TSLANet: Rethinking Transformers for Time Series Representation Learning

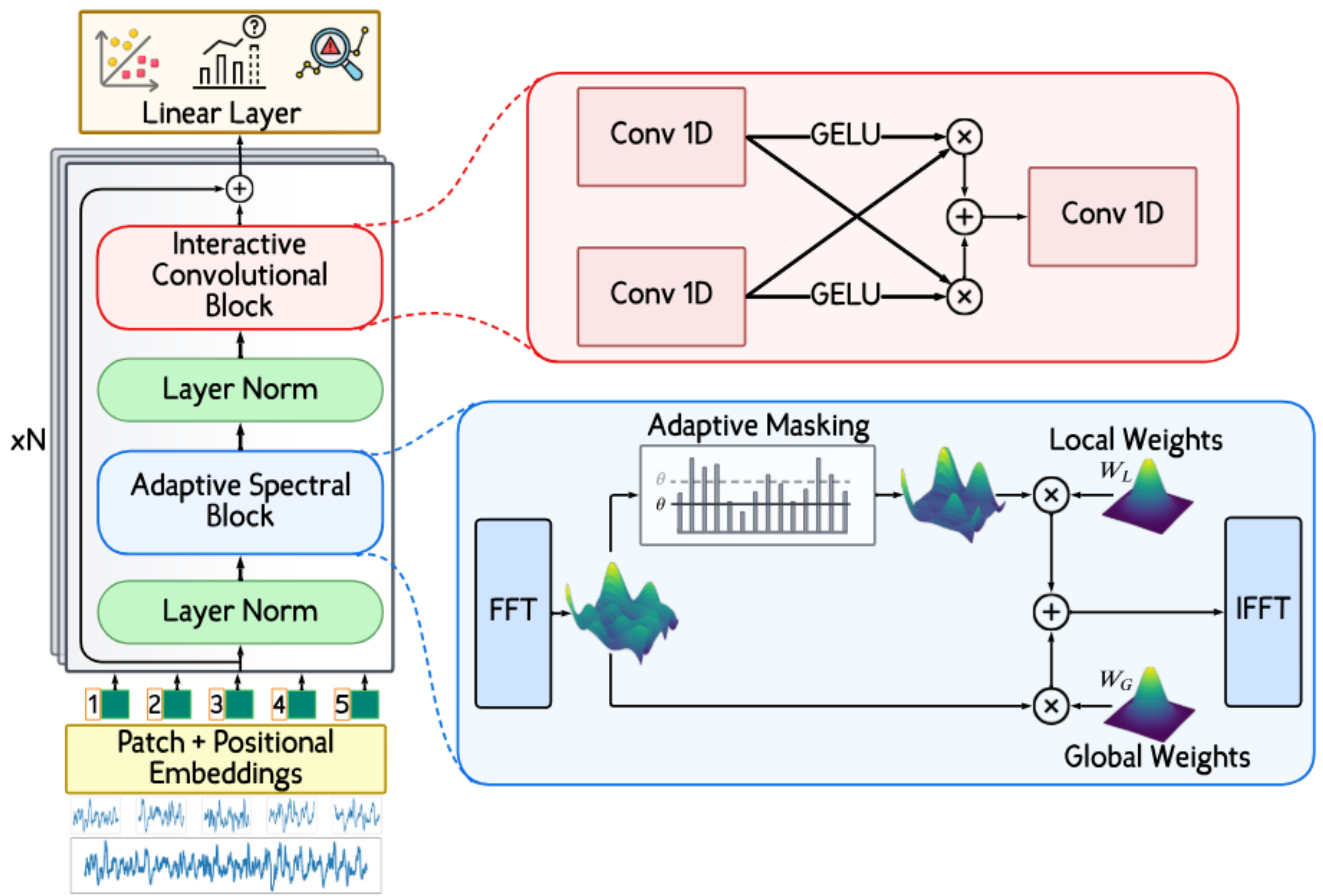
Emadeldeen Eldele[1], Mohamed Ragab[1,2], Zhenghua Chen[1,2], Min Wu[2], and Xiaoli Li[1,2]
[1]Centre for Frontier AI Research, Agency for Science, Technology and Research, Singapore
[2]Institute for Infocomm Research, A*STAR, Singapore

Why TSLANet?

- Traditional models, e.g., Transformers, struggle with capturing both long-range and local patterns effectively.
- Transformer-based models for time series excel in capturing long-range dependencies. However, they are:
  - parameter-heavy,
  - computationally intensive,
  - prune to overfitting on small datasets,
  - inefficient with noisy data or in learning local patterns.

What is TSLANet?

- A novel **T**ime **S**eries **L**ightweight **A**daptive **Net**work that combines the strengths of spectral analysis and CNN in a unified framework.
- A computationally efficient model that learns *both* long- and short-term relationships within the data.
- A flexible and scalable model that can be utilized across different time series tasks.
- We leverage Fourier transform alongside global and local filters to cover wider frequency spectrum, while adaptively removing noise.



Interactive Convolutional Block

- Integrates multiple convolutional layers that interact to refine feature extraction.
- The parallel convolutions have different kernel sizes to capture local features and longer-range dependencies.
- The output of each convolution in the first layer modulates the feature extraction of the other.

Adaptive Spectral Block

- Utilizes Fourier-domain processing to enhance feature representation by focusing on relevant frequency components.
- Two sets of learnable filters; a global filter and a local filter.
- Adaptive Removal of High-Frequency Noise: adaptive local filter to dynamically adjust the level of filtering according to the dataset characteristics.

Main Results (classification and forecasting)

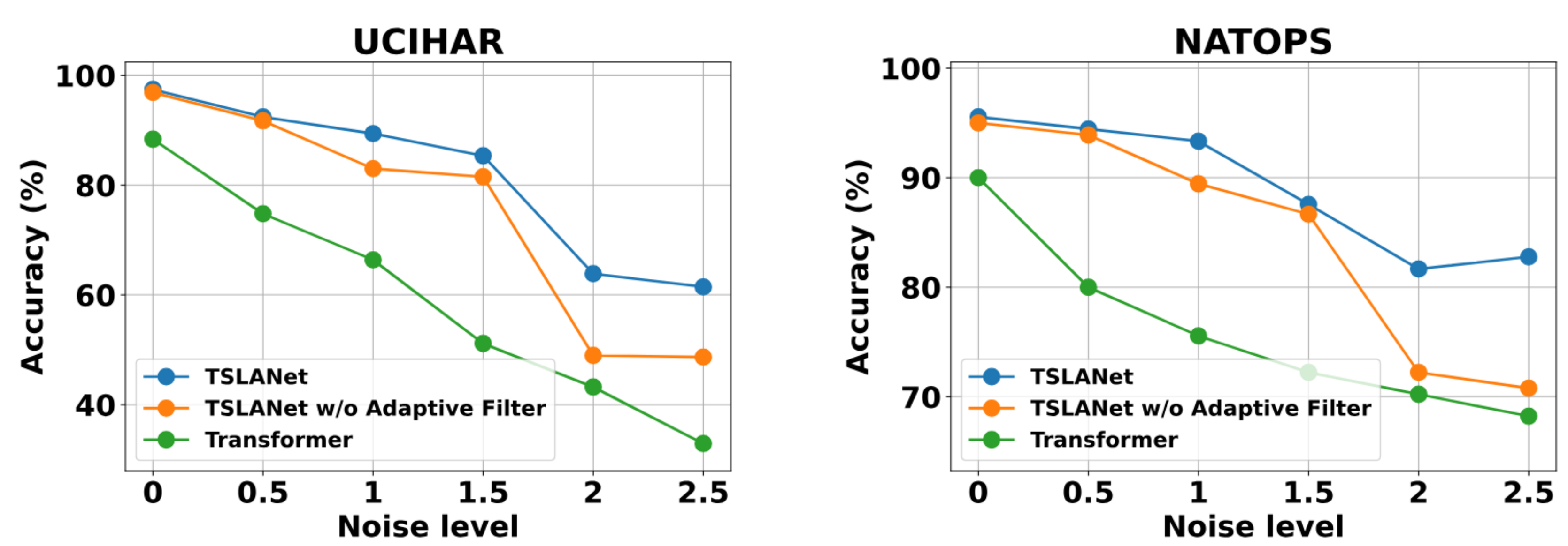
Table 2: Classification results in different datasets. Results are averaged across each subset of datasets. Results are in terms of accuracy (as %). **Blue**: best results, Purple: second best.

Methods	TSLANet (*Ours*)	GPT4TS (2023)	TimesNet (2023)	ROCKET (2020)	Crossformer (2023)	PatchTST (2023)	MLP (2023)	TS-TCC (2021)	TS2VEC (2022)
UCR repository (85 datasets)	**83.18**	61.58	65.27	81.42	73.47	71.84	69.68	75.07	81.42
UEA repository (26 datasets)	**72.73**	58.51	66.55	68.79	66.84	69.13	65.81	69.38	59.62
Biomedical signals (2 datasets)	90.24	87.04	87.10	87.20	70.82	83.87	70.63	**92.25**	86.31
Human activity recognition (3 datasets)	**97.46**	92.71	91.51	96.44	77.55	94.87	56.69	97.16	95.70
Average	**85.90**	74.96	77.61	83.46	72.17	79.93	65.70	83.55	80.76

Table 3: Multivariate forecasting results with prediction lengths $\in \{96, 192, 336, 720\}$. Results are averaged from all prediction lengths. *Avg means further averaged by subsets*. **Blue**: best results, Purple: second best.

Models	TSLANet (*Ours*)		Time-LLM (2024)		iTransformer (2024)		PatchTST (2023)		Crossformer (2023)		FEDformer (2022)		Autoformer (2021b)		RLinear (2023)		Dlinear (2023)		TimesNet (2023)		GPT4TS (2023)		SCINet (2022)	
Metric	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ECL	0.165	0.257	**0.158**	**0.252**	0.178	0.270	0.167	0.259	0.244	0.334	0.214	0.327	0.227	0.338	0.219	0.298	0.166	0.263	0.192	0.295	0.167	0.263	0.268	0.365
ETT (Avg)	0.337	0.377	**0.330**	**0.372**	0.383	0.399	0.347	0.378	0.685	0.578	0.408	0.428	0.465	0.459	0.380	0.392	0.369	0.398	0.391	0.404	0.350	0.382	0.689	0.597
Exchange	0.369	0.404	-	-	0.360	0.403	0.367	0.404	0.940	0.707	0.519	0.429	0.613	0.539	0.378	0.417	**0.297**	**0.378**	0.416	0.443	0.370	0.406	0.750	0.626
Traffic	0.396	0.271	**0.388**	**0.264**	0.428	0.282	0.420	0.277	0.550	0.304	0.610	0.376	0.628	0.379	0.626	0.378	0.433	0.295	0.620	0.336	0.414	0.294	0.804	0.509
Weather	0.228	0.264	**0.225**	**0.257**	0.258	0.279	0.238	0.268	0.259	0.315	0.309	0.360	0.338	0.382	0.272	0.291	0.246	0.300	0.259	0.287	0.237	0.270	0.292	0.363

Robustness against Noise



(a) Robustness against noise levels on UCIHAR dataset.

(b) Robustness against noise levels on NATOPS dataset.

Parameters Efficiency

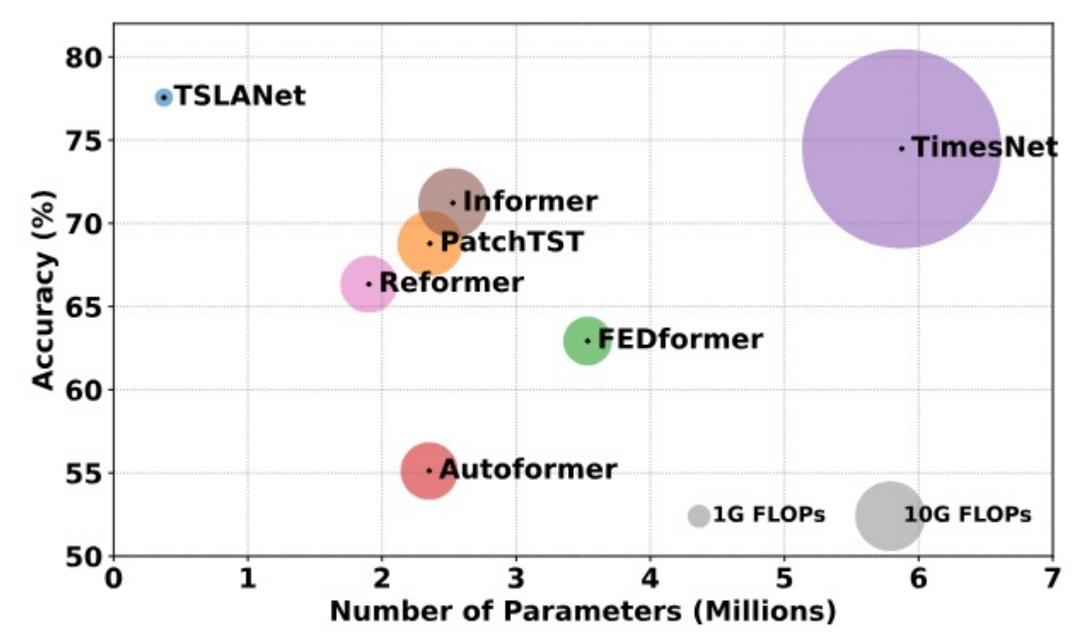


Figure 5: TSLANet vs. baselines in terms of the number of parameters and FLOPS count against the classification accuracy of the UEA Heartbeat dataset.

References

- Rao, Y. et al. “Global filter networks for image classification”. In NeurIPS, 2021.
- Wu, H. et al. “Timesnet: Temporal 2d-variation modeling for general time series analysis”. In ICLR, 2023.
- Zhou, T. “One fits all: Power general time series analysis by pretrained LM”. In NeurIPS, 2023.