 doi:10.20944/preprints202207.0370.v1

Dynamic Weight Agnostic Neural Networks and Medical Microwave Radiometry (MWR) for Breast Cancer Diagnostics

Jolen Li[1], Christoforos Galazis[2], Larion Popov[3], Lev Ovchinnikov[4], Sergey Vesnin[4], Alexander Losev[3], Igor Goryanin[1,5]

1. *School of Informatics, University of Edinburgh, Edinburgh, UK*
2. *Department of Computing, Imperial College London, London, UK*
3. *Volgograd State University, Volgograd, Russian Federation*
4. *Medical Microwave Radiometry Ltd, Edinburgh, UK*
5. *Okinawa Institute of Science and Technology, Okinawa, Japan*

Abstract

Background and Objective: Medical Microwave Radiometry (MWR) is used to capture the thermal properties of internal tissues and has usages in breast cancer detection. Our goal in this paper is to improve classification performance and investigate automated neural architecture search methods.

Methods: We investigate optimizing the weights of a weight agnostic neural network using bi-population covariance matrix adaptation evolution strategy (BIPOP-CMA-ES) once the topology is found. We compare it against a weight agnostic and cascade correlation neural network.

Results: The experiments are conducted on a breast cancer dataset of 4912 patients. Our proposed weight agnostic BIPOP-CMA-ES model achieved the best performance. It obtained an F1-score of 0.9225, accuracy of 0.9219, precision of 0.9228, recall of 0.9217 and a topology of 153 connections.

Conclusions: The results are an indication of the potential of MWR utilizing a neural network-based diagnostic tool for cancer detection. By separating the tasks of topology search and weight training, we are able to improve the overall performance.

Keywords: breast cancer; passive microwave radiometry (MWR); cascaded correlation neural network (CCNN); weight agnostic neural network (WANN); CMA-ES algorithm.

 doi:10.20944/preprints202207.0370.v1

1. Introduction

Medical Microwave Radiometry (MWR) is used to obtain internal tissue temperature of the body [1]. This is done by measuring the naturally omitted thermal radiation from the tissues. Due to the device's accuracy, non-invasive, non-ionizing and cost-effective characteristics, there are already multiple clinical applications using the temperature readings to identify various conditions [1-11], such as breast cancer that is investigated here. This is feasible because the growth rate of tumors is correlated with the tissues' temperature [12, 13]. In addition, we can derive from MWR the cancer cells' reproduction rate and mutagenesis risk levels [14].

MWR is a relatively new clinical imaging technique. To speed up its adoption we suggest the development of an artificial intelligent (AI) diagnostic tool. The diagnostic tool alleviates the need of extensive clinical training, prevents increased workload and provides accurate predictions. Thus, our first objective is to improve the diagnostic accuracy of the model. Furthermore, while we are focused on breast cancer, MWR has clinical applications for various anatomical locations and conditions [1, 14]. To reduce model development time, we explore adapting automatic machine learning (AutoML) techniques for MWR data.

Previous work on AutoML for MWR explored a Cascade Correlation Neural Network (CCNN) [15]. Improvements were introduced by expanding the available pool of layers and activation functions [15, 16]. However, it was not able to outperform predefined architectures [16]. It did result in a small network optimal for edge computing and constraint hardware. Other classification models have been explored, such as deep/convolutional neural networks, support vector machines, random forests [15, 16] and a rule-based model for interpretability [16].

In summary, our contributions in the field of MWR for breast cancer detection are two-fold. First, we evaluate the application of Weight Agnostic Neural Network (WANN) [17] on MWR data and compare it against the CCNN [15]. Secondly, we improve the WANN model for MWR classification. Once the topology of the network is found, we use the BI-Population Covariance Matrix Adaptation Evolution Strategy (BIPOP-CMA-ES) [18] to optimize the weights. Using WANN and BIPOP-CMA-ES strategies we obtain state-of-the-art classification performance on MWR breast cancer data.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2022 doi:10.20944/preprints202207.0370.v1

2. Methods

2.1 Cascade Correlation Neural Network

Cascade Correlation Neural Network (CCNN) is an early neural architecture search (NAS) technique for supervised tasks [19]. A CCNN starts with a minimum sized network, input and output layers, and adds one additional layer at a time until convergence. The steps of the algorithm are [19]:

1. Initialize network topology with input and output nodes.
2. Create a pool of candidate hidden layers with different initialized weights. The hidden layer node takes as input from all previous layers. Its output is connected to the output layer nodes. Each candidate layer is trained until convergence.
3. From the pool of candidates, select and add to the network the candidate node that maximizes the magnitude of correlation between the output and target on the validation set. The added hidden layer's weights are frozen.
4. If the correlation does not improve or improves by a small margin then terminate the algorithm. Otherwise, repeat from step two.

An example after adding the 3rd hidden layer can be seen in Figure 1.

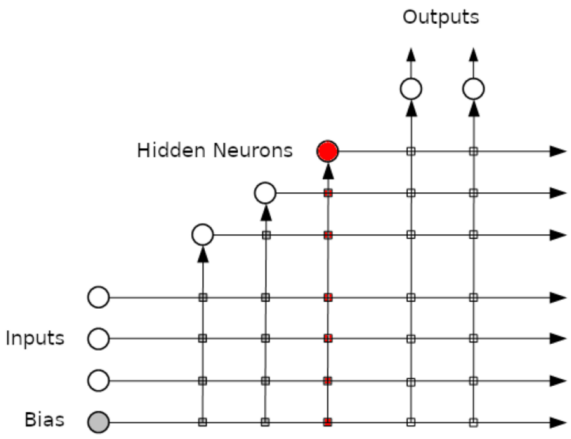

Figure 1: The cascade correlation neural network after adding three hidden layers.

The loss function we use is the cross-correlation loss and optimize using stochastic gradient descent. The optimizer's learning rate is set to $5x10^{-6}$. Each of the nodes uses a sigmoid activation function. Furthermore, the weights are sampled from a Gaussian distribution with a mean of 0 and 0.5 standard deviation, and the bias set to 0. A combination of batch normalization and dropout layers (rate of 0.5) are added after the node. The candidate node pool size is set to 30. Additionally, we reinitialize the weights of the output layers after each iteration to avoid bad local minima.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2022 doi:10.20944/preprints202207.0370.v1

2.2 Weight Agnostic Neural Network

Another NAS method is the Weight Agnostic Neural Network (WANN) approach [17]. Unlike the CCNN, the weights of the model are not trained. The idea behind this is to find architectures that have inductive biases and can perform well in their given task without training.

Inspired by genetic evolution, WANN starts with a small population of networks that consist of sparsely connected input and output layers. Following, a series of fixed shared value weights ({-2, -1.5, -1, -0.5, +0.5, +1, +1.5 +2}) are used to evaluate the performance of each topology based on the geometric mean. The topologies are ranked based on three criteria: mean performance across all fixed shared weights, best performance between any of the fixed shared weights and smallest number of connections. We use the Non-dominated Sorting Genetic Algorithm II (NSGA-II) [20] to sort the network topologies. Then, the highest-ranking topologies are selected for the next step using tournament algorithm [21].

Following, the selected topologies are altered to generate the next generation of population. We used three mutation operations. First, a node can be inserted between two connected nodes. Secondly, a new connection can be added between two existing nodes. Finally, the activation function of a node can be changed according to Table 1. This process of evaluating, ranking and generating new population is repeated until there is no longer improvement.

Table 1: The pool of activation functions the weight agnostic neural network samples from.

Name	Equation	Range
Linear	$f(x) = x$ (1)	$(-\infty, \infty)$
Binary step	$f(x) = \{1 \ for \ x \geq 0 \ 0 \ for \ x < 0$ (2)	$\{0, 1\}$
Sin	$f(x) = sin\,(x)$ (3)	$[-1, 1]$
Cosine	$f(x) = cos\,(x)$ (4)	$[-1, 1]$
Sigmoid	$f(x) = \frac{1}{1+e^{-x}}$ (5)	$(0, 1)$
Gaussian	$f(x) = e^{-x^2}$ (6)	(0,1]
TanH	$f(x) = tanh\,(x)$ (7)	$(-1,1)$
Inverse	$f(x) = -x$ (8)	$(-\infty, \infty)$
Absolute Value	$f(x) = abs(x)$ (9)	$[0, \infty)$
ReLu	$f(x) = \{x \ for \ x \geq 0 \ 0 \ for \ x < 0$ (10)	$[0, \infty)$
Squared	$f(x) = x^2$ (11)	$[0, \infty)$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2022 doi:10.20944/preprints202207.0370.v1

The hyperparameters for the WANN model are summarized in Table 2. We search 200 generations with each having a population size of 200. The initial connections between input and output layers are set to 0.2 rate. For topology variation, we set the likelihood of activation function change to 0.5, node addition to 0.25 and new node connection to 0.25. Finally, we set the size of the tournament algorithm to 4.

Table 2. The hyperparameter selection for the weight agnostic neural network

Hyperparameter	Value
Generations	200
Population Size	200
Change Activation Probability (%)	50
Add Node Probability (%)	25
Add Connection Probability (%)	25
Initial Active Connections (%)	20
Tournament Size	4

2.3 Weight Agnostic Neural Network BIPOP-CMA-ES

With MWR breast data we can determine patterns between the points to identify high-risk patients. Likewise, the WANN topology acts as a prior to identify similar or new properties. However, there are more subtle cases to distinguish between low- and high-risk, such as slow tumor growth rate. This can be achieved through weight optimization once the topology is determined.

While WANN performs better than chance in most cases it is not able to outperform weight-trained models [17]. However, optimizing the weights of a WANN model is not straight forward due to the difficult gradient traversal from using multiple activation functions [17]. Alternatively, the weights can be optimized through a randomized search method such as CMA-ES algorithm [17, 22]. The randomized search makes it suitable for a rigged landscape that has many bad local minima, discontinuities and noise. The steps of the CMA-ES algorithm are summarized in Figure 2.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2022 doi:10.20944/preprints202207.0370.v1

Step 1: Initialize the CMA-ES Parameters
D ← no. of dimensions
λ ← Offspring population size (4.0 + 3.0 Log(D))
μ ← Parent population for next generation (floor(λ/2))
σ_{Start} ← Initial standard deviation.
c_{cov} ← Covariance learning rate

Step 2: **While** stopping criterion is not met **do**
Step 3: Update the **Covariance Matrix** $C^{(g+1)}$ (see Covariance Matrix Adaption)

$$C^{(g+1)} \leftarrow (1 - c_{cov})C^{g} + \frac{c_{cov}}{\mu_{cov}} P_c^{(g+1)} P_c^{(g+1)^T} + c_{cov}\left(1 - \frac{1}{\mu_{cov}}\right) \times \sum_{i=1}^{\times} w_i \left(\frac{X_{1:\lambda}^{(g+1)} - M^{g}}{\sigma^{(g)}}\right)\left(\frac{X_{1:\lambda}^{(g+1)} - M^{g}}{\sigma^{(g)}}\right)^{T}$$

Step 4: Update the Step Size σ^{g} (see Step Size adaption)

$$\sigma_{g+1} \leftarrow \sigma_g \times \exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{||P_\sigma||}{E\,N(0,I)} - 1\right)\right)$$

Step 5: Generate Sample Population for generation g+1

$$x_k^{(g+1)} \sim N\left(M^{(g)}, (\sigma^{(g)})^2\, C^{(g)}\right) \quad \text{for k} = 1,\ldots,\lambda$$

Step 6: Update the mean for generation g+1

$$m^{(g+1)} \leftarrow \sum_{i=1}^{\mu} w_i x_{i:\lambda}^{(g+1)}$$

Step 7: Update best ever solution
Step 8: End **While**

Figure 2. The pseudo code of the CMA-ES algorithm [23].

We use a variant of the CMA-ES algorithm, the BI-Population Covariance Matrix Adaptation Evolution Strategy (BIPOP-CMA-ES) which uses a variable population size, as we found it to perform better [18]. It initially starts with a small population size, set to 50, and doubles after each restart. Additionally, to speed up convergence, we fine-tune the initial single shared weight of the model by linearly evaluating values between -2 and 2. Finally, we use cross-entropy loss to find best fit.

3. Results

3.1 Data

The MWR breast cancer data is captured using the RTM-01-RES (http://www.mmwr.co.uk) device across multiple clinics. Classification of 4912 patients as low- (4377) or high-risk (535) for breast cancer was done by clinicians using a combination of MWR, mammography and/or biopsy data. Subsequently, we class-balance split our data to train, validation and test sets using 60%, 20% and 20% respectively.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2022 doi:10.20944/preprints202207.0370.v1

The MWR data consists of temperature readings of the skin and at a depth of 3-5 cm at each point, resulting in total 44 points. At each gland, a point is recorded on the nipple and eight equidistant points around the nipple. Additionally, two reference points are captured just below the chest and two at the axillary regions. The positions described are shown in detail in Figure 3.

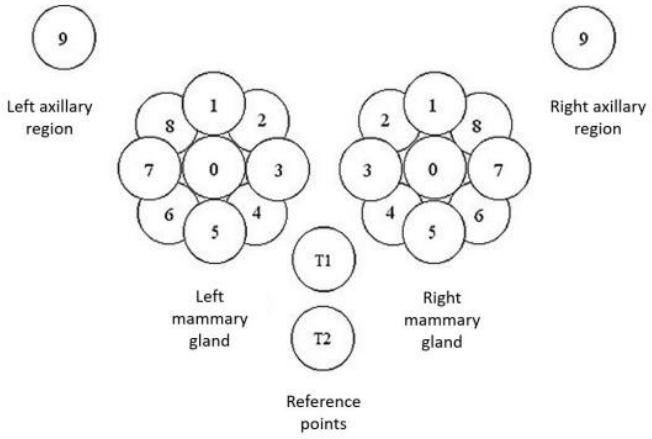

Figure 3. Capture positions for each mammary gland (0-8) and at the axillary point (9). Additional T1 and T2 reference points under the chest. Each position is captured on the skin and at a depth of 3-5 cm.

3.2 Experimental Results

The best performing model on the test set on all metrics is the proposed WANN BIPOP-CMA-ES. It obtains 0.9225 F1-score, 0.9212 accuracy, 0.9228 precision, 09217 recall and 153 total connections. The worst performing is the WANN model without weight optimization. The summary of the results is shown in Table 3. While the WANN BIPOP-CMA-ES obtained the best performance, it was the slowest to converge. This is primarily due to the BIPOP-CMA-ES algorithm. It requires a total of 10^6 function evaluations before converging on the validation set. The accuracy and loss during training can be seen in Figure 4.

Table 3. The summary of the results of the three models (CNN, WANN and WANN BIPOP-CMA-ES).

Model	F1-score	Accuracy	Precision	Recall	Connections
CCNN	0.8252	0.8309	0.8095	0.8415	662
WANN	0.6546	0.6947	0.7363	0.5892	153
WANN BIPOP-CMA-ES	0.9225	0.9212	0.9228	0.9217	153

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2022 doi:10.20944/preprints202207.0370.v1

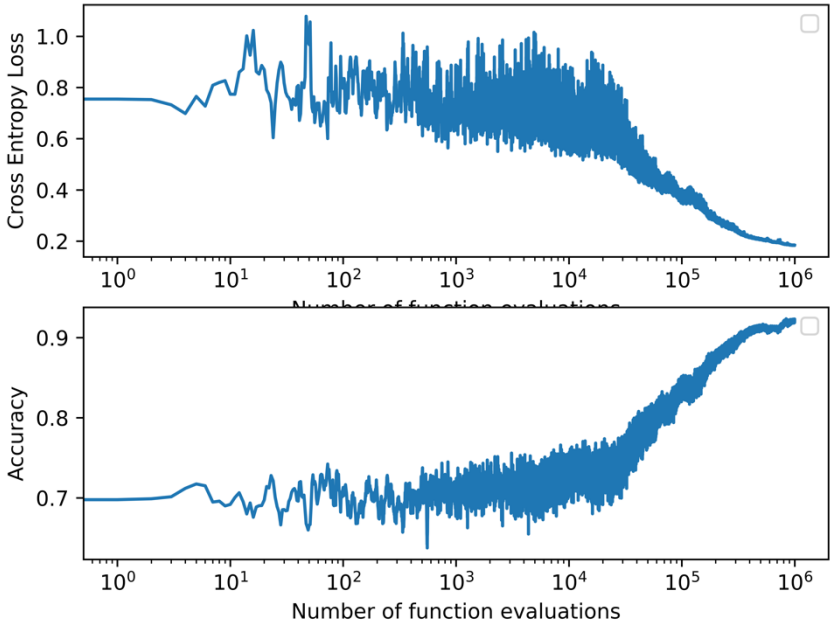

Figure 4. Cross-entropy loss and accuracy of the BIPOP-CMA-ES optimization algorithm during training of the best WANN topology.

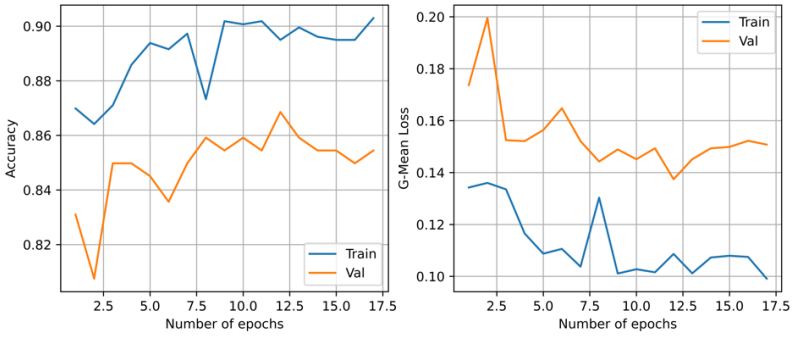

Figure 5. The geometric mean loss and accuracy of the training and validation set of the cascade correlation neural network.

The architecture and optimized weights obtained from the CCNN has a good performance on the F1-score with a value of 0.8252. It required a total of 17 iterations before terminating due to the validation loss not decreasing given a patience of 5 iterations. The accuracy and loss during training can be viewed in Figure 5. The model used 662 connections before converging.

Additionally, we evaluate the robustness of the WANN's generated topology to weight changes. Specifically, we evaluate the average performance across 10 runs of the network when using fixed, randomly generated and a fine-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2022 doi:10.20944/preprints202207.0370.v1

tuned shared weight. In Table 4, we show a summary of the results on the test set. By tuning the weights, we are able to obtain the best performance. In contrast, using random weights we obtain the worst performance.

Table 4. Performance of WANN on the test set using different weight schemes.

WANN	F1-score	Accuracy	Precision	Recall
Random weight	0.5209	0.5591	0.5571	0.4892
Shared weight	0.5979	0.6628	0.7212	0.5105
Tuned shared weight	0.6546	0.6947	0.7363	0.5892

4. Discussion

We have found that the WANN BIPOP-CMA-ES outperforms the other models which in turn has very promising clinical applications. Specifically, it obtains test accuracy of 0.9212, F1-score of 0.9225, precision of 0.9228 and recall of 0.9217. However, we are only able to obtain good performance using a random search evolution strategy rather than the simpler gradient descent-based optimizer for the weights.

AutoML, and specifically WANN BIPOP-CMA-ES, has clear benefits but is underutilized in the healthcare domain [24]. First, the generated small network allows it to be deployed on low-end devices and existing clinical hardware. This is particularly important for accessibility to low- and mid-income countries. Furthermore, clinicians can more effectively contribute their knowledge and experience to improve the diagnostic tool, as technical barriers are reduced. Finally, more specific to MWR, we can quickly adapt the model to other anatomical locations and conditions.

Regarding AutoML for MWR, we have explored two network architecture search methods, CCNN and WANN. For future work, we will look at improving our model by searching an optimal loss function [25], utilizing one-short learning search [26] and including an automatic hyperparameter search [27]. We will also compare against other NAS methods such as reinforcement learning-based searching methods [28, 29], ensemble methods [30] and transfer learning [31].

The mutation operations of WANN only increase the complexity of the network topology. We will expand it to include operations such as deleting nodes and connections so there is more flexibility. In addition, there is no crossover mutation operation in the WANN model, which will reduce population diversity. Hence, we will explore different crossover mutation operations [32, 33] and restart techniques to increase model performance [34, 35].

References

1. I. Goryanin, S. Karbainov, O. Shevelev, A. Tarakanov, K. Redpath, S. Vesnin and Y.Ivanov, Passive microwave radiometry in biomedical studies. Drug Discov. Today. 25 (2020) 757–763. DOI: 10.1016/j.drudis.2020.01.016.
2. Levshinskii, V., C. Galazis, A. Losev, T. Zamechnik, T. Kharybina, S. Vesnin, and I. Goryanin. Using AI and Passive medical Radiometry for diagnostics (MWR) of venous diseases. Computer Methods and Programs in Biomedicine (2021): 106611.
3. Shevelev, Oleg, Marina Petrova, Andrey Smolensky, Batyr Osmonov, Samatbek Toimatov, Tatyana Kharybina, Sergey Karbainov et al. Using medical microwave radiometry for brain temperature measurements. Drug discovery today (2021).
4. Rodrigues, Dario B., Paul R. Stauffer, Pedro JS Pereira, and Paolo F. Maccarini. Microwave radiometry for noninvasive monitoring of brain temperature. In Emerging Electromagnetic Technologies for Brain Diseases Diagnostics, Monitoring and Therapy, pp. 87-127. Springer, Cham, 2018.
5. Tarakanov, A. V., A. A. Tarakanov, S. Vesnin, V. V. Efremov, I. Goryanin, and N. Roberts. Microwave Radiometry (MWR) temperature measurement is related to symptom severity in patients with Low Back Pain (LBP). Journal of Bodywork and Movement Therapies 26 (2021): 548-552.
6. O. Batyr, L. Ovchinnikov, C. Galazis, B. Emilov, M. Karaibragimov, M. Seitov, Sergey Vesnin et al. Passive microwave radiometry for the diagnosis of coronavirus disease 2019 lung complications in kyrgyzstan. Diagnostics 11, no. 2 (2021): 259.
7. Kaprin, A. D., A. A. Kostin, M. I. Andryukhin, K. V. Ivanenko, S. V. Popov, P. V. Shegai, D. P. Kruglov, F. Sh Mangutov, V. Yu Leushin, and S. V. Agasieva. Microwave radiometry in the diagnosis of various urological diseases. Biomedical Engineering 53, no. 2 (2019): 87-91.
8. Zamechnik, Tatyana Vladimirovna, Alexander Georgievich Losev, and Alexander Yuryevich Petrenko. Guided classifier in the diagnosis of breast cancer according to microwave radiothermometry. Mathematical Physics and Computer Modeling 22, no. 3 (2019).
9. Losev, Alexander G., Denis A. Medevedev, and Andrey V. Svetlov. Neural networks in diagnosis of breast cancer. In Institute of Scientific Communications Conference, pp. 220-227. Springer, Cham, 2020.
10. Bardati, Fernando, and Santina Iudicello. Modeling the visibility of breast malignancy by a microwave radiometer. IEEE Transactions on Biomedical Engineering 55, no. 1 (2007): 214-221.
11. Levick, A., D. Land, and Jeffrey Hand. Validation of microwave radiometry for measuring the internal temperature profile of human tissue. Measurement Science and Technology 22, no. 6 (2011): 065801.
12. M. Gautherie, Temperature and blood flow patterns in breast cancer during natural evolution and following radiotherapy. Progr. Clin. Biol. Res. 107 (1982) 21–64.
13. M. Gautherie and C. M. Gros, Breast thermography and cancer risk prediction. Cancer. 45(1980) 51–56.
14. S.G. Vesnin, A. K. Turnbull, J. M. Dixon, and I. Goryanin. Modern microwave thermometry for breast cancer. Journal of Molecular Imaging and Dynamics, 7 (2017) 136. DOI: 10.4172/2155-9937.1000136
15. C. Galazis, S. Vesnin and I. Goryanin, Application of artificial intelligence in microwave radiometry (MWR), in: Proceedings of the 12-th International Joint Conference on Biomedical Engineering systems and Technologies (BIOSTEC 2019), E. De Maria, Hugo Gamboa, A. Fred (Eds), Scitepress, 2019 (4), pp. 112–122.
16. Vladislav Levshinskii, Christoforos Galazis, Lev Ovchinnikov, Sergey Vesnin, Alexander Losev, and Igor Goryanin. Application of data mining and machine learning in microwave radiometry (MWR). in: A.Rogue et al. (Eds.), International Joint Conference on Biomedical Engineering Systems and Technologies, Springer Nature, 2020, pp. 265–288.
17. Adam Gaier and David Ha, Weight agnostic neural networks, in: 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada, 2019. arXiv:1906.04358 (cs. LG).
18. Ilya Loshchilov. CMA-ES with restarts for solving CEC 2013 benchmark problems, in: 2013 IEEE Congress on Evolutionary Computation, Cancun, Mexico, IEEE, 2013, pp. 369–376. DOI:10.1109/CEC.2013.6557593.
19. S. E Fahlman, and C.Lebiere, The cascade-correlation learning architecture, in: Carnegie-mellon University Pittsburgh PA School of computer science, 1990, pp. 524- 532.
20. K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Transactions on Evolutionary Computation. 6 (2002) 182–197. Doi:10.1109/4235.996017
21. Xuan Fang, Tamas Tettamanti and Arthur Piazzi, Online calibration of microscopic road traffic simulator, in: SAMI 2020. IEEE 18th World Symposium on Applied Machine Intelligence and Informatics, 2020, Herl'any, Slovakia, 2020, pp. 275–280. DOI:10.1109/SAMI48414.2020.9108744.

22. Nikolaus Hansen, Sibylle D M¨uller, and Petros Koumoutsakos. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). Evolutionary computation. 11 (2003) 1–18.
23. Srikanth Reddy Gagganapalli, Implementation and evaluation of CMA-ES algorithm, Fargo, North Dakota, 2015.
24. Waring, Jonathan, Charlotta Lindvall, and Renato Umeton. Automated machine learning: Review of the state-of-the-art and opportunities for healthcare. Artificial Intelligence in Medicine 104 (2020): 101822.
25. Li, Chuming, Xin Yuan, Chen Lin, Minghao Guo, Wei Wu, Junjie Yan, and Wanli Ouyang. Am-lfs: Automl for loss function search. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 8410-8419. 2019.
26. Li, Xiang, Chen Lin, Chuming Li, Ming Sun, Wei Wu, Junjie Yan, and Wanli Ouyang. Improving one-shot nas by suppressing the posterior fading. In Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition, pp. 13836-13845. 2020.
27. Zela, Arber, Aaron Klein, Stefan Falkner, and Frank Hutter. Towards automated deep learning: Efficient joint neural architecture and hyperparameter search. arXiv preprint arXiv:1807.06906 (2018).
28. Zoph, Barret, and Quoc V. Le. Neural architecture search with reinforcement learning. arXiv preprint arXiv:1611.01578 (2016).
29. Baker, Bowen, Otkrist Gupta, Nikhil Naik, and Ramesh Raskar. Designing neural network architectures using reinforcement learning. arXiv preprint arXiv:1611.02167 (2016).
30. Zaidi, Sheheryar, Arber Zela, Thomas Elsken, Chris Holmes, Frank Hutter, and Yee Whye Teh. Neural ensemble search for performant and calibrated predictions. arXiv preprint arXiv:2006.08573 2 (2020): 3.
31. Zhao, Yiyang, Linnan Wang, Yuandong Tian, Rodrigo Fonseca, and Tian Guo. Few-shot neural architecture search. In International Conference on Machine Learning, pp. 12707-12718. PMLR, 2021.
32. Lu, Zhichao, Ian Whalen, Vishnu Boddeti, Yashesh Dhebar, Kalyanmoy Deb, Erik Goodman, and Wolfgang Banzhaf. Nsga-net: neural architecture search using multi-objective genetic algorithm. In Proceedings of the genetic and evolutionary computation conference, pp. 419-427. 2019.
33. Hassanat, Ahmad, Khalid Almohammadi, Esra Alkafaween, Eman Abunawas, Awni Hammouri, and V. B. Prasath. Choosing mutation and crossover ratios for genetic algorithms—a review with a new dynamic approach. Information 10, no. 12 (2019): 390.
34. Molina, Daniel, Antonio LaTorre, and Francisco Herrera. An insight into bio-inspired and evolutionary algorithms for global optimization: review, analysis, and lessons learnt over a decade of competitions. Cognitive Computation 10, no. 4 (2018): 517-544.
35. Yu, Haibo, Ying Tan, Chaoli Sun, and Jianchao Zeng. A generation-based optimal restart strategy for surrogate-assisted social learning particle swarm optimization. Knowledge-Based Systems 163 (2019): 14-25.