Internal Global Adaptive Filtering Layer for Computer Vision

Final Project in Digital Signal Processing

Nikita Kurdiukov Pavel Tikhomirov

Skolkovo Institute of Science and Technology

December 22, 2023

Skoltech

Table of Contents

Skoltech

Problem statement: Challenge of increasing quality of ANNs' predictions

Having machine learning task and benchmarks to measure quality of solutions, we want to improve this quality by upgrading generalization ability of neural network with less costs. Possible solutions:

1. More parameters, larger models - need a lot of compute !
2. Develop efficient/task-specific modules - research direction

It is useful to add inductive bias - specific techniques/modules that make model less mutable, but task specific. Usually it allows the model to be more efficient for that exact task.

Skoltech

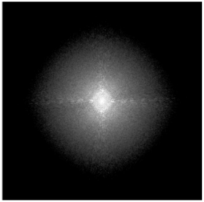
Table of Contents

Skoltech

Non-spatial domains in imaging

Fourier transform

Real Space



Fourier Space

$$\mathcal{F}(k, l) = \frac{1}{MN} \sum_{m=0}^{M-1} f[m, n] e^{-j2\pi\left(\frac{k}{M}m + \frac{l}{N}n\right)}$$

Skoltech

Non-spatial domains in imaging

Wavelet transforms

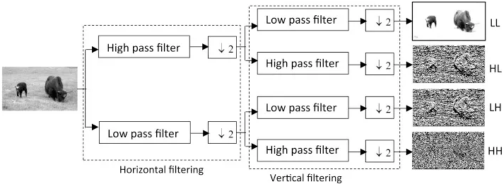


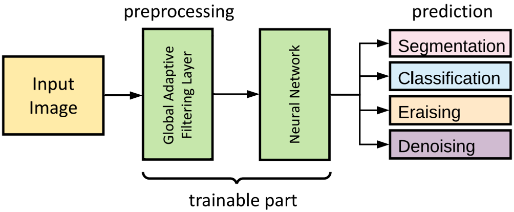
Figure: **Decomposition of an image with 2D discrete wavelet transform**. Source — *Parida et al. , 2017. Wavelet based transition region extraction for image segmentation. Future Computing and Informatics Journal.*

$$LL_{j+1} = \sum_m \sum_n h[m]h[n] \cdot f_j[2m, 2n] \qquad HL_{j+1} = \sum_m \sum_n h[m]g[n] \cdot f_j[2m, 2n+1]$$

$$LH_{j+1} = \sum_m \sum_n g[m]h[n] \cdot f_j[2m+1, 2n] \qquad HH_{j+1} = \sum_m \sum_n g[m]g[n] \cdot f_j[2m+1, 2n+1]$$

Skoltech

Global Adaptive Filtering Layer



Input: I - Initial image.
$\mathcal{F}$ - Fast Fourier operator.

$W_1, W_2, B_1, B_2 = ReLU(W_1, W_2, B_1, B_2)$;
$F = \mathcal{F}\mathrm{I}$;
$S = W_2 * \sigma(W_1 * |F| + B_1) + B_2$;
$S = S * F/|F|$;
$\mathrm{I}' = \mathcal{F}^{-1}S$;

Output: I' - Image after global frequency filtering.

The solution is to automate the search for the optimal weights in the frequency spectrum until the desired metric of a given network is maximized for each CV task

Skoltech

GAFL configurations

Some words about log and general and phase

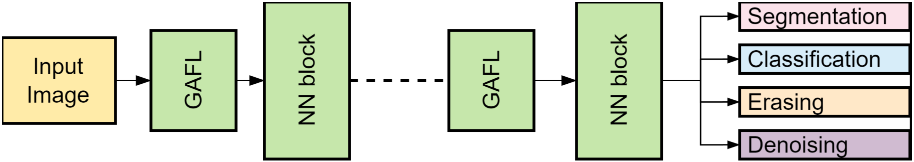
- GAFL Linear: $|FI| \leftarrow W \cdot |FI|$
- GAFL Linear log: $|FI| \leftarrow exp\left[W \cdot log(1 + |FI|)\right] - 1$
- GAFL General: $|FI| \leftarrow W_2 \cdot \sigma(W_1 \cdot |FI| + B_1) + B_2$
- GAFL General log: $|FI| \leftarrow W_2 \cdot \sigma(W_1 \cdot log(1 + |FI|) + B_1) + B_2$
- GAFL Phase: The weight matrix remains the same size as in the Linear configuration. Each weight encodes the angle for which the rotation operator is constructed for each frequency.

Skoltech

Table of Contents

Skoltech

Internal GAFLs



Filtration of frequencies not only for the input image, but also for deeper feature maps and combination of GAFLs at different depths.

Skoltech

Experiments - Segmentation

We experiment with two CV tasks on medical images:

1. Classification with ResNet on BUSI (Breast Ultrasound Images) dataset.
2. Segmentation with UNet on BUSI and BPUI (Brachial Plexus Ultrasound Images) datasets.

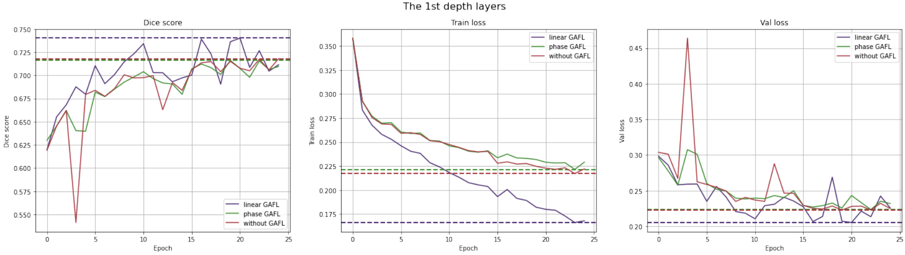
We reproduce article and compare our results against inserting GAFL only befor main neural network.

Our goal is to observe whether adding GAFLs deeply does any better in terms of metric and convergence.

Skoltech

Experiments - Segmentation

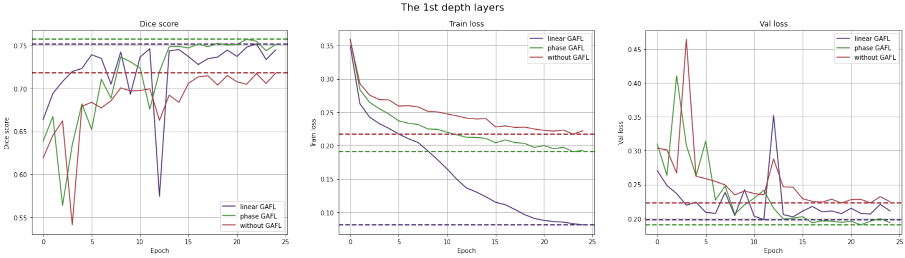
depth=0 (approvment of article results)



Skoltech

Experiments - Segmentation

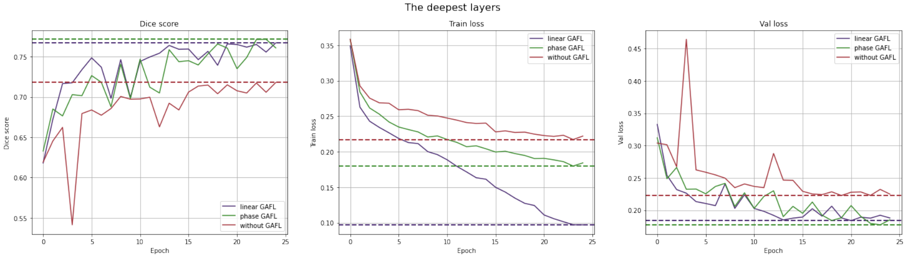
depth=1



Skoltech

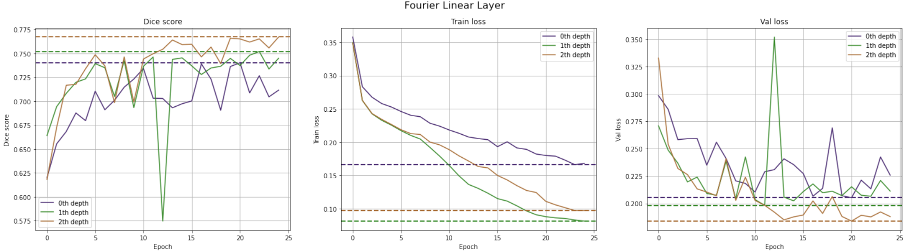
Experiments - Segmentation

depth=2



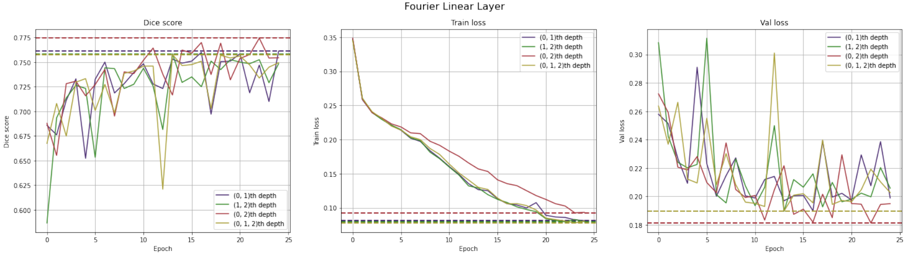
Experiments - Segmentation

GAFL among different depths



Experiments - Segmentation

Multiple GAFLs



Skoltech

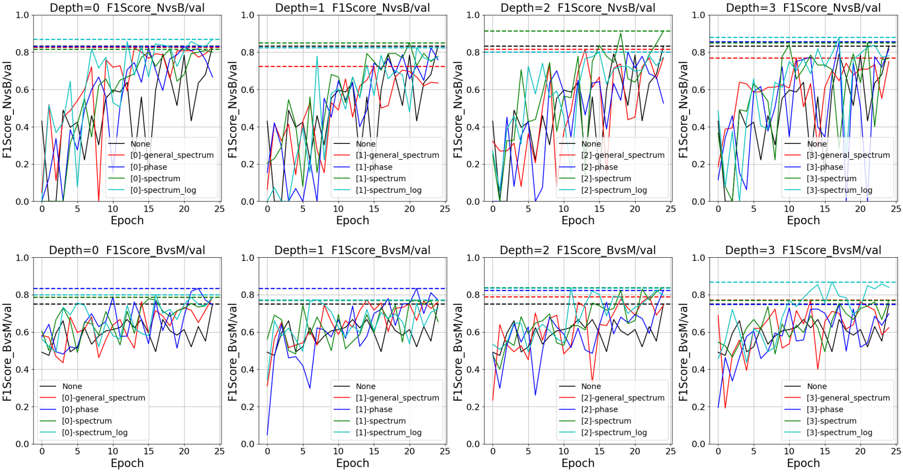
Experiments - Segmentation

Model	BPUI	BUSI
U-Net	0.64	0.72
+ 0-1 GAFL Linear	0.65	0.76
+ **0-1-2 GAFL General**	**0.66**	**0.77**
+ 0-1-2 GAFL Phase	**0.66**	0.76
+ **0-2 GAFL General**	**0.66**	**0.77**
+ **0-2 GAFL Phase**	**0.66**	**0.77**
+ 0 GAFL General	0.65	0.73
+ 0 GAFL Phase	0.64	0.72
+ 1-2 GAFL General	0.65	0.76
+ 1-2 GAFL Phase	**0.66**	0.76
+ 1 GAFL General	**0.66**	0.75
+ 2 GAFL General	**0.66**	0.76
+ **2 GAFL Phase**	**0.66**	**0.77**

Skoltech

Experiments - Classification

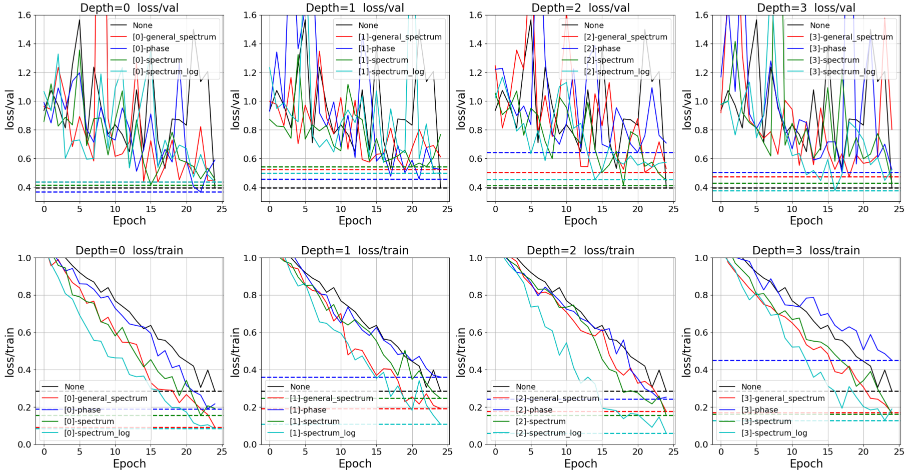
Comparison of 1 GAFL on every position vs None. Binary F1Score.



Skoltech

Experiments - Classification

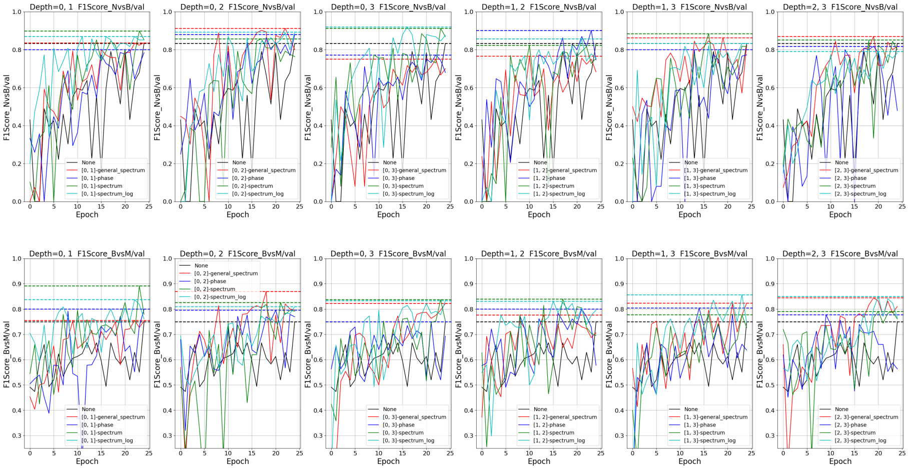
Comparison of 1 GAFL on every position vs None. Loss.



Skoltech

Experiments - Classification

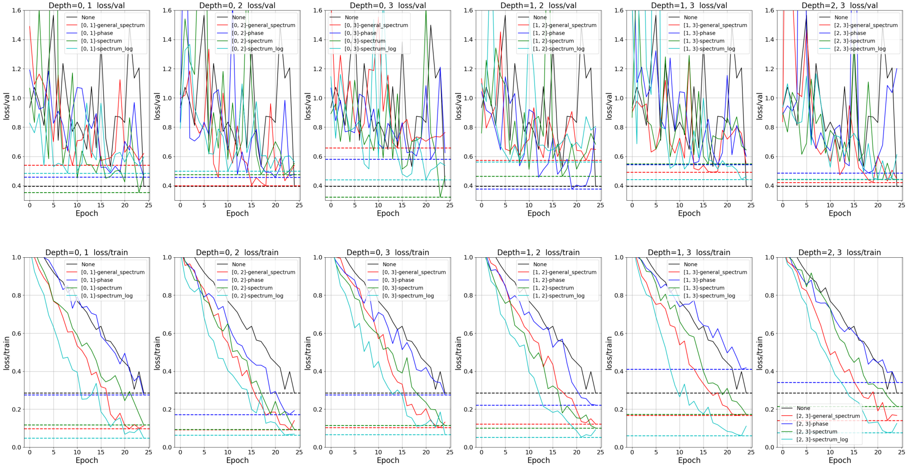
Comparison of 2 GAFLs on every position vs None. Scores.



Skoltech

Experiments - Classification

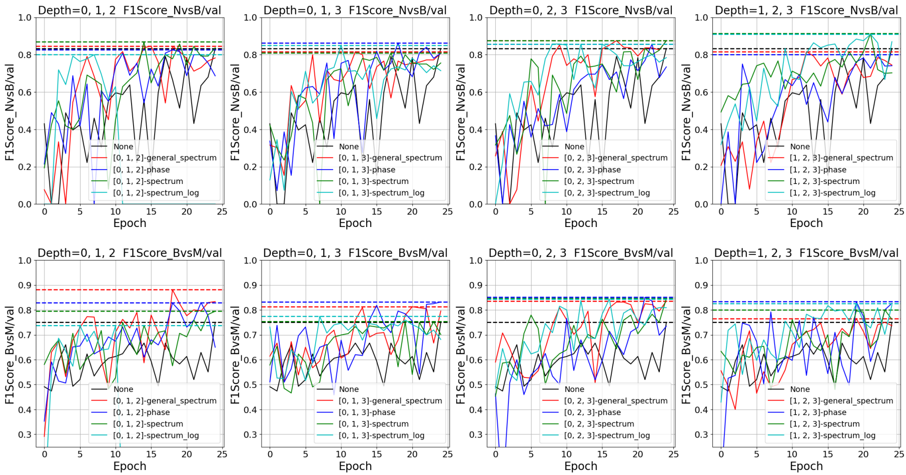
Comparison of 2 GAFLs on every position vs None. Loss.



Skoltech

Experiments - Classification

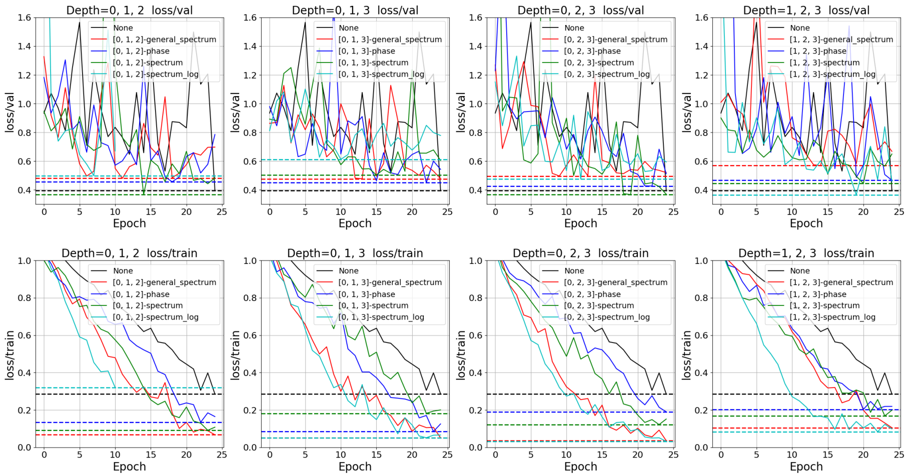
Comparison of 3 GAFLs on every position vs None. Scores.



Skoltech

Experiments - Classification

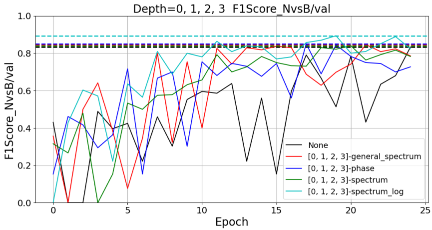
Comparison of 3 GAFLs on every position vs None. Loss.

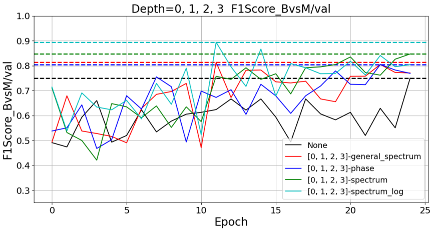


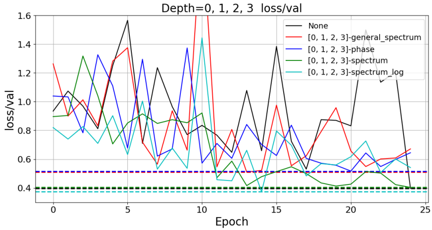
Skoltech

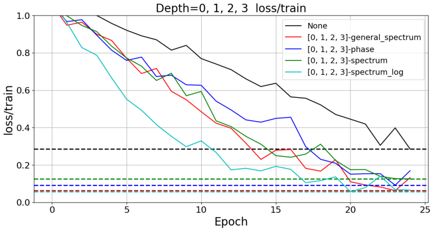
Experiments - Classification

Comparison of 4 GAFLs on every position vs None. Scores and Loss.









Experiments - Classification

Max Binary F1-score Normal vs Benign - min/mean/max values.

Model + diff. pos. GAFLs	**Linear**	**Linear log**	**General**	**Phase**
ResNet (3 blocks)	0.83/0.83/0.83	0.83/0.83/0.83	0.83/0.83/0.83	0.83/0.83/0.83
+ 1 GAFL	0.82/0.86/0.91	0.80/0.84/0.88	0.72/0.78/0.83	0.80/0.83/0.86
+ 2 GAFLs	0.82/0.87/0.91	0.79/0.86/0.92	0.75/0.83/0.91	0.77/0.83/0.90
+ 3 GAFLs	0.81/0.87/0.91	0.80/0.85/0.91	0.82/0.84/0.88	0.80/0.84/0.86
+ 4 GAFLs	0.84/0.84/0.84	0.89/0.89/0.89	0.85/0.85/0.85	0.85/0.85/0.85

Max Binary F1-score Benign vs Malignant - min/mean/max values.

Model + diff. pos. GAFLs	**Linear**	**Linear log**	**General**	**Phase**
ResNet (3 blocks)	0.75/0.75/0.75	0.75/0.75/0.75	0.75/0.75/0.75	0.75/0.75/0.75
+ 1 GAFL	0.77/0.79/0.84	0.77/0.82/0.87	0.77/0.78/0.79	0.75/0.81/0.83
+ 2 GAFLs	0.78/0.83/0.89	0.81/0.84/0.86	0.76/0.82/0.87	0.75/0.79/0.80
+ 3 GAFLs	0.75/0.80/0.85	0.74/0.80/0.84	0.77/0.82/0.88	0.83/0.84/0.85
+ 4 GAFLs	0.85/0.85/0.85	0.89/0.89/0.89	0.81/0.81/0.81	0.80/0.80/0.80

Skoltech

Table of Contents

Conclusion

- Adding more GAFL layers helps with convergence and metrics. However, it is dubious, whether the regular layer wouldn't yield the same results.
- Perhaps, big model are capable of learning Fourier Transform or useful analogies by themselves, if they need it.

Skoltech

Future directions

Practically:

- Control experiments - preserve number of parameters when adding GAFLs.
- Multiple iterations of the experiments - estimate confidence intervals.
- Longer, more precise training training.
- Analyzing learnt filters of GAFL layers.

Conceptually:

- Experiment with the impact of GAFL layer for smaller/less capable models.
- Analyze learnt filters of GAFL layers.

Acknowledgements

We would like to thank:

- **Dmitry Dylov** for proposing the project idea and providing with the links.
- **Viktor Shipitsin**, first author of the GAFL paper, for shedding more light onto his work.

Skoltech

Our team

Nikita Kurdiukov
IoT and Wireless Technologies

Pavel Tikhomirov
Data Science

Skoltech

Thank you for your attention!

https://github.com/ocenandor/GAFL

Skoltech