International Information and Engineering Technology Association

Advancing the World of Information and Engineering

An Efficient Deep Neural Framework for Nucleus Semantic Segmentation with Enhanced U-Net

Check for updates

Rahul Sawhney[1], Shilpi Sharma[1*], Purvit Vashishtha[1], Arun Kothari[1], Vibhu Gupta[1], Tanupriya Choudhury[2]

[1] Department of CSE, Amity University, Uttar Pradesh 201313, India
[2] Informatics Cluster, School of Computer Science, University of Petroleum and Energy Studies (UPES), Dehradun 248007, Uttarakhand, India

Corresponding Author Email: ssharma22@amity.edu

ABSTRACT

With the advancement of medicine, there is a terrific requirement to process an increasing number of medical imaging where image segmentation comes into play. Single cell division is generally among the first as well as most essential tasks in image-based cell analysis. The identification of the nuclei allows pathologists to determine each cell in the sample, and by measuring how cells respond to various treatments, they can comprehend the fundamental biological processes in work. In this paper we gathered a deep learning network which detects and parts the cell microscopy image. Highly advanced performance is achieved in image segmentation tasks through deep learning-based techniques. These procedures are complex and need the support of compelling computational resources. This paper emphasizes the basic principles of the methods used to segment an image. We have implemented U-Net for the semantic segmentation of nucleus. A successful implementation will aid researchers immensely in their fight to find pharmaceutical solutions to medical crises while saving both valuable research time and funding.

1. INTRODUCTION

Microscopy is a central technology of specialty clinical research discussed in nearly a million scientific articles indexed in Pub-Med to date (Pub-Med search 'microscopy OR magnifier OR microscopic,' consulted seven months of the Gregorian calendar 2018). The photographs taken are increasingly being evaluated quantitatively [1-3]. Different research techniques enable the recording of structural and practical properties of biological model systems and cells, tissues and dead-end organoids. The cell nucleus could be a reliable reference compartment for characteristic single cells in imaging research. The segmentation of cells or cell nuclei is an essential step in analyzing medical research images [4]. Assuming the correct segmentation of cells or cell nuclei, multiple biological or medical analyses can then be performed along with cell-type classification [5], explicit cell count [6], cell composition analysis [7], providing valuable diagnostic data for doctors and researchers. Though common image processing strategies are presently still in use and labor-intensive tasks are often unable to deliver the products for optimized performance for several reasons, we appreciate the limited ability to manage multiple images [4].

The primary is that the beginning of numerous microscopic cell analyses. The proper location of the core is the basis for a variety of quantitative measurements through crucial step within the boundaries of the individual cells for several extra analyses. The predominant approaches for this task were classic image process algorithms in the current scenario [8], which were usually guided by form and special background [9]. It should be readjusted as before due to the fact that the speculative condition alters. Deep learning has eventually upgraded a variety of picture study aims, ranging across image categorization [10] to detect faces [11] and scene partitioning [12].

In 2018 it was hosted by Kaggle, Booz Allen Hamilton, and therefore the Broad Institute, the Data Science Bowl (DSB) confronted attendees to drive advances in core segmentation. It exceeds the characteristics of the conditions and {and in connection with} different cell lines, remedies, and the types of solar microscopy. Many knowledgeable scientists participated in the challenge from all over the world. - We do not present the most effective ancient performance methods hierarchically to more than 1,000 of 3,891 submissions in level 1. The fundamental forms based on deep learning were based on only a few of the many designs; the aspects that participants frequently believed influence their rankings were associated to the data: the quantity of data, pre-processing, and data upgrading strategies.

Many scientists began exploring the prospective uses of deep learning (DL) in cell or cell nuclei segmentation as the domain of deep learning (DL) grew. Here we tend to study the construction of a segmentation model that automatically determines the nucleus through multiple sets of two-dimensional (2D) light analysis footage that has been stained without any kind of human interaction. Such a model would potentially enable future automated microscopes to enable some extensive biological uses by understanding and imaging nuclei in excessive amounts across cell varieties, staining types and magnifications despite experimental variations. In research images, classic algorithms for characteristic kernels follow similar machine strategies with different parameters or configurations. Our goal was to see whether a fancy solution that take account high-capability deep learning models would

provide a unified solution without the need for continuous configuration. Biological image segmentation supported by machine learning already be present in simple software comparable to Ilastik and ImageJ, and recent studies guarantee the quality of this method. Deep learning has great ability for solving powerful cellular image study, and neural network models also exist. Current solutions, on the other hand, demand users to design bespoke models, annotations, model training, and portion algorithms for each experiment. Instead, we tend to develop a generic, reproducible model that "can be trained once, published, and used in a variety of research projects without any explicit user intervention." We tend to employ illustrated core segmentation software tools that are used as face indicators in natural images with constant ease and power; they perform without the need for users to direct models or build sets, and they work in a variety of light and stage situations.

2. LITERATURE REVIEW

In paper [13], Benchmark image segmentation dataset was used on which image augmentation technique was performed using GANs and then late nuclei segmentation. Though there is still no intrinsic metric evaluation present which contributes for better model training and for generating complex outputs.

According to the study [14], for fragmentation of cell image multiple datasets were used so as to apply deep learning. It follows the methodology utilizing a CNN firstly along with about a million path-images in order to recognize different cell regions. Its limitation was that it discards images which were giving results as low chances of containing abnormal cells.

Pap smear dataset was used by Ronneberger et al. [15] for implementation of another upgraded version of U-net model that is progressive growing PGU-net model. This model implemented two patterns for image extraction features at different scales but still it is unable to extract some of the complex image- derived features which contributes for more accurate image segmentation.

In Ref. [16], segmentation of nucleus image was performed on DBS 2018 (IV- A) dataset using U-net Network approach. Also in this process, SAP and CE modules were utilized which is then for segmenting nucleus image, is added into the U-net network. For this experiment, offline method was used which means that a larger storage space is required for this method.

As suggested by the study [17], U-net like decoders on Data Science Bowl (2018) dataset were used in this research. FPN decoding scheme is used while encoders included ResNET, Inception- ResNET and Dual Path Networks. Data availability is a limitation for this experiment which results in performance issues.

Hollandi et al. [18] presented 53 images out of total dataset used is 3426 nuclei. Techniques used were convolutional neural networks along with U-Net with a ResNET34 backbone. It also includes CNN with an attached region proposal and segmentation network.

In the study [19], dataset used is prepared with the help of hospital. Methodology is that after preprocessing stage, segmentation is performed on the image. Then various types of clustering algorithms are applied and then at last post processing is done.

In Ref. [20], Unet model is enhanced to light weight model with modified enhanced branch so that it would potentially be able to work with low-resources computing. This model is

then applied to Data science bowl 2018 dataset and the limitation is that masks needed to be constructed for removal of isolated objects and small holes within the image.

Punn and Agarwal [21] used the dataset provided by the National Library of Medicine on which deep learning based nuclei segmentation is performed. The methodology used is firstly clustering approach by k-means++ and then later segmentation method by gathering localized information. For this method, limitation was that complete merging of some of the pills were coming which were having color similar to the background which resulted in complete black mask.

3. METHODOLOGY

In our work, survey of Scopus and web of science indexed journals research papers from 2010 to 2021 were considered. The methodology used for selecting literature is shown in Algorithm 1.

Algorithm 1 Paper search strategy

---

procedure TOPIC (Nuclei Image Segmentation)
   SearchDatabases ← IEEEXplore, GoogleScholar, ScienceDirect
   SearchYear←2010-2021.
► *counter←*1
   Initialize count
► *NIS ←* 15
   NIS is the Number of search database

   for *i ≤ NIS* **do**
     keyword ← Nuclei Segmentation, Machine Learning, Segmentation
     if *SearchLink* ϵ *SearchDatabases* and *Year* ϵ *SearchYear* **then**
       Search (Nuclei Segmentation **AND** Segmentation **AND** machinelearning)
     end if
   end for

   if *NumberofPaper ≥* 0 **then**
     Refine Paper
     applyinclusioncriteria ← IC1, IC2
     applyexclusioncriteria ← EC1, EC2, EC3
   end if
end procedure

3.1 Paper selection

Also, the paper Inclusion Criteria (IC) and Exclusion Criteria (EC) are indicated on Table 1.

Table 1. Inclusion criteria and exclusion criteria for paper selection

IC	EC
IC1: Peer reviewed paper.	EC1: Different databases have duplicate studies.
IC2: papers published in either Scopus or web of science indexed journals	EC2: MTech, MSc and PhD papers

3.2 Dataset

The dataset which is used in our paper is an assortment of cell images that differ throughout several cell type, image magnification and imaging methods. The training set is composed of 670 distinct raw images while the test set contained 65 images. All images in the test set were not labeled, but however, for every training example, there exists the raw image which are taken of a cell and a set of pixel masks that each displays the exact position of one cell nucleus in the original picture. Since, each mask is unique, hence there is no overlap between different nuclei masks. Figures 1 and 2 illustrates an example of a training pair. It also contains a csv file of the run-encoded pixel values for each nucleus' mask.
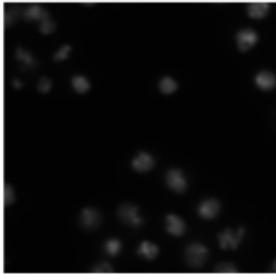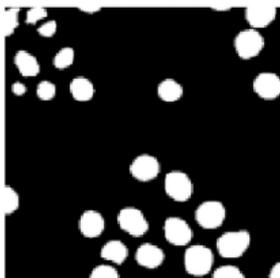

Figure 1. Cell microscopy image

Figure 2. Cell microscopy mask

3.3 Data manipulation

To train our model, firstly we have resized all the images to (256×256) pixels to ensure the identical input which allows faster training of our model; In a similar way, we converted colored images to grayscale to ensure all the training images had the same format. In particular, we used two main kind of transformations namely *horizontal* and *vertical flips* with likelihood of 0.25 and then at last, we converted all the greyscale images to *FloatTensor* with range [0, 1] and combined all of the individual image masks into a single binary image file.

3.4 Model architecture

Figure 3 is describing the UNet architecture. For our paper, we began by creating a UNet model that is based on an earlier architecture known as the FCN, is a CNN that substitutes fully connected layers with an inverted convolutional layer that upsamples the feature map according to the dimensions of the initial input image while keeping the location data [20, 21]. Unfortunately, due to excessive downsampling, the FCN's final layer suffers from information loss. Due to high information loss, upsampling becomes difficult, resulting in a shoddy output.

Net builds on the FCN by employing a design that is like the FCN but differs in key aspects instead of using a single transposed convolutional operation at the network's end, FCN uses comparable set of convolutional operations to downscale the input data to feature space, then transposed convolutional operations to upscale the mappings to initial source image dimensions. There are also skipped connections which help to keep track of data that otherwise would have been lost while encoding.

Contracting path: The image which is inserted is originally passed through a sequence of convolution layers which lowers down the dimensions in a way increasing the number of channels, so as to create a contracting route (encoder with down sampling steps). Crop function: To build a skip connection. This phase crops image coming through down sampling route and appends with the present picture on up sampling route Expanding path (Decoder with up-sampling steps): The expanding path reverses the contracting path's process, enlarging the picture to its initial dimensions whilst progressively reducing the number of channels. After that, the clipped feature map from the downsampling route is fused. Last Block Feature Maps: a 1×1 convolutional operation is applied in the top layer in order to transfer each constituent feature vector to desired no. of classes [22].
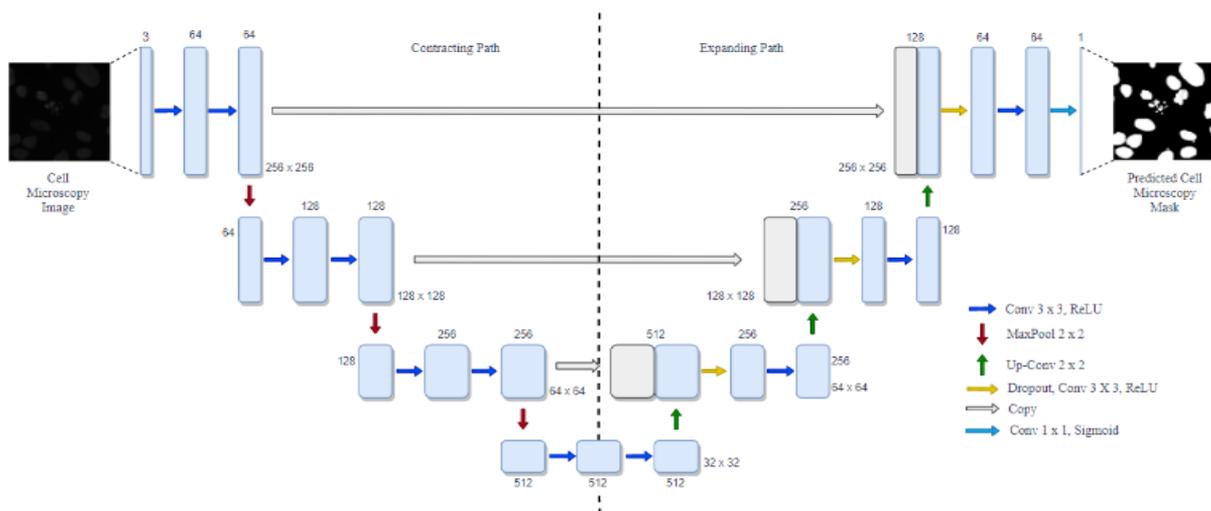

Figure 3. UNet Architecture

3.5 Model optimizer and LR scheduler

In our paper, we have used AdamW that is a modified version of weighted decaying [23]. In paper [24], dataset used was provided by the National Library of Medicine on which deep learning based nuclei segmentation is performed. The methodology used is firstly clustering approach by k-means++ and then later segmentation method by gathering localized information. For this method, limitation was that complete merging of some of the pills were coming which were having color similar to the background which resulted in complete black mask. In general implementation of the weighted decaying is automatically connected to L.R in the Adam optimizer, which implies that if we tune the L.R, we'll have to discover a better weighted decaying for every L.R; we try. Weighted decaying is decoupled in the AdamW optimizer, which means the weight decay and L.R can be tuned separately. i.e., they no longer influence each other. This results in improved generalization performance.

$$w_t = w_{t-1} - \eta \frac{m_t}{\sqrt{v_t} + \epsilon} \tag{1}$$

$$w_{t+1} = (1 - \lambda)w_t - \eta \triangledown f_t(w_t) \tag{2}$$

$$w_t = w_{t-1} - \eta \left(\frac{m_t}{\sqrt{v_t} + \epsilon} + \lambda w_{t-1} \right) \tag{3}$$

where, w_t=weight decay at time 't'; η=learning rate or step size at time t; m_t=aggregate of gradients at time t (bias corrected); v_t=sum of square of past gradients (bias corrected); ϵ=small positive constant to avoid '*division by 0*' error when v_t=0.

In addition to this, we have also used Lambda's Learning Rate schedular, which is used to tune the learning rate while training by increasing or decreasing the learning rate according to a defined schedule. In Pytorch's Lamba lr schedular we have tuned the learning rate of each parameter by multiplying the initial learning rate with a custom defined function applied on each epoch. This gives us the flexibility on the function that we have change according to the model's performance.

$$lr_{epoch} = lr_{initial} * Lambda(epoch) \tag{4}$$

4. RESULTS AND EVALUATION

4.1 Evaluation metric

For evaluation of our model, we have used IoU metric. Intersection Over Union (IoU) [25] also known as Jaccard index which is a metric used to judge the percentage of overlap between our predicted mask and the original mask. In simple terms it is the measure of total number of pixels that are common between out predicted mask and the original mask divided by the total number of pixels that are present. The Kaggle competition refers to its evaluation metric as an LB score, which quantifies model performance by measuring the mean across all test images of an average precision for each cell nucleus detection. This precision can be calculated over a given set of IoU threshold and averaged across the set of threshold values. For a threshold t in a list of thresholds [0.5→0.95], we consider our model to be accurate at that threshold t if the IoU score meets or exceed the certain threshold.

$$IoU(\hat{Y}, Y) = \frac{Y \cap \hat{Y}}{Y \cup \hat{Y}} \tag{5}$$

The Kaggle competition also defines a true positive as the model predicting a specific nucleus mask with IoU score that is greater than or equal to the threshold, Kaggle defines the precision over a set of thresholds T of a predicted set of masks $\hat{Y}$ with regards to a ground truth set of nucleus masks *masks* as:

$$P(\hat{Y}, masks) = \frac{1}{|T|} \sum_t \frac{TP(t)}{TP(t) + FP(t) + FN(t)} \tag{6}$$

where, the label is a set of masks *masks* and the entire score is averaged over all threshold values.

$$TP(\hat{Y}, mask, t) = \sum_{\substack{nucleus \in \hat{Y} \\ > t}} 1(IoU(nucleus, mask) \tag{7}$$

$$FN(\hat{Y}, mask, t) = |mask| - \left(\left| \hat{Y} - FP(t) \right) \right) \tag{8}$$

$$FN(\hat{Y}, mask, t) = \sum_{\substack{nucleus \in \hat{Y} \\ < t}} 1(IoU(nucleus, mask) \tag{9}$$

4.2 Loss function

1) Dice Loss: The term dice loss comes from the Strensen-Dice coefficient, a statistic established in the 1940s to determine how comparable two samples are. It was introduced in the computer vision community in 2016 for 3D medical image segmentation. It is also known as F1 score metric which is similar to IoU and used to find the similarity between two samples. It does exceptionally well in case of class imbalance in the dataset. The Dice coefficient loss is widely used to calculate the similarity between two images. In the Dice Loss, in addition, 1 is added to the numerators & denominators to guarantee whether the expression is defined in gray areas like Y equals 0 [26].

$$Dice\ Coefficient(\hat{Y}, Y) = 2 * \frac{Y \cap \hat{Y}}{Y \cup \hat{Y}} \tag{10}$$

$$Dice\ Loss(\hat{Y}, Y) = 1 - \frac{2 * (Y * \hat{Y}) + 1}{Y + \hat{Y} + 1} \tag{11}$$

2) DiceBCELoss: This loss blends Dice loss and the usual binary cross-entropy (BCE) loss, which is commonly used in segmentation. Using both approaches provide us variability in loss while maintaining BCE's consistency. The equation for multi-class BCE by itself will be familiar to anyone who has studied logistic regression.

$$(\hat{Y}, Y) = \frac{-1}{N} \sum_{n-1}^{N} [y_n log_n \hat{y}_n + (1 - y_n) * log_n (1 - \hat{y}_n)] \tag{12}$$

In order to estimate and favor the acceptance of our organized approach, we intend to direct this procedure ample amount of times by just changing and adjusting the hyperparameters of the model. Each and every time various results comes. The most appropriate results are revealed by this research paper. Figure 4 shows the generated model loss and accuracy plots illustrates the efficient segmentation of cell microscopy images keeping in mind that they are much harder to detect.
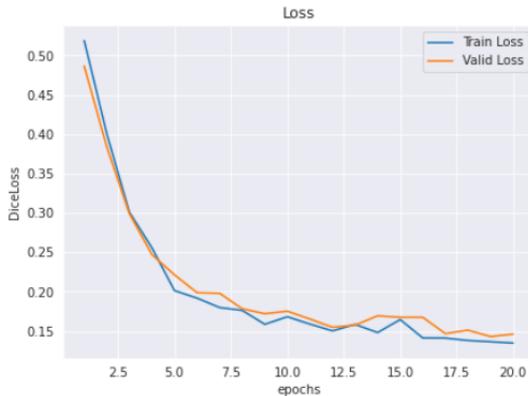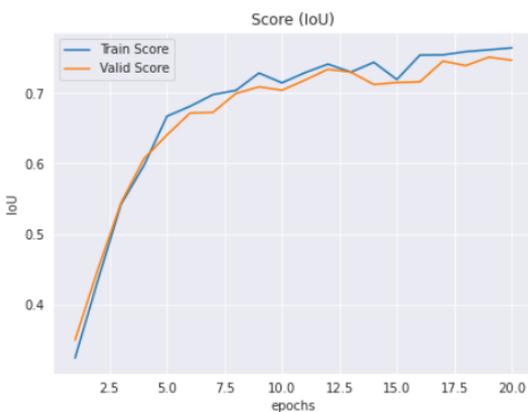

Figure 4. Model loss plot

Figure 5. Model accuracy plot

Our deep neural network is immune to noise as because it is highly robust to the fluctuating environment. We have depicted the model loss plot. So as to learn complex data patterns, many hidden layers help the model. From the Figure 5 we can see that model has achieved 78% and 76% accuracy in the *training* and the *validation* set.

Our results reveal considerable as shown in Figures 6, 7 and 8 respectively and it provides an assurance, as well as at a basic level the algorithm works well, which is shown by the success of the initial prediction.
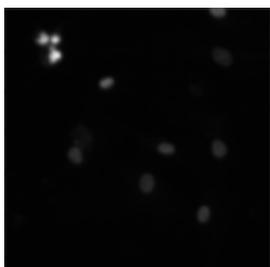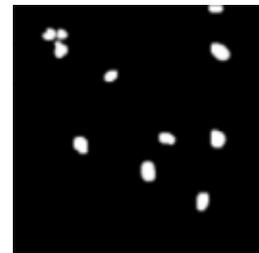

Figure 6. Training image
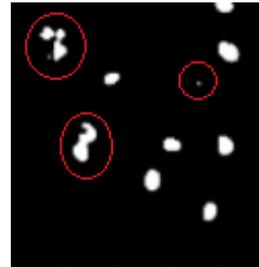

Figure 7. Image masking
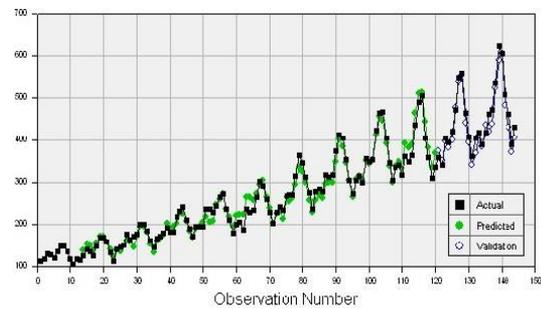

Figure 8. Predicted image mask

Figure 9. Time series values of patients

Figure 9 specifies the number of observations at the end of the series to use for the validation. The model uses only the observations prior to these held-out observations. This approach can be used on a time series that has more than one value, or so-called multivariate time series. However, there are certain defects in the prediction of our model as we can see that model is having hard time in predicting the masks that are conjoined. It is because the margin which divides them are too slim.

5. CONCLUSION AND FUTURE WORKS

We know that the beginning point for the majority of the cancer diagnostics is the identification of cells as most of the human body is taken into consideration to have trillions of cells containing a nucleus full of DNA, the hereditary code that is responsible for coding of each cell. The identification of the nuclei allows pathologists to determine each cell in the sample, and by measuring how cells respond to various treatments, they can comprehend the fundamental biological processes in work. In this research paper we gathered a deep learning network which detects and parts the cell microscopy image. Our segmentation network learned from given image/mask pairs. Our approach helped the network to adapt to a diverse set of test data outside the domain of the training data and surpasses many other classical methods. Our model has achieved 78% and 76% accuracy in the *training* and the

validation set. In addition, the deep learning [27-29] methodologies strategically developed and the methodology accuracy also helped to justify the model architecture.

Traditionally diagnostic tests, cytological tests are conducted through pathologists since they are easy, cheap, less intrusive, and more efficient. However, the process is time consuming and prone to inter- and intra-observer variations. To ease the workload & to accelerate the diagnosis procedure a computer aid-diagnostics system has been made for cancer cells, cellular nucleus segmentation is required since tumor cells are primarily identified by their morphology characteristics in the nucleus. Consequently, it is essential to choose a precise as well as efficient cell nuclei division techniques that can assist precisely delineate the nuclei shapes.

REFERENCES

[1] Sommer, C., Gerlich, D.W. (2013). Machine learning in cell biology–teaching computers to recognize phenotypes. Journal of Cell Science, 126(24): 5529-5539. https://doi.org/10.1242/jcs.123604

[2] Boutros, M., Heigwer, F., Laufer, C. (2015). Microscopy-based high-content screening. Cell, 163(6): 1314-1325. https://doi.org/10.1016/j.cell.2015.11.007

[3] Usaj, M.M., Styles, E.B., Verster, A.J., Friesen, H., Boone, C., Andrews, B.J. (2016). High-content screening for quantitative cell biology. Trends in Cell Biology, 26(8): 598-611. https://doi.org/10.1016/j.tcb.2016.03.008

[4] Caicedo, J.C., Roth, J., Goodman, A., et al. (2019). Evaluation of deep learning strategies for nucleus segmentation in fluorescence images. Cytometry Part A, 95(9): 952-965. https://doi.org/10.1002/cyto.a.23863

[5] Liu, Y., Long, F. (2019). Acute lymphoblastic leukemia cells image analysis with deep bagging ensemble learning. In ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging, pp. 113-121. https://doi.org/10.1007/978-981-15-0798-4_12

[6] Salovarda, M., Bolkovac, I., Domitrovic, H. (2005). Estimating perceptual audio system quality using PEAQ algorithm. In 2005 18th International Conference on Applied Electromagnetics and Communications, Dubrovnik, Croatia, pp. 1-4. https://doi.org/10.1109/ICECOM.2005.205017

[7] Bougen-Zhukov, N., Loh, S.Y., Lee, H.K., Loo, L.H. (2017). Large-scale image-based screening and profiling of cellular phenotypes. Cytometry Part A, 91(2): 115-125. https://doi.org/10.1002/cyto.a.22909

[8] Pandey, S., Singh, P.R., Tian, J. (2020). An image augmentation approach using two-stage generative adversarial network for nuclei image segmentation. Biomedical Signal Processing and Control, 57: 101782. https://doi.org/10.1016/j.bspc.2019.101782

[9] Araújo, F.H., Silva, R.R., Ushizima, D.M., Rezende, M.T., Carneiro, C.M., Bianchi, A.G.C., Medeiros, F.N. (2019). Deep learning for cell image segmentation and ranking. Computerized Medical Imaging and Graphics, 72: 13-21. https://doi.org/10.1016/j.compmedimag.2019.01.003

[10] Zhao, J., Dai, L., Zhang, M., Yu, F., Li, M., Li, H., Wang, W., Zhang, L. (2019). PGU-net+: Progressive growing of U-net+ for automated cervical nuclei segmentation. In International Workshop on Multiscale Multimodal Medical Imaging, pp. 51-58. https://doi.org/10.1007/978-3-030-37969-8_7

[11] Fang, J., Zhou, Q., Wang, S. (2021). Segmentation technology of nucleus image based on U-net network. Scientific Programming, 2021: 1892497. https://doi.org/10.1155/2021/1892497

[12] Kromp, F., Fischer, L., Bozsaky, E., Ambros, I., Doerr, W., Taschner-Mandl, S., Ambros, P., Hanbury, A. (2019). Deep Learning architectures for generalized immunofluorescence based nuclear image segmentation. arXiv preprint arXiv:1907.12975. https://arxiv.53yu.com/abs/1907.12975.

[13] Win, K.Y., Choomchuay, S., Hamamoto, K., Raveesunthornkiat, M. (2018). Comparative study on automated cell nuclei segmentation methods for cytology pleural effusion images. Journal of Healthcare Engineering, 2018: 9240389. https://doi.org/10.1155/2018/9240389

[14] Sornapudi, S. (2017). Nuclei segmentation of histology images based on deep learning and color quantization and analysis of real world pill images. Missouri University of Science and Technology.

[15] Ronneberger, O., Fischer, P., Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234-241. https://doi.org/10.1007/978-3-319-24574-4_28

[16] Jung, C., Kim, C., Chae, S.W., Oh, S. (2010). Unsupervised segmentation of overlapped nuclei using Bayesian classification. IEEE Transactions on Biomedical Engineering, 57(12): 2825-2832. https://doi.org/10.1109/TBME.2010.2060486.

[17] Long, F. (2020). Microscopy cell nuclei segmentation with enhanced U-Net. BMC Bioinformatics, 21(1): 1-12. https://doi.org/10.1186/s12859-019-3332-1

[18] Hollandi, R., Szkalisity, A., Toth, T., et al. (2019). A deep learning framework for nucleus segmentation using image styletransfer. bioRxiv, 580605. https://doi.org/10.1101/580605

[19] Englbrecht, F., Ruider, I.E., Bausch, A.R. (2021). Automatic image annotation for fluorescent cell nuclei segmentation. PloS One, 16(4): e0250093. https://doi.org/10.1371/journal.pone.0250093

[20] Lv, G., Wen, K., Wu, Z., Jin, X., An, H., He, J. (2019). Nuclei R-CNN: Improve mask R-CNN for nuclei segmentation. In 2019 IEEE 2nd International Conference on Information Communication and Signal Processing (ICICSP), pp. 357-362. https://doi.org/10.1109/ICICSP48821.2019.8958541

[21] Punn, N.S., Agarwal, S. (2021). Modality specific U-Net variants for biomedical image segmentation: A survey. arXiv preprint arXiv:2107.04537.

[22] Loshchilov, I., Hutter, F. (2018). Fixing weight decay regularization in Adam. https://doi.org/10.1007/978-3-319a

[23] Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., Savarese, S. (2019). Generalized intersection over union: A metric and a loss for bounding box regression. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 658-666.

[24] Jadon, S. (2020). A survey of loss functions for semantic segmentation. In 2020 IEEE Conference on Computational Intelligence in Bioinformatics and

Computational Biology (CIBCB), Via del Mar, Chile, pp. 1-7. https://doi.org/10.1109/CIBCB48159.2020.9277638

[25] Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., Savarese, S. (2019). Generalized intersection over union: A metric and a loss for bounding box regression. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 658-666. https://doi.org/10.1109/CVPR.2019.00075

[26] Jadon, S. (2020). A survey of loss functions for semantic segmentation. 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), pp. 1-7. https://doi.org/10.1109/CIBCB48159.2020.9277638

[27] Mishra, M., Sarkar, T., Choudhury, T. et al. (2022). Allergen30: Detecting food items with possible allergens using deep learning-based computer vision. Food Anal. Methods. https://doi.org/10.1007/s12161-022-02353-9

[28] Arunachalaeshwaran, V.R., Mahdi, H.F., Choudhury, T., Sarkar, T., Bhuyan, B.P. (2022). Freshness classification of hog plum fruit using deep learning. 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1-6. https://doi.org/10.1109/HORA55278.2022.9799897

[29] Choudhury, T., Mahdi, H.F., Agarwal, A., et al. (2022). Quality evaluation in guavas using deep learning architectures: An experimental review. 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1-6. https://doi.org/10.1109/HORA55278.2022.9799824