schmidt, D.F., Berretti, S., Weber, J., Devanne, M., Forestier, G., Webb, G.I.: An Approach to Multiple Comparison Benchmark Evaluations that is Stable Under Manipulation of the Comparate Set, http://arxiv.org/abs/2305.11921, (2023).

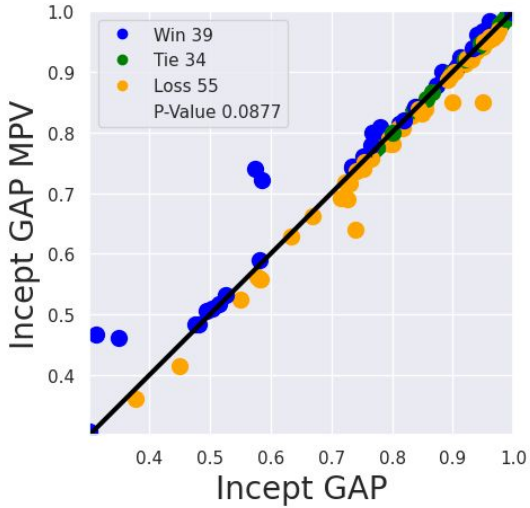
GAP vs Feature

- No alternatives outperform GAP
  - Selection of the best candidates



GAP vs GAP + Feature

- GAP outperform most combination
- Better than a single feature



Mean-Accuracy	Incept GAP MPV 0.8328	Incept GAP 0.8319	Incept GAP MPV ↑PPV 0.8306	Incept GAP ↑PPV 0.8294	Incept ↑PPV MPV MIPV 0.7845
Incept GAP 0.8319	-0.0009 55 / 34 / 39 0.0877	Mean-Difference r>c / r=c / r<c Wilcoxon p-value	**0.0013** **64 / 30 / 34** **0.0043**	0.0024 53 / 35 / 40 0.2294	**0.0474** **101 / 13 / 14** **≤ 1e-04**

UCR

If in bold, then p-value < 0.05

−0.04 0.00 0.04

Mean-Accuracy	Incept GAP 0.7006	Incept GAP MPV 0.6975	Incept GAP ↑PPV 0.6967	Incept GAP MPV ↑PPV 0.6901	Incept ↑PPV MPV MIPV 0.6811
Incept GAP 0.7006	Mean-Difference r>c / r=c / r<c Wilcoxon p-value	0.0031 10 / 8 / 8 0.6889	0.0039 8 / 9 / 9 0.9689	0.0105 11 / 8 / 7 0.2558	**0.0195** **16 / 5 / 5** **0.0107**

UEA

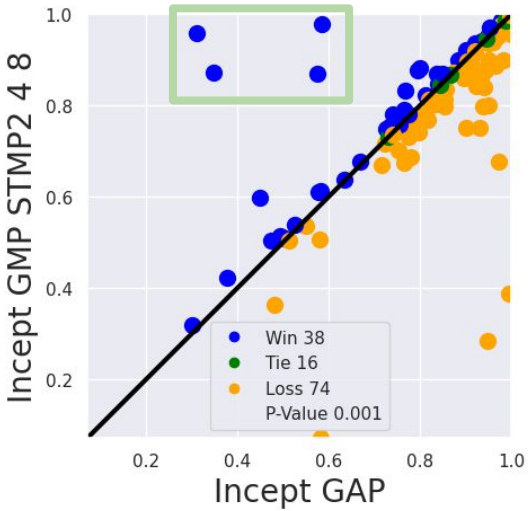
If in bold, then p-value < 0.05

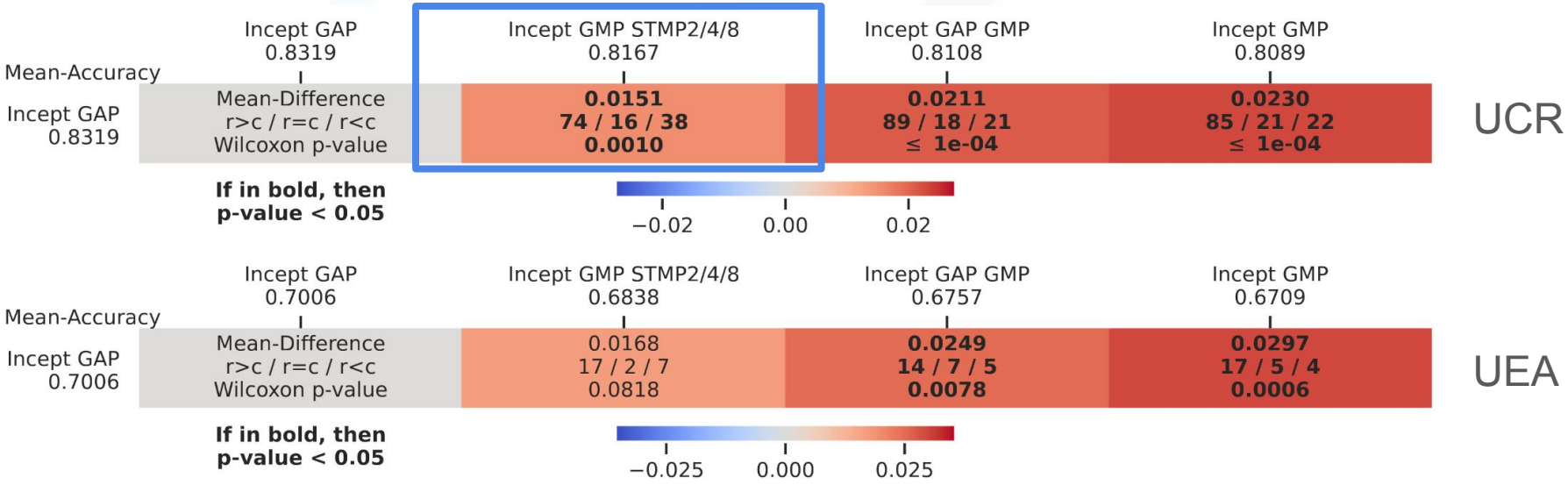
−0.02 0.00 0.02

No GAP

GAP vs Max Poolings

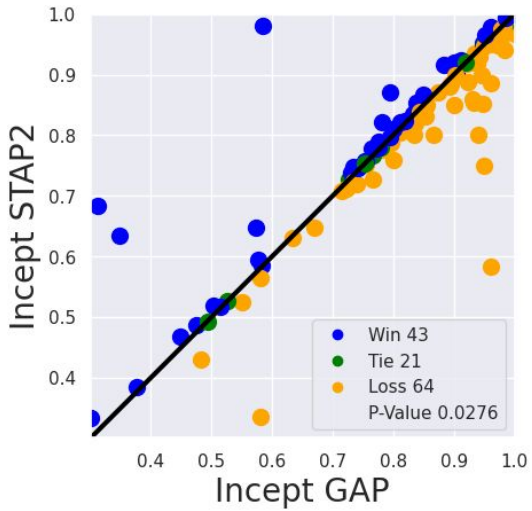
- GMP STMP2/4/8 (2, 4 and 8 window size, not stride)
  - 60% common wins between Inception and LITE





GAP vs Average Poolings

- No average accuracy improvement
- GAP winrate vs STAP2 : 50% / 54%



All combined

Mean-Accuracy	Incept GAP 0.8319	Incept STAP2 0.8277	Incept STAP4 0.8192	Incept GAP STAP2/4/8 0.8121	Incept STAP8 0.8097	
Incept GAP 0.8319	Mean-Difference r>c / r=c / r<c Wilcoxon p-value	**0.0042** **64 / 21 / 43** **0.0276**	**0.0127** **65 / 24 / 39** **0.0030**	**0.0197** **79 / 13 / 36** **0.0001**	**0.0221** **76 / 17 / 35** **0.0001**	UCR

If in bold, then p-value < 0.05

−0.02 0.00 0.02

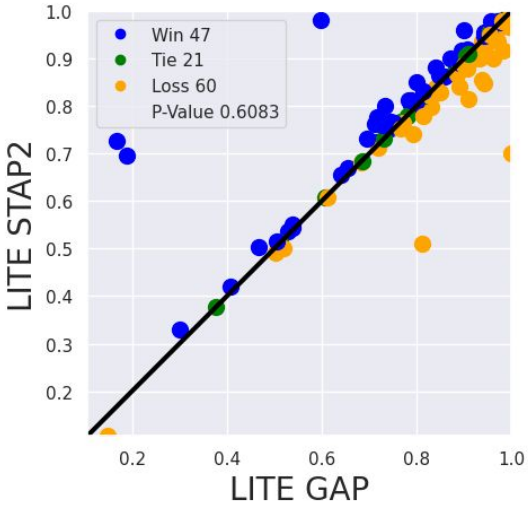
Mean-Accuracy	Incept GAP 0.7006	Incept STAP2 0.6915	Incept STAP4 0.6912	Incept GAP STAP2/4/8 0.6868	Incept STAP8 0.6827	
Incept GAP 0.7006	Mean-Difference r>c / r=c / r<c Wilcoxon p-value	0.0091 14 / 5 / 7 0.2616	0.0094 14 / 4 / 8 0.2572	0.0138 16 / 3 / 7 0.1336	0.0179 13 / 6 / 7 0.0913	UEA

If in bold, then p-value < 0.05

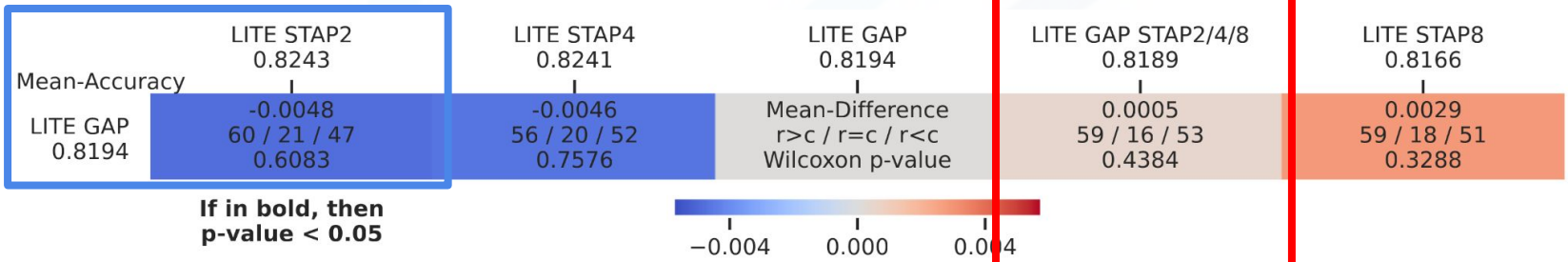
−0.015 0.000 0.015

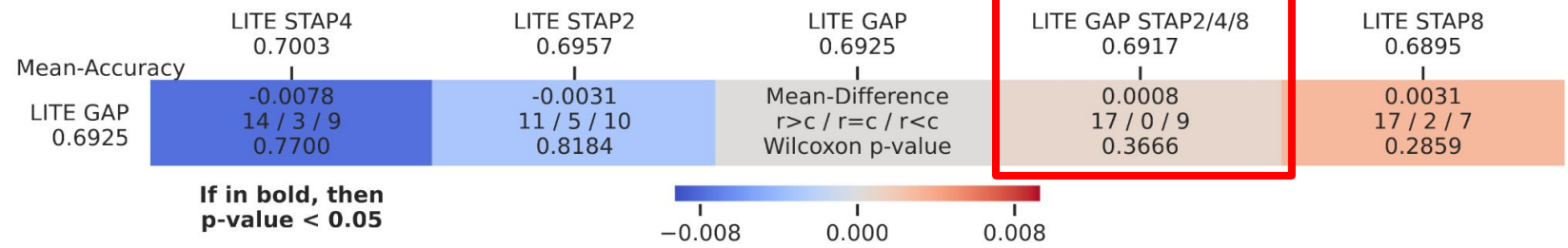
GAP vs Average Poolings (LITE)

- STAP2 wins : 47
- STAP4 wins : 52
- Common wins : 39



All combined





Conclusion & Perspectives

- GAP is a very good default choice
  - Differences are often very close
- Alternatives candidates for specific domain / dataset should not be ignored
- **Never assume** that **adding a layer** will **improve** or be **ignored**

- Larger datasets (e.g. MONSTER)
- Other alternatives
  - Convolutional pooling, Median pooling, Std. pooling, etc.
  - RNN’s with pooling / attention mechanism

Questions

- Check the paper for a lot more ;)
- And the GitHub of course !

- Feel free to contact me
  cyril.meyer@uha.fr
  Linkedin @cyril-meyer

README MIT license

PoolParty-4-TSC

Source code and experimental results for the paper "A Deep Dive into Alternatives to the Global Average Pooling for Time Series Classification" AALTD 2025.

A study on candidates to replace Global Average Pooling (GAP) in neural network architectures for time series classification.

github.com/MSD-IRIMAS/PoolParty-4-TSC

Now, we have the answer to the question :
Why are we using GAP in TSC DL model ?

UNIVERSITÉ HAUTE-ALSACE

IRIMAS UNIVERSITÉ HAUTE-ALSACE
Institut de Recherche en Informatique, Mathématiques, Automatique et Signal