Learning Methods and Parameters used in Neural Architecture Search for Image Classification

S. Jeevidha
Research Scholar
Department of Information Technology
Puducherry Technological University

S. Saraswathi
Professor
Department of Information Technology
Puducherry Technological University

ABSTRACT
NAS -Neural architecture search is a part of the deep learning mechanism that is widely adopted in a variety of autonomous systems, object detection, image classification, and gaming techniques. The entire performance of the Neural architecture depends only upon the learning methods and parameters, Earlier the data scientist chooses a neural architecture and the enhancement parameters based upon the model family and constructs an ensemble of models and they will perform training with the domain knowledge. In these new decades, the end-users who are familiar with the data at hand do not necessarily have the experience in designing Neural Architecture. For any image classification process, the best learning methods and parameters with less iteration should be chosen in order to reduce time and computing resources. As a result, there is increasing interest in enhancing the parameters and learning methods of Neural architecture search. Learning methods and parameters are used to control the learning rate that occurred in thrust and build the model based on type in which it gives more accuracy This paper provides a review of learning methods and parameters used in NAS techniques such as learning rates, batch normalization, batch sizes, data augmentation and evaluation parameters implemented in various Image classification techniques.

Keywords
Learning Methods, Neural Architecture Search, Enhancement Parameters, Convolutional Neural networks, Image Classification

1. INTRODUCTION
Deep learning models are difficult to train, understand, and are frequently so complicated that they can only be trained on very large GPU clusters. [11]There is a change in neural architecture Search that incorporates the existing benchmark datasets, although there are some effects between both the best chosen architecture and the best chosen enhancement parameter configuration. [8] Existing methods for dealing with translation, language processing, classification of images, and speech recognition problems.[9].Automatically extracting Textual and image features on data are widely credited with their success. [3] The manual feature engineering for unstructured- data approach is currently being replaced by network components and architectures for deep learning approaches, as seen in this study. While changes to architecture enhance the accuracy of deep learning methods, finding the right architecture consumes time, tedious, and error-prone effort in and of itself. However, even with identical precision, both of these search techniques cost hundreds of GPU hours in their separate computations and had varied computational efficiency on different hardware architectures.[1] As a result, many of the ensuing studies focus on ways for the order to reduce computational strain. The successful approach in this line of research takes the use of the notion of reusing learned model parameters, which is important to note[18][6]. The fact that many excellent CNN models are expertly created over a prolonged period to accomplish a satisfactory outcome, and As a result, normal developers are unable to design a good architecture, which necessitates creating new CNNs manually[24] Automated core network searches have mostly relied on neuro-evolution to build a network design, which mimics the evolutionary biology process in nature, for the past many years[29]. The current NAS research is divided into three groups:, NAS space definition, Model evaluation optimizing methods and NAS search algorithm[4][22] are described. (see Fig 1.).The NAS search space specifies the set of steps and connections to cast network architecture and learns new topology to achieve best performance on a certain task[37]. Representing a layer in a network requires specialized knowledge in network serialization because each process is linked to specific layer-specific general parameters number of filters involved, anchor point, filter-height, filter-width, stride-height, and stride-width.[6] .In cell-based search, the same cell can be replicated multiple times with several computations. By using sequential layer operations, it needs a high GPU, or search space should be restricted to enhance the search. [32].

The cell-based architecture enables transferability between datasets. By altering the number of cell repetitions, one can easily balance the load of the model size.[13] It has a general cell in the input, output feature maps which have the same dimension. In the reduction cell, The output feature map has the width and height reduced by one-half and network motifs can also be used to reduce search space, and the drop path technique used enable transferability between the datasets.[10] [32].In hierarchical NAS replication of complex structure can be obtained, Tree structure search space can be obtained for allocation and merge scheme. The representation of memory is made up of several memory blocks that are accessed and written[16].The goal of each layer operation is to read data from a group of memory blocks, compute the data, and then write the data to another subset of memory blocks. [8]. Since creating a search space computation involves a number of layer-specific parameters, including layer inputs, depth dimensions, and skip connections, these associations are hardcoded [31].In recent years, the approaches used construct search spaces are macro search space, micro and hierarchical search space[23].A population of child networks is selected by NAS search algorithms. [2]Performance evaluation is done among the child model as a reward parameter and it is used to learn high-performance architecture. It is useful in hyperparameter search.[15] In reinforcement learning Human Designs Search space-agent (controller, RNN) outputs a string that describes a neural network configuration in the designed space. The output configuration is trained and evaluated on the given task. [37].The reward is recorded and the RL agent is trained to produce configurations that maximize the expected

reward. This process is repeated from RL Agent.[27][20] In Bayesian optimization, The attainment of learning function for a surrogate model for guiding the evaluation parameter. Gaussian process-based approach for NAS in MLP&CNN was implemented in BO technique to calculate distance metric to navigate search space. The few-shot architecture was also proposed based on Bayes learning in which modeling architecture parameters using Hierarchical automatic relevance determination.[10]

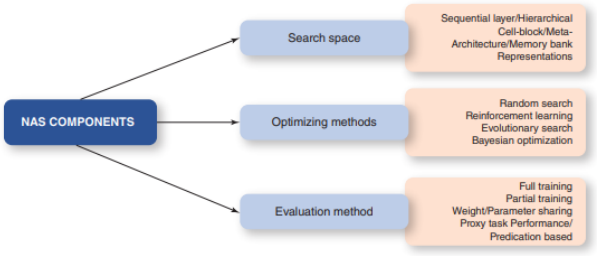

Fig 1: Components of Neural Architecture search-NAS

In Evolutionary Algorithms, number of approaches are used to optimize the architecture for deciding the genetic parameters and determining the optimum parameters, and how incisively the topologies of neural networks are represented [2][33].To increase the chance of finding new architecture and allow parallelism. Highlighted the power to vindicate diversity in genetic algorithms,**.** This method predicts the performance and number of enhancement parameters in a single run, for quantifiability, high veracity, and low resource consumption **G**enetic algorithms Given an environment, it aims to replicate how populations acquire in nature to optimize subroutines and evaluate the fitness of each solution Select the population fittest to create offspring Create offspring and mutations, recapitulate the process with the new propagations of solutions.[2] by having the overview neural architecture and can confirm that how learning methods and parameters are important for classification techniques. There are many learning methods and parameters summarized in this paper such as convolution auto encoders, transfer learning methods, early stopping, Hand crafted and SGD.

2. LITERATURE SURVEY

In research fields, such as supervising in and IoT applications .Image classification intends to extract features from an original image and convert them into useful features that are lower in dimension and have less noise. [18]The Random Forest (SVM) or other supervised learning algorithms are used to solve classification issues or train classifiers; nevertheless, these methods have limitations: i.e. they have a low generalization rate and high cost[18]. It is very essential to break through an efficient and extremely precise classification algorithm. Classification of images depend on two requisites that is feature extraction from images and the utilization of the extracted features to classify the category curbed in the images[14]. In the CNN layer, the extracted features from the image are fed into the classifier as input. After passing the image through successive layers such as convolutional, nonlinear, pooling, and fully connected layers, the output is obtained [19].The image contains matrix and pixel values are filled in. At the upper left of the image, and read the input window matrix. As a result, the software selects a smaller window matrix in that area, known as a filter. In this step, the input image is convoluted by the filter, which moves along it. In order to multiply the value, the filter multiplies it by the value of the pixel from which it comes. The final step is to add up all these multiplications.[1] Finally, an image appears. The filter moves 1 unit to the right and repeats the procedure because it has only received the image in the upper left corner. Each convolution process is followed by the addition of a nonlinear layer. Because the activation function has a nonlinear property [12]. A network would not be able to model the response variable if it lacked this characteristic (as a class label). Following the nonlinear layer, the pooling layer is applied. It performs a down sampling operation on the image's width and height.[36] .The image is reduced and there is a convolution operation where the detailed image is no longer expected for any other augmentation and processing techniques which can be compressed toward the expanded image.[12] It takes the output data from convolutional networks in a layer. "A fully- connected layer is connected to the network's end, it acquires an n-dimensional vector, where n comprises the size of classes from which the model selects the desired class"[2]. The main purposes of classification algorithms are to categorize any given data collection and forecast the result for the absolute data.

A supervised classification technique of machine learning, It provides labels for the class and identifying features for classification which increases the level of accuracy in identifying images which dataset is trained along with the human intervention. There are two general categories in supervised classification strategies

1) Building a separate model by using a trained dataset

2) Constructing a model directly with a good decision.

The most common algorithm used in the supervised classification- support vector machine, logistic regression and decision tree. In Unsupervised classification, the techniques in machine learning are calculated through automated software and it does not provide labels and finds all unknown patterns in data and finds characteristics that can be useful for categorization,

In unsupervised classification algorithm is trained using unlabeled data. A) Clustering B) Association. Unsupervised learning is used to explore the information structure and identify distinct patterns and extract valuable insights and implement the operation to increase efficiency The most commonly used algorithm is k means- clustering, hierarchical clustering, apriori algorithm. The neural network used in CNN is also called shift-invariant artificial neural networks based on their weight shared mechanism implemented in their structure which consists of one or more convolutional layers and it is accompanied by fully connected layers.[3] The input to the CNN layer is the height and width of the image and c is the number of channels. The convolutional layers have k filters of size sh*sw*sc which may be lesser than dimension of the input image or the no of channels vary for each kernel [17] which is mentioned in the following Table1.

Table 1. Parameters used in Convolutional Neural Network

Parameters	**Description**
h*w*c	Height/weight/channel input of the convolutional layer

sh*sw*sc	Height/weight/channel input of the filter
S	Stride size

The size of the filter gives locally connected structure of the model which is convolved with the image in order to produce a feature map which is described in the above Table1.Parameters used in Convolutional Neural network. Layer by layer, CNNs ascertains the non-linear mapping between features and classes, extracting some higher level features from lower-level data gained in previous levels.[35] Image classification is a technique used in a variety of fields to classify data analysis of data collected from sensing devices, characteristics of the data taken, parameters applied to data, and the number of outputs generated from data.[21] .Because convolution is a linear process, non-linearity must be added to network for the CNN to appropriately capture non-linear relationship between the input image and output features. This is accomplished by using an activation layer with an element-wise activation function.[35] .The convolution procedure fixes the problem by lowering the no of parameters, enabling the network depth to grow while using fewer. [14].Normalization layers find things simpler to resolve the network training optimization problem by normalizing the inputs and gradients that propagate along with a network. Batch normalization layers are used between convolution layer and nonlinearities to increase the speed in network training and lower the susceptibility to initialization of the network. the output of neuron clusters in one layer to create a single cell in the next layer in the pool layer [6].

3. COMPARATIVE ANALYSIS

The comparison of various learning methods and enhancement parameters is given in the following Table 2. The search space in image classification includes the primitive operations used at each searching phase, as well as the previous anchor architecture is used to represent the outer network.[27].In the search strategy, it explains how to look for information in a given area. In the paper [26] Swarm optimization technique was adopted by using convolution auto-encoders learning methods and size of code, no of layers and loss function was considered to automatically discover the optimal architecture for traffic classification. Parameter such as Learning rate, mini batch size and Transfer learning method is adopted for fundus image classification.[5]. Hence by using the above parameters the accuracy has been increased. In the images available in Freiburg dataset [7] Early stopping learning method is used and epoch size and layer size parameters are considered which focus on detection. The target of Neural Architecture Search is to find architectures that can predict unknown data well. To speed up using thousands of GPUs and massively parallelize the whole search/training process. 2. Using Evaluation Metric as the reward and using proxy datasets function. [13]NAS is just a branch of AutoML (auto-machine learning) that involves a vast area with hyper parameter optimization and meta-learning.

4. METHODS AND DISCUSSION

From the below -mentioned Comparative study in Table 2, to validate the performance of any image classification model, the learning method and enhancement parameter need to choose wisely. The purpose of NAS is to uncover designs that perform effectively whenever handling with unknown type of data in which it splits datasets into training and validation, and search architecture on training but assessed by validation. The NAS has received a lot of attention, with the majority of it focusing on picture classification. [4]The 'architecture' of a CNN is defined as the precise assembling of layers, parameters that make up its structure. As the depth of a network's desired depth grows, the number of design options available grows dramatically. The above-mentioned network topologies were mostly created by hand, with expert knowledge used to tune parameters and enhancement parameters to produce the best possible outcomes.[13].Suppose take an example for parameter sharing of models performing classification task with same set of classes [18] but with different type of input distributions where x and y are the two models with different input parameters I1 & I2

$$Model\ x=I1(x) \quad (1)$$

$$Model\ x=I2(y) \quad (2)$$

The two model maps the input into two different but related outputs

$$O(x)=f'(I1(x), z) \quad (3)$$

$$O(y)=g'(I2(y), z) \quad (4)$$

From the above equation1 and 2, if the task performed during training the two different model are similar, then assume the model parameters will be close to each other for the equation 3 and 4. Thus, the parameter sharing concept is widely adopted in convolutional neural networks (CNN) because there are many statistical properties which are invariant to translation.

Therefore, for any input parameter for any type of image classification will produce a related output.

$$I1(x)= I2(y) \quad (4)$$

Table 2. Dataset/Algorithm/ Learning methods/parameters involved in various types of image classification

Reference	Algorithm/ Techniques	Dataset	Learning Method	Parameters	Merits	Demerits
[26]	Swarm optimization (PSOA)	STL-10 and Caltech 101	Convolutional AutoEncoders	Size of _code Num_of_layers Loss function	Automatically discover the optimal architecture without manual intervention	High computational time

[5]	Stochastic Gradient Descent with Momentum (SGDM)-Optimizers	Resnet18& Google net (Fundus images dataset	Transfer learning method (Depth concatenation layer)	Learning rate Mini_batch_size	Learning rate and mini-batch size are tuned	Overfitting
[7]	Cross-validation-voting (CVV) Ensemble learning	Freiburg dataset	Early stopping	Epoch-size Layer size	Focus on object detection than classification	Lack of generalization
[25]	Cascaded Ensemble learning	HAM10000 Dataset	Handcrafted features based on	Model loss No_of_epochs	High precision	Need more computational time
[34]	Deep marginal learning algorithm (DMLA)-Dark zone classification Snake Search Algorithm (SSA)-margin slices	STAND BESAR	Stochastic Gradient Descent	Learning rate Drop factor Drop period Max_epochs	Time Efficiency Classification accuracy	Targeted to only specific scenes

Expert learning is used in a vast number of machine learning applications to preprocess the data, accurately identify features, choose a suitable model, and optimize enhancement parameters. It is important to reselect the network's enhancement parameters for diverse datasets or even start from scratch.[2]It is not possible to introduce NAS directly from image classification to image feature extraction: First, the categorization task's search space differed greatly from that of segmentation; second, the concept of domain adaptation from low to high picture resolutions was unexpected.[28]. In the classification chore, feature- selection is an important out growth, robust features selection can greatly improve the classification accuracy Due to non-differentiable and non-convex properties in practice, determining the optimal number of convolutional layers and pooling layers is difficult, which is connected to neural network architecture optimization. There are three distinct types of algorithms for automatically finding the best neural network topologies.[26] 1)search space 2)search algorithm 3)Search optimization. There are also benchmark datasets for image classification tasks such as NASBENCH101, NASBENCH 201, and DARTS.A wide range of Evolutionary based Algorithms, and data processing techniques used in benchmark datasets with the parameters such as Learning rate policy, Optimizers, No of epochs, and Batch size are used.[20]Further, It can also prove the encoding strategies used in the NAS algorithm for the various types of Image classification**.** The fitness value often represents DNN performance, such as image classification task accuracy. The primary public architecture dataset for NAS Research is NASBench-101 [11]. It built a search space and used graph isomorphisms to find 423K different convolutional architectures. Then, on CIFAR-10, all of these architectures were trained, evaluated, and combined three times into a huge dataset of more than 5 million trained architectures. [30] Without large-scale processing, each architecture can query the corresponding metrics, such as test accuracy and training time, directly in the dataset NAS-Bench-201, based on cell-based technique of encoding, and was recently proposed.

In Searching of efficient architecture for each iteration

$$I(i) = \omega(n) * S = Z \quad (5)$$

Where n is the number of parameters, S is the search space, and Z is the output, and n is equal to 0....n is the no of parameters, and i is the no of iterations needed to detect new pathways. The various image classification methods use different learning rates and parameters involved are described in the Table 1. The parameters such as size of code, layer size, loss function, learning rate, epoch size. Thus the Parameter sharing in CNNs involves reusing the same weights across different parts of the input, enabling the model to detect features like edges and shapes consistently across the image. This property makes CNNs translation invariant and reduces the number of unique parameters, minimizing the risk of over fitting and computational complexity., translation Invariance: The convolution operation, using shared weights, allows for consistent feature detection regardless of location. Efficient Learning: Shared parameters reduce the number of learnable parameters, enhancing efficiency. Empirical Evidence: Models trained on different datasets with similar tasks often learn similar parameters, as seen through the closeness of their outputs.

5. SUGGESTIONS

In traditional image classification methods, a dataset is chosen and split into training and test data. The images are then subjected to various augmentation techniques and fed into the model. By increasing the size of the training and test data, a confusion matrix is developed, and model accuracy is defined. In this paper, different samples of images are considered, and

their learning methods and parameters, based on the dataset and algorithms, are studied. The results show that input parameters are correlated with any model to increase the efficiency of the image classification model. The wisely chosen algorithm, enhancement of learning methods, and calibration of the dataset need to be effectively managed to develop an accurate model.

6. CONCLUSION

This parameter sharing involves identifying various types of images using neural architecture search (NAS) with enhancement parameters. Different scenarios and several image classification approaches are discussed. Additionally, we examined various scenarios for different image classification techniques, learning methods, and parameters, as well as the benefits and drawbacks of each. Consequently, this study will assist in selecting an appropriate classification technique, learning method, and parameters from the many available options. This approach requires the coordination of multiple search algorithms that operate in the neural architecture space, weight space, and enhancement parameters, making it highly complex. Multi-objective NAS methods result in cost and time savings by evaluating multiple enhancement parameters. As a result, this research provides valuable insights into selecting the most suitable classification technique, learning method, and parameters. However, multi-objective NAS methods offer cost and time savings by efficiently evaluating multiple enhancement parameters, making them a practical choice for implementing robust image classification systems.

7. REFERENCES

[1] Achararit, P., Abdullah Hanif, M., Wicaksana Putra, R. V., Shafique, M., Hara-Azumi, Y. 2020. APNAS: Accuracy-and-performance-aware neural architecture search for neural hardware accelerators. IEEE Access, 8, 165319–165334. doi: 10.1109/ACCESS.2020.3022327

[2] Wang, X., et al. 2021. Evolutionary Algorithm-Based and Network Architecture Search-Enabled Multiobjective Traffic Classification. IEEE Access, 9, 52310-52325. doi: 10.1109/ACCESS.2021.3068267

[3] Zhu, N., Yu, Z., Kou, C. 2020. A New Deep Neural Architecture Search Pipeline for Face Recognition. IEEE Access, 8, 91303-91310. doi: 10.1109/ACCESS.2020.2994207

[4] Chi, M., Tang, S., Teoh, S. O. O. S., Member, S., Ibrahim, H., Member, S., Embong, Z. 2022. A Deep Learning Approach for the Detection of Neovascularization in Fundus Images Using Transfer Learning. IEEE Access, 10, 20247–20258. doi: 10.1109/ACCESS.2022.3151644

[5] Chou, F. I., Tsai, Y. K., Chen, Y. M., Tsai, J. T., Kuo, C. C. 2019. Optimizing Parameters of Multi-Layer Convolutional Neural Network by Modeling and Optimization Method. IEEE Access, 7, 68316–68330. doi: 10.1109/ACCESS.2019.2918563

[6] Domingo, J. D., Aparicio, R. M., Miguel, L., Rodrigo, G. 2022. Cross Validation Voting for Improving CNN Classification in Grocery Products. IEEE Access, 10, 20913–20925. doi: 10.1109/ACCESS.2022.3152224

[7] Dudziak, L., Chau, T., Abdelfattah, M. S., Lee, R., Kim, H., Lane, N. D. 2020. BRP-NAS: Prediction-based NAS using GCNs. Adv Neural Inf Process Syst, 2020-Decem, 1–11.

[8] Elsken, T., Metzen, J. H., Hutter, F. 2019. Neural Architecture Search. J Mach Learn Res, 20, 1--21.

[9] Elsken, T., Staffler, B., Metzen, J. H., Hutter, F. 2020. Meta-Learning of Neural Architectures for Few-Shot Learning. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit, 12362–12372. doi: 10.1109/CVPR42600.2020.01238

[10] Elsken, T., Zela, A., Metzen, J. H., Staffler, B., Brox, T., Valada, A., Hutter, F. 2022. Neural Architecture Search for Dense Prediction Tasks in Computer Vision.

[11] Feng, J., Chen, J., Liu, L., Cao, X., Zhang, X., Jiao, L., Yu, T. 2019. CNN-Based Multilayer Spatial-Spectral Feature Fusion and Sample Augmentation with Local and Nonlocal Constraints for Hyperspectral Image Classification. IEEE J Sel Top Appl Earth Obs Remote Sens, 12, 1299–1313. doi: 10.1109/JSTARS.2019.2900705

[12] Fielding, B., Zhang, L. 2018. Evolving Image Classification Architectures with Enhanced Particle Swarm Optimisation. IEEE Access, 6, 68560–68575. doi: 10.1109/ACCESS.2018.2880416

[13] Ganorkar, S. R., Shujaur Rahman, M. Iris Recognition System And Analysis Using Neural Networks.

[14] Guerrero-Viu, J., Hauns, S., Izquierdo, S., Miotto, G., Schrodi, S., Biedenkapp, A., Elsken, T., Deng, D., Lindauer, M., Hutter, F. 2021. Bag of Baselines for Multi-objective Joint Neural Architecture Search and Hyperparameter Optimization.

[15] Gupta, S., Member, S., Imani, M., Member, S. 2019. NNPIM: A Processing In-Memory Architecture for Neural Network Acceleration. IEEE Transactions on Computers, 9340. doi: 10.1109/TC.2019.2903055

[16] Juneja, M., Joshi, S., Singla, N., Ahuja, S., Saini, S. K., Thakur, N., Jindal, P. 2022. Denoising of computed tomography using bilateral median based autoencoder network. Int J Imaging Syst Technol, 32, 935–955. doi: 10.1002/IMA.22668

[17] Lee, S. J., Chen, T., Yu, L., Lai, C. H. 2018. Image Classification Based on the Boost Convolutional Neural Network. IEEE Access, 6, 12755–12768. doi: 10.1109/ACCESS.2018.2796722

[18] Liu, C., Zoph, B., Neumann, M., Shlens, J., Hua, W., Li, L. J., Fei-Fei, L., Yuille, A., Huang, J., Murphy, K. 2017. Progressive Neural Architecture Search.

[19] Mehta, Y., White, C., Zela, A., Krishnakumar, A., Zabergja, G., Moradian, S., Safari, M., Yu, K., Hutter, F. 2022. NAS-Bench-Suite: NAS Evaluation is (Now) Surprisingly Easy.

[20] Nath, S. S., Mishra, G., Kar, J., Chakraborty, S., Dey, N. 2014. A survey of image classification methods and techniques. 2014 Int Conf Control Instrumentation, Commun Comput Technol ICCICCT 2014, 554–557. doi: 10.1109/ICCICCT.2014.6993023

[21] Oloulade, B. M., Gao, J., Chen, J., Lyu, T., Al-Sabri, R. 2022. Graph Neural Architecture Search: A Survey.

[22] Sekanina, L. 2021. Neural Architecture Search and Hardware Accelerator Co-Search: A Survey. IEEE Access, 9, 151337–151362. doi: 10.1109/ACCESS.2021.3126685

[23] Sensing, R., Scene, I. 2022. RS-DARTS: A Convolutional Neural Architecture Search for Remote Sensing Image Scene Classification.

[24] Sharma, A. K., Member, S., Tiwari, S. 2022. Dermatologist-Level Classification of Skin Cancer Using Cascaded Ensembling of Convolutional Neural Network and Handcrafted Features Based Deep Neural Network. IEEE Access, 10. doi: 10.1109/ACCESS.2022.3149824

[25] Sun, Y., Xue, B., Zhang, M., Yen, G. G. 2019. A Particle Swarm Optimization-Based Flexible Convolutional Autoencoder for Image Classification. IEEE Trans Neural Networks Learn Syst, 30, 2295–2309. doi: 10.1109/TNNLS.2018.2881143

[26] Wang, X., Wang, X., Jin, L., Lv, R., Dai, B., He, M., Lv, T. 2021. Evolutionary Algorithm-Based and Network Architecture Search-Enabled Multiobjective Traffic Classification. IEEE Access, 9, 52310–52325. doi: 10.1109/ACCESS.2021.3068267

[27] Weng, Y., Zhou, T., Li, Y., Qiu, X. 2019. NAS-Unet: Neural Architecture Search for Medical Image Segmentation. IEEE Access, PP:1. doi: 10.1109/ACCESS.2019.2908991

[28] Weng, Y., Zhou, T., Liu, L., Xia, C. 2019. Automatic Convolutional Neural Architecture Search for Image Classification under Different Scenes. IEEE Access, 7, 38495–38506. doi: 10.1109/ACCESS.2019.2906369

[29] White, C., Zela, A., Ru, B., Liu, Y., Hutter, F. 2021. How Powerful are Performance Predictors in Neural Architecture Search?

[30] Zela, A., Elsken, T., Saikia, T., Marrakchi, Y., Brox, T., Hutter, F. 2019. Understanding and Robustifying Differentiable Architecture Search.

[31] Zela, A., Klein, A., Falkner, S., Hutter, F. 2018. Towards Automated Deep Learning: Efficient Joint Neural Architecture and Hyperparameter Search.

[32] Zhou, X., Qin, A. K., Gong, M., Tan, K. C. 2021. A Survey on Evolutionary Construction of Deep Neural Networks. IEEE Trans Evol Comput, 25, 894–912. doi: 10.1109/TEVC.2021.3079985

[33] Zhu, H. 2021. Dark Area Classification From Airborne SAR Images Based on Deep Margin Learning. IEEE Access, 9, 139823–139841. doi: 10.1109/ACCESS.2021.3119076

[34] Zhu, N., Yu, Z., Kou, C. 2020. A New Deep Neural Architecture Search Pipeline for Face Recognition. IEEE Access, 8, 91303–91310. doi: 10.1109/ACCESS.2020.2994207.

[35] Zhu, P., Isaacs, J., Fu, B., Ferrari, S., Member, S. 2017. Deep Learning Feature Extraction for Target Recognition and Classification in Underwater Sonar Images.