International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol. 9, Issue.4, April 2021 E-ISSN: 2347-2693

Image Segmentation In The Framework of Deep Learning – A Comprehensive Survey

Bolla Ramesh[1*], S. Kiran[2]

[1,2]Department of CSE, YSR College of Engineering, Proddatur, India

DOI: https://doi.org/10.26438/ijcse/v9i4.3540 | Available online at: www.ijcseonline.org

Received: 16/Apr/2021, Accepted: 21/Apr/2021, Published: 30/Apr/2021

***Abstract*—**Machine learning is one of the prime aspects that is used in various applications of artificial intelligence and is widely used. In the area of machine learning, deep learning is observed to be of great interest to the researchers with the improvement of computer-based data processing. Recent works in this area of deep learning have paved way to the new innovations in science, technology and applied research which has a wide application for identification and classification of images. Such a classification is quite imperative when security is of prime concern. Deep learning is basically an artificial intelligence-machine learning hybrid. This has provided a versatile and precise model that can result in better accuracy. However, theoretical designs and experiments that are existing till date are very complex. So, there is a need to develop techniques that reduce the computational complexity. Deep learning can be used to solve various problems in the study of images and patterns. Image Segmentation is one of such applications. This paper explores the recent work that is carried out in image segmentation using Deep Learning. Many methods that are introduced for image segmentation are based on supervised classification. In general, such methods work well if the training set are representative of the test images in the segment. However, issues can occur in the course of training and test results, due to the impairment in the hardware and the concerned protocols that are existing in various distributions. The weights that are assigned to the features need to be adaptively chosen for proper classification of the segmentation area. This further improves the processing capability of the algorithm so developed.

***Keywords*—**Deep Learning, Artificial Intelligence, Data Science, Image Segmentation, Supervised Classification

I. INTRODUCTION

Deep learning is a machine learning subset. It is more appropriatefor massive data analysis relatedto conventional machine learning. The efficiency of the algorithm increases with the increase in volume of the data. In comparison to conventional machine learning, this does not rely on the artificial determination of the application. Instead, it tries to achieve higher features directly from the data and a deep-level machine learning model through many function transformations[1].Deep learning, also called the Deep Neural Network, consists of several layers in each layer with a number of neurons. These layers may range from a few to thousands and may contain thousands of neurons in each layer (processing unit). The simplest feature of a neuron is to multiply the input values by the weight assigned to each input and to summarize the output. This finding is further investigated by activation function. This increases the precision of the deep learning model[2].

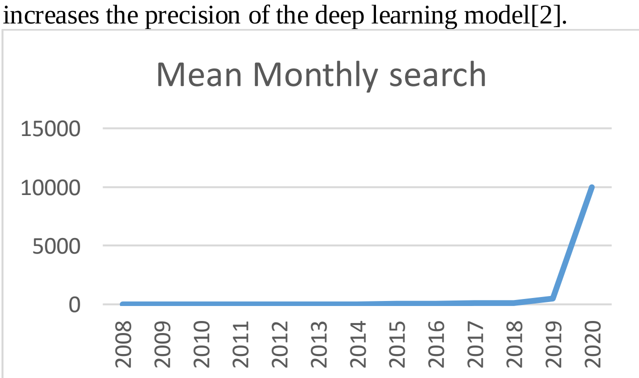

Fig.1. Deeplearning based google trends.

Figure 1 shows the google trends in deep learning. It can be observed that there is a steep increase in the recent era. With currentdevelopments in artificial intelligence, deep learning is smearedto various areas, often with total supervision, including speech recognition and visual recognition. A standard deep learning visual recognition architecture is based on a convolutionary neural network (CNN). Provided highly sized training data and with the capability of high-performance computing infrastructure, with thousands of categories, deep learning is able to achieve significant improvements in visual recognition. It is observed that the DL is providing high level of performance when supervised learning is applied and still this research is on early stage which means that there is a weak supervision and the common frameworks such as human labelling still is having a vital role in the analysis.

Big data is a high volumetric data with no or noisy tags, which is the very essence of large-scale web or real-world data. The advent of image search engines such as open and free based social network sites provides vision researchers with ample visual information; sadly, there are far shorter supplies of strong labels for these images. Unmonitored or weakly supervised approaches are also especially preferredbecause they can allow greater use of the large-scale web tools available.

In general, weakly supervised learning can be observed as a method for learning from sparse or noisy labels. In order to understand successful visual representations for classification, recognition, and segmentation, these weakly supervised approaches have been successfully used, all using weak labels alone, because web data typically will be of heterogeneous in nature with a high level of noise. In terms of visual recognition, apart from conventional exhaustive searches, individuals suggested a range of methodsthat are relevant to the current trend. Such strategies either implement saliency data, train systems with the models that are developed, or migrate to more sophisticated system that are having the capability of adaptive learning the segmentation algorithms that further reduce the search space. In general, these existing algorithms have very high recalls and also tend to observe least accuracies, which means that the developed techniques are not quite effective in the sense of noisy[3].

II. DEEP LEARNING ARCHITECTURES

A. Deep Belief Network

Deep Belief Networks (DBN) are an inherently generative graphical representation, i.e. It generates all distinct outcomes that can be generated in the case in question. DBNs consist of several layers with weights or values or costs based on the application, in which the layers are connectedd. The main aim or the prime function of the system is to assist the system in classifying the given input into various classes such that the output represents a class of the trained data. Figure 2 shows the DBN architecture.

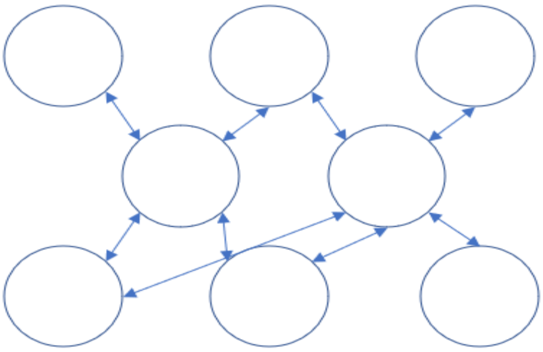

Fig. 2. Deep Belief Network Architecture

Neural networks of first generation used perceptions that depict a specific property or weight of the object that is of prime concern. This can be efficient only at a fundamental level and not useful in the later stages. The next generation of neural networks, are based on the principle of back propagation to overcome theproblems of the previous generation such that the obtained output to the expected output and reduces the value of error to almost zero. Trajectorybased machines enable the development and comprehension of more test cases that refer to previous trialcases. The next item was a recurringgraph known as creed networks that aid to resolve issues associated with interpretationand crammingissues. These networks also store the values in the leaf nodes that are not biased. Initially, the control signal from the pixels is taken directly and the values on each layer are considered next. Whenever another layer of properties and features is applied to the credentials network, the lower bound will increase the log likelihood of the training data set.

B. Convolutional Neural Networks (CNN)

A feedforward neural network has a depth based structure that involves convolution operation and is one of the noteworthy algorithm.This is observed to be the widely applied algorithm in the field of field of image recognition and classification. The prime points of this network are the local sensor region, weight sharing and the pooling layer. The main layer is the convolution layer that identifies the features and also classifies them for further recognition. The number of weights is adaptively aligned so that the complexity of the entire process of image extraction to the reconstruction phase is reduced. As the number of layers determine the accuracy of the system these are adaptively changed based on the information provided. When high quality images are processed, a large number of parameters need to be processed by completely connected neural networks so all the neurons will interact in the upper and lower layers. This also adopts the sparse linkapproach and adopts the weight sharing network structure via the "convolution kernel" as the intermediary. This also have an advantage of reduction in the number of parameters that are optimized from layer to layer and based on the image.

The basic structure of the CNN involves the layer of convolution, the pooling layer and the complete layer of link. Multiple coil units are made up of convolutional layer and pooling layer to remove features layer by layer, and ultimately classifies the given object.. The kernel which is used for the convolution operation is used as a filter in the process of image classification such that it filters out the negligible regions develop the unique values based on these regions that are having discardable values. The picture after convolution is still very large because of the dimensionrestriction of the convolution kernel. It is required to pool the picture at this moment, that is, it must be sampled such that it is modified or scaled down, hence, making the data dimension to be decreased. Various phases of CNN training is shown in Fig.3.

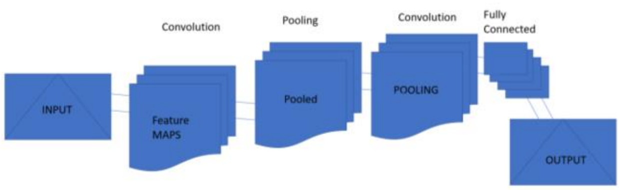

Fig. 3. Architecture of CNN

Batch Normalization (BN) aims at stopping the convolutionary nerve network gradually shifting or modifying the distribution of data input to the active layer during neural network training so as to allow a slower network convergence, i.e. the layer will accelerate convergence, in the case of back propagation. At the same time, the problem of insensitivity to the initialization weight of the convolution network can also be reduced and the fitting problem can be regulated[9].

$$F(x) = \gamma * \frac{x - E(x)}{\sqrt{\sigma^2 + \varepsilon}} + \beta \text{ [9]}$$

Where $E(x)=\frac{1}{n}\sum_{i=1}^{n}x_i$ and $\sigma=\frac{1}{n}\sum_{i=1}^{n}(x_i-E(x))^2$, the training parameters are β and γ.

Rectified Linear Unit (ReLu), which is a nonlinear operational activation function on batch normalization results which is given by the formulation:

$$F(x) = ma(0,x)$$

As the input of any CNN model is the image patch in this case, the centering on the feature point to reduce the size of an image patch such that it represents the features points that can be utilized for the input of the model.

TABLE I. CNN ALGORITHM

Algorithm for training of CNN
Step 1. Input the image/object for training.
Step 2. Initialization of the weight.
Step 3. Calculate the output of each layer.
Step 4. Solve for the error between the current expected output
Step 5. Check for the error margin.
Step 6. If error is within the margin then end the training process.
Step 7. If error is not within the margin, then calculate the error gradient, update the weights and move to step3.

C. Recurrent Neural Network

Recursive neural network is an artificial neural network with a tree-like hierarchical structure, in which nodes of the network heuristically develop a sequence of input based on the link and is one of the applications of the deep learning algorithms. The hidden layer is observed to have an addition bias which is to be of feedback type to the input layer, output layer and hidden layer. It can be understood that the description of the present input and the past output is the input to the hidden layer. The temporal and structural NN are divided into recursive neural networks. In order to form a directed graph, the neurons of the temporal recursive neural network are related, and a complex structure is developed using structural network. By using both supervised and unsupervised learning theories, recurrent neural networks can be educated. Weights are modified by the use of back propagation algorithms during supervised learning. For functional learning of structural knowledge, unsupervised learning is used. The significant changebetween the recurrent and the feedforward neural network is that there is a certain' memory' in the recurrent neural network. This has the advantage that input data is related, can be processed and input that is not inherently large can be handled. It requires several training parameters, however, and has no distinctive learning capacity. Figure 4 shows the architecture of recurrent neural networks.

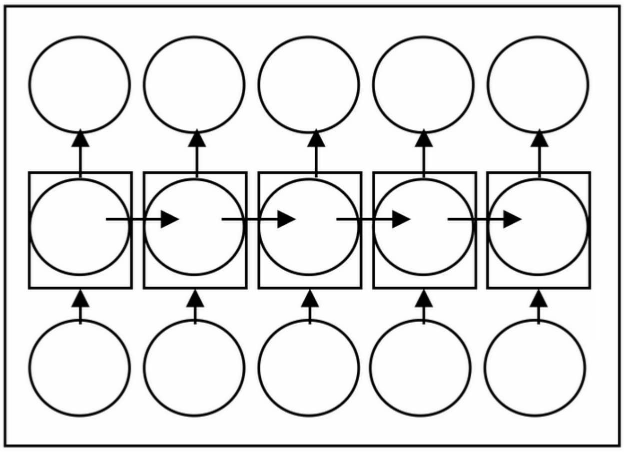

Fig. 4. Architecture of recurrent neural networks

III. IMAGE SEGMENTATION USING DEEP LEARNING

Figure 5 shows the basic block diagram of image segmentation using deep learning approach. The input image for processing is pre-processed for any changes in the input that is required for standardization. Standardization is required for the further process as it may reduce the computational complexity as well as all the objects that are being processed will be of the same standard norm which is quite required for the purpose of computation. The standardized object which is normalized in size and resolution wise is processed using feature extraction process. Feature extraction process is quite imperative as this stage defines the accuracy of the proposed technique.

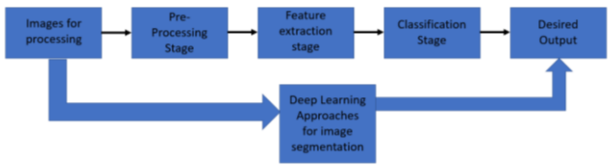

Fig. 5. Block Diagram of Image segmentation using Deep Learning Approach

It is very well understood that as the number of features increase, the overhead on the computation process also increases. Hence, optimization techniques are proposed in the literature. The next stage is the crucial stage which is the classification stage. Here, the traditional way of classification is compared with that of the deep learning-based strategy and the desired output is compared with that of the results of the deep learning output.

A critical step for both medical science and clinical practice is the dissectionof biomedical images into different parts such that they can be easily identified. Automatic segmentation is important because it is very laboriousand vulnerable to the subjective classification and to manually segment three-dimensional images. Many methods of automated segmentation are deployed based on the known or trained values and are further classification. The final decision that the class to which it belongs (which tissue or structure) is made per voxel by a decision-based classification on a manually annotated training collection. For instance, in brain MRI, such techniques have been used for whole-brain segmentation, segmentation of brain tissue, segmentation of white matter lesions, and segmentation of brain structure.

The prime task of any biomedical image processing application is to initially segment the given image such that the basic structure of the image need to be revealed for proper guidance for surgery and for further diagnosis. Researchers have developed many automated approaches to segmentation over the past decades. However, in order to be clinically useful, fully automated segmentation methods seldom produce sufficiently accurate and robust performance. Usually, this occurs because of low quality images that are impaired due to various factors, inhomogeneous appearances, and heterogeneity of protocols among users that lead to the development of a new structure. Interactive segmentation methods are desirable to resolve the shortcomings of automated segmentation approaches, as they permit greater precision and robustness in many applications. The system is observed to have high level of time complexity and also having labor intensive procedures for annotating manually the segmented parts.

For as few user interactions as possible, a good synergistic segmentation approach can achieve effective and efficient results, contributing to interaction performance. There are mainly two variables that have a crucial effect on its efficiency and usefulness for such a system. The first is the type of user experiences used as the input of the methodand the other is the underpinning model of the algorithm. Most of them are faced with the need for a huge number of user interferences and long time that a user handles, or restricted learning ability despite the large number of current interactive segmentation techniques[14].The next section describes the applications of deep learning in image segmentation.

IV. APPLICATIONS

D. Segmentation of parts of the eye

In recent decades, automated optic disk (OD) and optic cup (OC) segmentation techniques have been thoroughly researched. Early attempts to segment OC and OD in fundus images are based on hand-craft characteristics such as color, contrast thresholding, level-set approach, methods based on clustering and others. Manually built features, however, lack adequate discriminative capacity to have a high performance impact on the imaging conditions and the complexity of pathological regions. Deep learning techniques, on the other hand, learn comprehensively from the labelled training set and quickly become a common methodology for analyzing medical images. Convolutional neural networks, in particular, expose the compromises many tasks such as segmentation, classification and soon becoming a common method for segmentation of OD and OC, results[11].

For all trio apps: voxelwise brain tissue, white matter lesion (WML), and hippocampus segmentation, image weighting substantially enhanced efficiency over weighting all training images equally. This persuasive advantage of picture weighting correlates to previous results. Kullback-Leibler (KL) weighting worked much better than other weighting strategies for the WML segmentation, which had extremely unbalanced classes. KL weighting presumably outperformed other weighting strategies here, since it gives greater meaning to parts of the distribution with a limited previous distribution P (x). The best techniques varied for brain tissue and hippocampus segmentation that were relatively having minor differences and inconsistencies between the three methods of weighting[12].Super-resolution can provide more accurate images that will help enhance the precision of segmentation, while label maps in the segmentation dataset will lead to finer edges during the process of super-resolution[13].In order to obtain an initial automatic segmentation, a 2-stage framework is proposed with a P-Net and an R-Net to improvethe result based on user experiences that are converted into geodesic distance maps and then incorporated into the R-Net input. With dilated convolution for dense prediction, a resolution-preserving network structure is proposed and the current RNN-based CRF is expanded so that it can learn pair-wise freeform capacity and take advantage of user interactions as hard constraints. Placenta segmentation outcomes from 2D fetal MRI and brain tumors from 3D FLAIR images indicate that better results than automatic CNNs are obtained by our proposed process. This needs much fewer users. Compared to conventional interactive approaches, time achieves greater precision for 3D segmentation of brain tumors [14].An annotation-free segmentation approach increases segmentation accuracy. A target image conformingto the conventional unevensegmentation effectively using adversarial training is synthesized. The images in synthesis and such process helps us to train the models of segmentation better. The experimental results in the segmentation tasks of both cells and vessels indicate a substantial improvement, and also demonstrates the successful generalization and efficacy. For instance, in some applications, our technique is slightly substandardto the fully supervised techniques that require additionally high image detail accuracy[15].Segmentation based on color, watershed algorithm, Itti Koch, visual saliency based on graph, binary image for microscopic image segmentation was investigated[16]. The outcome of the system segmentations is centered on a threshold, where the output isbetter. The first step for the classification of the cell type is the segmentation of the nucleus. When the features of the nucleus and cytoplasm are removed, the prediction is so simple. Morphological analysis of WBC can detect blood cancer[17].A new approach to segment differential interference contrast (DIC) images with high precision is proposed based on a mixtureof deep learning and the transformation of the watershed. The key clueof this technique is to train CNN so that to detect both cellular markers and regions and to separate the individual cells using the transformation of the watershed based on these predictions. Based on the images of dense HeLa cell populations included in the Cell Tracking Challenge database, the strategy was developed[18].

A transformed low-rank and structured sparse decomposition (TLS2D) method is proposed for the concurrent segmentation and restoration of pathological MR brain images. For extracting pathological regions and while restoring quasi-normal MR images, the proposed method is robust. Experiments on synthetic and real MR brain tumor images show that both image recovery and tumor segmentation can effectively provide satisfactory output in our process. TLS2D is a popular technique for evaluating MR brain tumor scans and is useful for improving the diagnosis and follow-up of individual patients with cancer[19].

The training performance and efficiency of deep convolutionary neural networks (DCNNs) have been significantly improved by residual networks (ResNet) and densely connected networks (DenseNet), mainly for object classification tasks. Hence, by considering the benefits of both networks, an optimal network architecture is suggested. For medical image segmentation, the proposed method is incorporated into an encoder-decoder DCNN model[20].

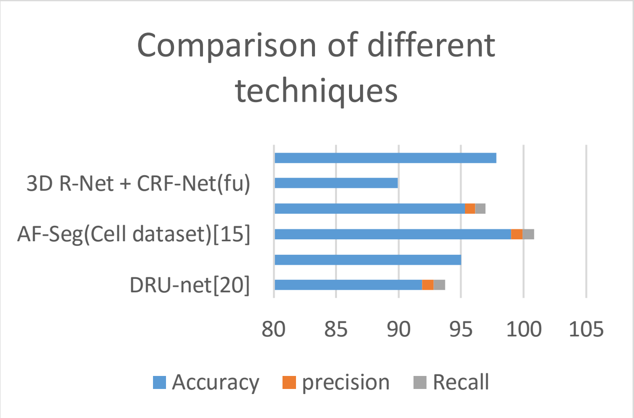

Fig. 6. Comparison of Accuracy of different techniques.

Figure 6 shows the comparative analysis of the various algorithms that are used by different researchers in the recent past.

V. CONCLUSION

Current trend of learning depth-based image classification algorithm, especially based on the convolution of the neural network algorithm to deal with some simple image classification, can achieve a very good effect, but in some complex image processing, there is a certain change in particular, such as face recognition, face will be like time, light, angle, constant change, and how to improve. The present situation of tagged data training still occupies a dominant role in the field of feature learning. With the rising volume of data, however, applying tags to all the data is becoming more and more impractical. Therefore, marking the data automatically or training the network by unmarked data is becoming more and more urgent. On the other hand, how to increase the training pace while maintaining the accuracy of recognition is also a hot subject of current research in the field of image recognition and classification. Currently, the setting of such training parameters for a convolutionary neural network is focused on individual experience or experiments, and there is no systemic mechanism of control. It is hoped that there will be an automated adjustment method for the related parameters in the network in the future, so that the network can fully achieve global optimization and reduce the time needed for network preparation. Based on this, a more systematic and in-depth analysis of the algorithm for image classification based on deep learning must be carried out.

REFERENCES

[1] Y. -n. Dong and G. -s. Liang, "Research and Discussion on Image Recognition and Classification Algorithm Based on Deep Learning," 2019 International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), Taiyuan, China, 2019, pp. 274-278, doi: 10.1109/MLBDBI48998.2019.00061.

[2] O. Sharma, "Deep Challenges Associated with Deep Learning," 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India, 2019, pp. 72-75, doi: 10.1109/COMITCon.2019.8862453.

[3] J. Wu, Yinan Yu, Chang Huang and Kai Yu, "Deep multiple instance learning for image classification and auto-annotation," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 3460-3469, doi: 10.1109/CVPR.2015.7298968.

[4] S. Khan, E. Ahmed, M. H. Javed, S. A. A Shah and S. U. Ali, "Transfer Learning of a Neural Network Using Deep Learning to Perform Face Recognition," 2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), Swat, Pakistan, 2019, pp. 1-5, doi: 10.1109/ICECCE47252.2019.8940754.

[5] L. Khelifi and M. Mignotte, "Deep Learning for Change Detection in Remote Sensing Images: Comprehensive Review and Meta-Analysis," in IEEE Access, vol. 8, pp. 126385-126400, 2020, doi: 10.1109/ACCESS.2020.3008036.

[6] H. Ucuzal, Ş. YAŞAR and C. Çolak, "Classification of brain tumor types by deep learning with convolutional neural network on magnetic resonance images using a developed web-based interface," 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 2019, pp. 1-5, doi: 10.1109/ISMSIT.2019.8932761.

[7] L. Zhuang and Y. Guan, "Deep Learning for Face Recognition under Complex Illumination Conditions Based on Log-Gabor and LBP," 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chengdu, China, 2019, pp. 1926-1930, doi: 10.1109/ITNEC.2019.8729021.

[8] J. Latif, C. Xiao, A. Imran and S. Tu, "Medical Imaging using Machine Learning and Deep Learning Algorithms: A Review," 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 2019, pp. 1-5, doi: 10.1109/ICOMET.2019.8673502.

[9] Y. Liu, X. Xu and F. Li, "Image Feature Matching Based on Deep Learning," 2018 IEEE 4th International Conference on Computer and Communications (ICCC), Chengdu, China, 2018, pp. 1752-1756, doi: 10.1109/CompComm.2018.8780936.

[10] I. Karabayir, O. Akbilgic and N. Tas, "A Novel Learning Algorithm to Optimize Deep Neural Networks: Evolved Gradient Direction Optimizer (EVGO)," in IEEE Transactions on Neural Networks and Learning Systems, doi: 10.1109/TNNLS.2020.2979121.

[11] P. Yin, R. Yuan, Y. Cheng and Q. Wu, "Deep Guidance Network for Biomedical Image Segmentation," in IEEE Access, vol. 8, pp. 116106-116116, 2020, doi: 10.1109/ACCESS.2020.3002835.

[12] A. Van Opbroek, H. C. Achterberg, M. W. Vernooij and M. De Bruijne, "Transfer Learning for Image Segmentation by Combining Image Weighting and Kernel Learning," in IEEE Transactions on Medical Imaging, vol. 38, no. 1, pp. 213-224, Jan. 2019, doi: 10.1109/TMI.2018.2859478.

[13] S. Lei, Z. Shi, X. Wu, B. Pan, X. Xu and H. Hao, "Simultaneous Super-Resolution and Segmentation for Remote Sensing Images," IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019, pp. 3121-3124, doi: 10.1109/IGARSS.2019.8900402.

[14] G. Wang et al., "DeepIGeoS: A Deep Interactive Geodesic Framework for Medical Image Segmentation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 7, pp. 1559-1572, 1 July 2019, doi: 10.1109/TPAMI.2018.2840695.

[15] F. Yu et al., "AF-SEG: An Annotation-Free Approach for Image Segmentation by Self-Supervision and Generative Adversarial Network," 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 2020, pp. 1503-1507, doi: 10.1109/ISBI45749.2020.9098535.

[16] S. Mishra, P. Liang, A. Czajka, D. Z. Chen and X. S. Hu, "CC-NET: Image Complexity Guided Network Compression for Biomedical Image Segmentation," 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 2019, pp. 57-60, doi: 10.1109/ISBI.2019.8759448

[17] S. M. Sundara and R. Aarthi, "Segmentation and Evaluation of White Blood Cells using Segmentation Algorithms," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2019, pp. 1143-1146, doi: 10.1109/ICOEI.2019.8862724.

[18] F. Lux and P. Matula, "DIC Image Segmentation of Dense Cell Populations by Combining Deep Learning and Watershed," 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 2019, pp. 236-239, doi: 10.1109/ISBI.2019.8759594.

[19] C. Lin, Y. Wang, T. Wang and D. Ni, "Segmentation and Recovery of Pathological MR Brain Images Using Transformed Low-Rank and Structured Sparse Decomposition," 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 2019, pp. 1878-1881, doi: 10.1109/ISBI.2019.8759441.

[20] M. Jafari, D. Auer, S. Francis, J. Garibaldi and X. Chen, "DRU-Net: An Efficient Deep Convolutional Neural Network for Medical Image Segmentation," 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 2020, pp. 1144-1148, doi: 10.1109/ISBI45749.2020.9098391.