International Journal on Emerging Technologies 11(3): 731-737(2020)

ISSN No. (Print): 0975-8364
ISSN No. (Online): 2249-3255

# Nuclei Segmentation in Cell Images using Fully Convolutional Neural Networks

*Siddharth Swarup Rautaray[1], Sohom Dey[2], Manjusha Pandey[3] and Mahendra Kumar Gourisaria[4]*
[1]*Associate Professor, School of Computer Engineering, KIIT Deemed to be University (Odisha), India.*
[2]*B.Tech Student, School of Computer Engineering, KIIT Deemed to be University (Odisha), India.*
[3]*Associate Professor, School of Computer Engineering, KIIT Deemed to be University (Odisha), India.*
[4]*Assistant Professor, School of Computer Engineering, KIIT Deemed to be University (Odisha), India.*

*(Corresponding author: Siddharth Swarup Rautaray)*
*(Received 10 March 2020, Revised 06 May 2020, Accepted 08 May 2020)*
*(Published by Research Trend, Website: www.researchtrend.net)*

**ABSTRACT: Nuclei detection in microscopy images is a major bottleneck in the discovery of new and effective drugs. Researchers need to test thousands of chemical compounds to find something of therapeutic efficacy. Nucleus being the most prominent part of a cell helps in the identification of individual cells in a sample and by analyzing the cell's reaction to various treatments the researchers can infer the underlying biological process at work. Automating the process of nuclei detection can help unlock cures faster and speedup drug discovery. In this paper, we propose a custom encoder-decoder style fully-convolutional neural network architecture with residual blocks and skip connections which achieves state-of-the-art accuracy. We use spatial transformations for data augmentation to prevent our model from over-fitting. We also use a combination of binary cross-entropy and dice-loss for our loss function which handles the class imbalance problem of the dataset. The proposed model is capable of segmenting nuclei effectively across a wide variety of cell types and experimental systems. The proposed model achieved a mean IOU of 94.86 on the BBBC038v1 dataset. Automated nuclei detection is projected to improve throughput for research in the biomedical field by saving researchers several hundred thousand hours of effort every year.**

**Keywords:** Artificial Intelligence, Biomedical Image Processing, Computer Aided Analysis, Medical Expert Systems, Neural Networks, Nuclei Detection, Cell Images.

## I. INTRODUCTION

Search for new and effective drugs require trial of thousands of chemical compounds and observing the reactions for each to arrive at an inference. For medical analysis, batches of cells are prepared and the reaction of the cells is observed after adding different chemical compounds to each batch of cells. Preparing batches of cells and testing with different chemicals can be done on a large scale after robotic automation replaced manual labor. A major delay in the pipeline is analyzing the huge amount of cell images for various characteristics, for which we certainly need software aid. The first and the most effective approach for cell analysis is most often the detection of the nuclei. From there various properties of the cell can be calculated to find out their disease state.

Let's explain the current pipeline followed by a scientist. When the nuclei are more-or-less round and easily distinguishable from each other, a classical computational algorithm can satisfactorily segment the nuclei. But the software tends to fail if the cell images are complex and involve tissue samples, because then it becomes hard to distinguish each nucleus as they have complicated shapes and are closer to each other, sometimes even overlapping. In these cases, the scientist has to analyze each sample by eye and this cost a significant amount of time and effort. Imagine manually analyzing thousands of images to arrive at a conclusion.

An accurate software model capable of nuclei identification in medical images without any arbitration will push the boundaries of biomedical image analysis and drug discovery and shorten the timespan to market a new drug. Classical image processing techniques require manual configuration, and existing models mainly specialize on specific types of cells. A single model intelligent enough to detect nuclei in different context and varying experimental system would save researchers a significant amount of time and effort and speed up the analysis by a huge margin.

Nuclei Detection from cell images is a segmentation problem, and U-Nets are known to excel in medical segmentation tasks. In this paper we propose a fully convolutional custom U-Net based architecture with redesigned encoder and decoder. We optimize a BCE-Dice loss function and train the model on the BBBC038v1 dataset with spatial data augmentation. This allows a lighter model to achieve higher accuracy

## II. RELATED WORKS

With the recent advancements in the artificial intelligence domain, neural networks are being widely used in medical image analysis and have proven to give better results than most classical image processing algorithms. Research in the field of biomedical image segmentation has become more demanding as more powerful neural architectures and deep learning techniques are emerging every year. In this section we discuss the recent advances in this field related to nuclei segmentation for cell analysis.

Nurzynska *et al.*, (2018) proposed a technique for searching the best parameters for color normalization for the task of segmenting the nucleus. Monte Carlo Simulation was used to search for the optimal parameters for color normalization which lead to better performance in segmentation [2].

Narotamo *et al.*, (2019) proposed a combined approach of using a Fast YOLO architecture and U-Net model for detection and segmentation respectively. The authors trained their model on 2D fluorescence microscopy images. They showed that their model is more computationally effective against Mask R-CNN while sacrificing some performance. Their proposed model is

9 times faster than Mask R-CNN on image size of 1388 × 1040 [3].

Chen *et al*., (2019) proposed a model for segmentation of Caudate Nucleus in MRI scans of brain based on a distance regularized level-set evolution [4].

Pan *et al*., (2019) proposed a model based on deep semantic network for segmentation of nuclei from pathological images. The authors used atrous depth wise separable convolution layers for their model (AS-UNet) which increases the receptive field of the model. It extracts and combines features of multiple scales so that the model can perceive both small and large cells. Their model achieves promising performance [5].

Mahbod *et al.,* (2019) proposed a U-Net architecture with two stages for segmentation of touching nuclei in sections of hematoxylin and eosin stained tissue. Semantic segmentation with U-Net was followed by the creation of a distance map with a regression U-Net model. Based on the segmentation mask and distance map a watershed algorithm is used for instance segmentation. Their model achieves a Jaccard index of 56.87% [6].

Zeng *et al.,* (2019) proposed a U-Net based model for nuclei segmentation which used residual blocks, and multi-scale feature and channel attention mechanism. Their model RIC-UNet achieves a Jaccard index of 0.5635 while the original U-Net achieves 0.5462 on the Cancer Genomic Atlas (TCGA) dataset [7].

Li *et al*., (2019) proposed a U-Net based model which utilizes boundary and region information, which provides a huge performance boost on overlapping glioma nuclei samples. They used a classification model to predict the boundary and the distance map is predicted by a regression model. These are further used to obtain the final segmentation mask. Their proposed architecture achieves a mean IOU of 0.59 on multi-organ nuclei segmentation open dataset (MoNuSeg) [8]. Sharma *et al*., (2020) have proposed a CNN based model for the prediction of paddy crop disease [18].

Zhou *et al*., (2019) proposed their (CIA-Net) for robust instance segmentation of nuclei. They used two separate decoders for separate tasks and a multi-level information aggregation module to captures the dependencies (spatial and texture) between the nuclei and the contour. This improved the generalizability of their model [9].

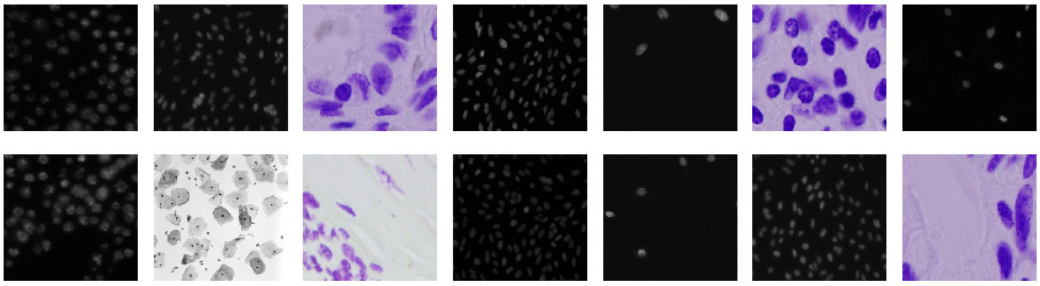

**Fig. 1.** Original Images.

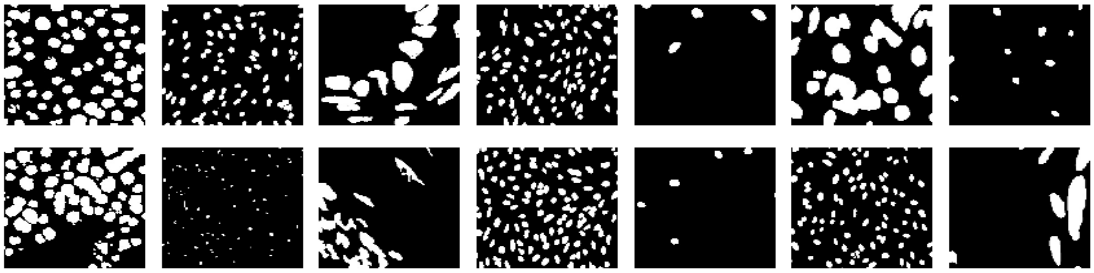

**Fig. 2.** Original Masks.

## III. PROPOSED METHOD

### *A. Dataset Used*

The BBBC038v1 dataset [1] is used for this experiment, which is accessible from Broad Bioimage Benchmark Collection Ljosa *et al*., Nature Methods, 2012]. The dataset contains 670 training images with more than twenty thousand annotated nuclei. The images were gathered from various sources including biomedical professionals in hospitals and industries and researchers in various universities. The dataset has a lot of variance as the cells belong to various animals and the imaging of the treated cells has been done in different experimental systems which involves variation in lighting conditions, microscope magnifications and histological stains.

### *B. Data Augmentation Used*

Deep-learning based approaches require a lot of input data, but it is difficult to find such huge amount of data in the medical field. The dataset we are using contains 670 images which are not sufficient for training a robust model, so we used specific data augmentation techniques to prevent our model from overfitting and make them generalize better and improve performance. In the case of medical images, spatial level transformations have already proven to give better results since they augment the data very close to real images. Especially elastic deformations and optical distortions work like charm while training a segmentation network. Shift and rotation invariances also work well with microscopy images. We used a lot of heavy augmentations: Horizontal Flip, Random Contrast, Random Gamma, Random Brightness, Elastic Transform, Grid Distortion, Optical Distortion, Shift Scale Rotate, etc.

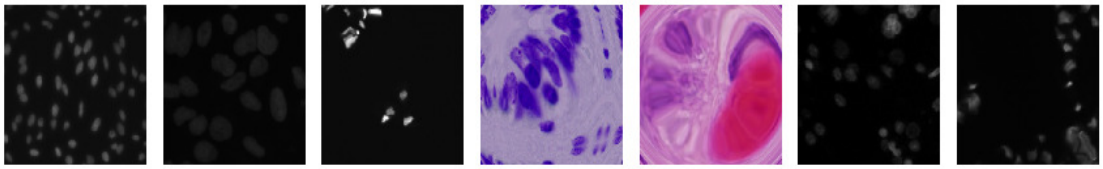

**Fig. 3.** Augmented Images.

**Fig. 4.** Augmented Masks.

*C. Model Architecture*

We used the semantic segmentation approach for our intended task of nuclei detection. Two of the most popular architectures in this domain are the Mask-RCNN [10] and the FCN [11] (Fully convolutional neural network) based segmentation net-works. FCN being a one-stage segmentation network is mostly preferred over two stage networks like Mask-RCNN for its simplicity and computational efficiency. The U-Net architecture [12] based upon the FCN architecture has been one of the most popular architecture for medical image segmentation recently. Our model is an improvement over the U-Net architecture.

FCN based segmentation networks replace the fully connected layers of a conventional CNN architecture with fully-convolutional layers [15-17]. It uses an encoder-decoder architecture to learn the segmentation mask from the input image. The encoder learns the contextual information and the decoder learns the spatial information. Skip connections help the decoder network to use the spatial information from the higher layers of the encoder network and fuses them with the up-sampled features to learn the precise location of the nuclei in the images. This method gives fine grained segmentation masks.

We use a 17 layers encoder network with residual blocks [13] which down-samples the feature map. We use convolution layers with a stride of 2 to down sample the images instead of using max-pooling. We only use max-pooling once at the beginning of the network. The decoder network uses transposed convolution layers to up-sample the feature maps, then concatenates features from encoder layers through skip connections, followed by residual blocks in each stage. Residual blocks allow easier optimization of deep networks while simple skip connections from encoder to decoder enable fine grained segmentation maps to be generated using in-formation from the previous layers of the encoder.

The resnet blocks helps in the gradient flow and thus allows us to make the UNet deeper. This in turn allows the model to learn better features. In this proposed method, we also use resnet blocks in the decoder network which increased the accuracy. The combination of Binary Cross entropy and Dice loss allows this deep model to be trained without being affected by the class imbalance. The spatial data augmentation prevents the model from being overfitted and helps the model to generalize better.

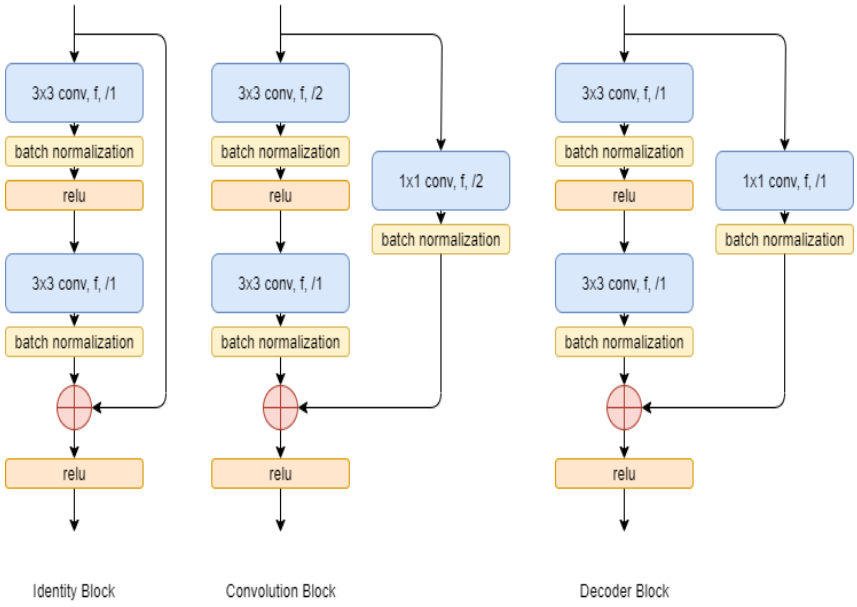


**Fig. 5.** Residual Bottleneck Blocks.

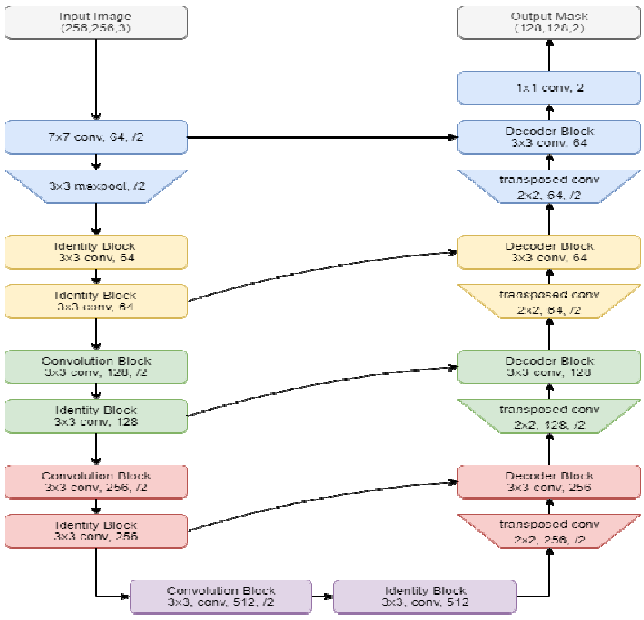


**Fig. 6.** Model Architecture.

**Table 1: Encoder Details.**

| Layers | | | Filters | Input Shape | Output Shape |
|---|---|---|---|---|---|
| Input | | | — | 256, 256, 3 | |
| **Stage 1** | Layer 1 | 7 × 7 conv | 64 | 256, 256, 3 | 128,128, 64 |
| | Layer 2 | 3 × 3 maxpool | — | 128, 128, 64 | 64, 64, 64 |
| **Stage 2** | Block 1 | 3 × 3 conv<br>3 × 3 conv | 64 | 64, 64, 64<br>64, 64, 64 | 64, 64, 64<br>64, 64, 64 |
| | Block 2 | 3 × 3 conv<br>3 × 3 conv | 64 | 64, 64, 64<br>64, 64, 64 | 64, 64, 64<br>64, 64, 64 |
| **Stage 3** | Block 1 | 3 × 3 conv, /2<br>3 × 3 conv | 128 | 64, 64, 64<br>32, 32, 128 | 32, 32, 128<br>32, 32, 128 |
| | Block 2 | 3 × 3 conv<br>3 × 3 conv | 128 | 32, 32, 128<br>32, 32, 128 | 32, 32, 128<br>32, 32, 128 |
| **Stage 4** | Block 1 | 3 × 3 conv, /2<br>3 × 3 conv | 256 | 32, 32, 128<br>16, 16, 256 | 16, 16, 256<br>16, 16, 256 |
| | Block 2 | 3 × 3 conv<br>3 × 3 conv | 256 | 16, 16, 256<br>16, 16, 256 | 16, 16, 256<br>16, 16, 256 |
| **Stage 5** | Block 1 | 3 × 3conv,<br>/2<br>3 × 3 conv | 512 | 16, 16, 256<br>8, 8, 512 | 8, 8, 512<br>8, 8, 512 |
| | Block 2 | 3 × 3 conv<br>3 × 3 conv | 512 | 8, 8, 512<br>8, 8, 512 | 8, 8, 512<br>8, 8, 512 |

**Table 2: Decoder Details.**

| Layers | | | Filters | Input Shape | Output Shape |
|---|---|---|---|---|---|
| **Encoder Output** | | | — | 8, 8, 512 | |
| **Stage 6** | Layer 1 | transposed conv | 256 | 8, 8, 512 | 16, 16, 256 |
| | Layer 2 | Concatenate ( [Stage 4 Block 2, Stage 6 Layer 1] ) | | | |
| | Block 1 | 3 × 3 conv<br>3 × 3 conv | 256 | 16, 16, 256 | 16, 16, 256 |
| **Stage 7** | Layer 1 | transposed conv | 128 | 16, 16, 256 | 32, 32, 128 |
| | Layer 2 | Concatenate ( [Stage 3 Block 2, Stage 7 Layer 1] ) | | | |
| | Block 1 | 3 × 3 conv<br>3 × 3 conv | 128 | 32, 32, 128 | 32, 32, 128 |
| **Stage 8** | Layer 1 | transposed conv | 64 | 32, 32, 128 | 64, 64, 64 |
| | Layer 2 | Concatenate ( [Stage 2 Block 2, Stage 8 Layer 1] ) | | | |
| | Block 1 | 3 × 3 conv<br>3 × 3 conv | 64 | 64, 64, 64 | 64, 64, 64 |
| **Stage 9** | Layer 1 | transposed conv | 64 | 64, 64, 64 | 128, 128, 64 |
| | Layer 2 | Concatenate ( [Stage 1 Layer 1, Stage 9 Layer 1] ) | | | |
| | Block 1 | 3 × 3 conv<br>3 × 3 conv | 64 | 128, 128, 64 | 128, 128, 64 |
| **Stage 10** | Layer 1 | 1 × 1 conv | 2 | 128, 128, 64 | 128, 128, 2 |
| Outputs | | | — | 128, 128, 2 | |

*D. Lost Function Used*

The most commonly used loss function for segmentation models is pixel wise cross-entropy loss which compares the class predictions for each pixel individually. Another very popular loss function used in biomedical image segmentation in soft-dice loss [14] which measures the overlap between two samples. For our task we optimize a BCE-Dice loss function which is basically binary cross-entropy added to soft-dice loss, which resulted in better performance and early convergence. The proposed model took only 60 epochs of training before early stopping.

*Binary Cross-Entropy Loss:*

$$-(ylog(p) + (1-y)\log(1-p)) \quad (1)$$

*Soft-Dice Loss:*

$$\frac{2|A \cap B|}{|A|+|B|} \quad (2)$$

*BCE-Dice Loss:*

$$-(ylog(p) + (1-y)\log(1-p)) + \frac{2|A \cap B|}{|A|+|B|} \quad (3)$$

*E. Evaluation Metric Used*

The most commonly used evaluation metrics are: pixel-wise accuracy and the Jaccard index also known as the mean IOU. We used mean IOU as our evaluation metric which calculates the overlap between the target and prediction masks. We also chose this metric since it is closely related to dice coefficient used in the dice loss.

$$\text{IOU} = \frac{\text{Target} \cap \text{Prediction}}{\text{Target} \cup \text{Prediction}} \quad (4)$$

We also use other famous metrics mostly used for classification tasks like: Precision, Recall and F1-score to evaluate our model to get more insight about the performance of our model.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5)$$

Precision tells us that, of all the pixels classified as belonging to a nucleus by the model, how many actually belonged to a nucleus.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (6)$$

Recall tells us that, of all the pixels that actually belonged to a nucleus how many did the model classify as belonging to a nucleus.

$$F1 - \text{score} = \frac{2*(\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}} \quad (7)$$

F1 score is the combination of precision and recall and is a better metric to judge classification tasks with imbalanced classes. A high F1-score means lower false positives and lower false negatives predicted by the model. This implies that the model is correctly classifying pixels in the nuclei region as 1 and pixels not in the nuclei region as 0.

*F. Algorithm*

**Table 3: Results.**

| Metric | Value |
|---|---|
| Precision | 0.9734 |
| Recall | 0.9738 |
| F1-score | 0.9736 |
| IoU | 0.9486 |

**Table 4: Comparison.**

| Method | IoU |
|---|---|
| U-Net | 90.77 |
| Wide U-Net | 90.92 |
| UNet++ | 92.63 |
| Our Model | 94.86 |

**Algorithm1: Model Training**

1: Build computational graph (model)
2: Initialize model weights
3: For I=0 to n_epochs do
4: Forward propagate through the computational graph to calculate prediction (p)
5: Calculate loss:

$$-(ylog(p) + (1-y)\log(1-p)) + \frac{2|y \cap p|}{|y|+|p|}$$

6: Calculate accuracy metrics:

$$\frac{y \cap p}{y \cup p}$$

7: Append loss in dictionary History['loss'] ← Loss
8: Append metric in dictionary History['metric] ← IOU
9: Back propagate gradients through the computational graph
10: Update weights:

$$w = w - \alpha \frac{\partial L}{\partial w}$$

11: If (Loss did not improve since last 10 epoch)
12: Break Loop
13: End If
14: If (Loss did not improve since last 5 epoch)
15: α = α x 0.1
16: End If
17: End For
18: Return Model, History

## IV. EXPERIMENTS AND RESULTS

We resize the input images to 256 × 256 before feeding them into the network. Our network outputs masks of dimension 128 × 128. Since the proposed model is considerably deep, we use data augmentation to prevent overfitting thus increasing the generalizability of the model and improve overall performance. We used Adam optimizer and auto-reduced the learning-rate when the learning plateaued out. The model reached a validation IOU of 0.9486 with just 25 epochs of training before being early-stopped. Using SGD optimizer gives a smooth training curve but takes 500 epochs to converge, while Adam takes 25 epochs but the initial training curve is quite abrupt. Fig. 7 and 8 shows the IOU and loss function curves for the training and validations sets for both the optimizers. Table 3 shows the results on the validation set and Table 4 compares our model with the top 3 state-of-the-art models for this specific task, and our model performs significantly better.

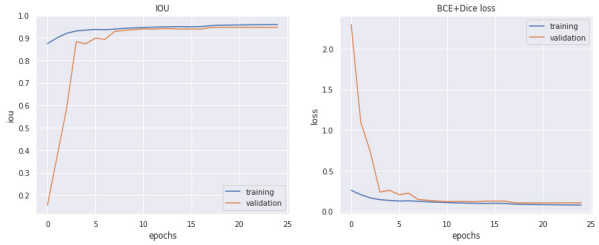


**Fig. 7.** Accuracy metric and loss function plots for training and validation using adam optimizer.

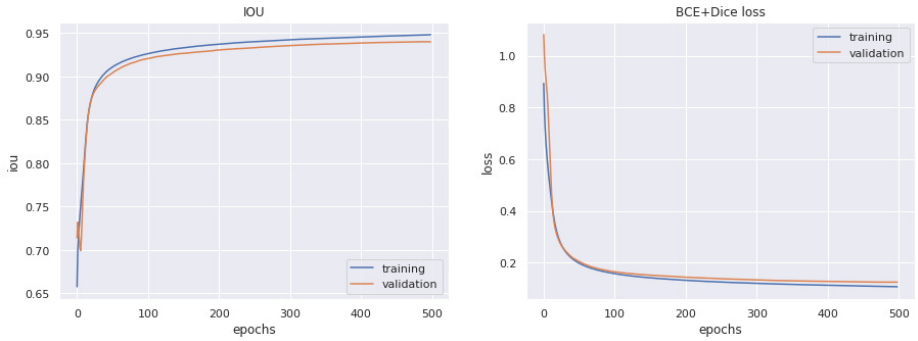


**Fig. 8.** Accuracy metric and loss function plots for training and validation using sgd optimizer.

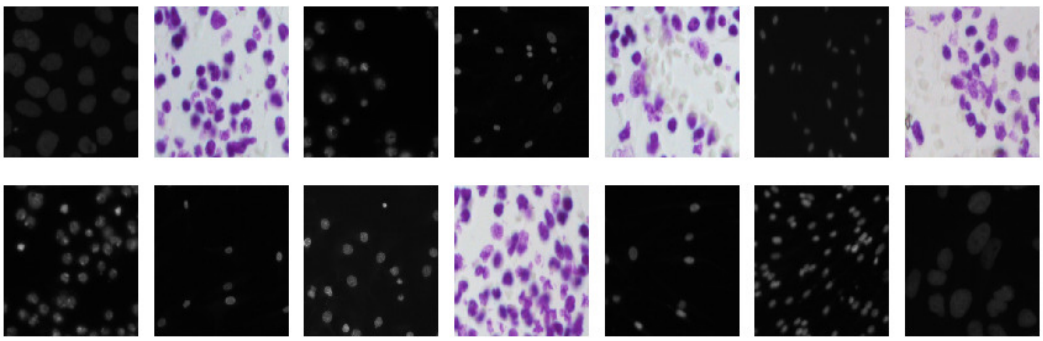

**Fig. 9** Testing Images.

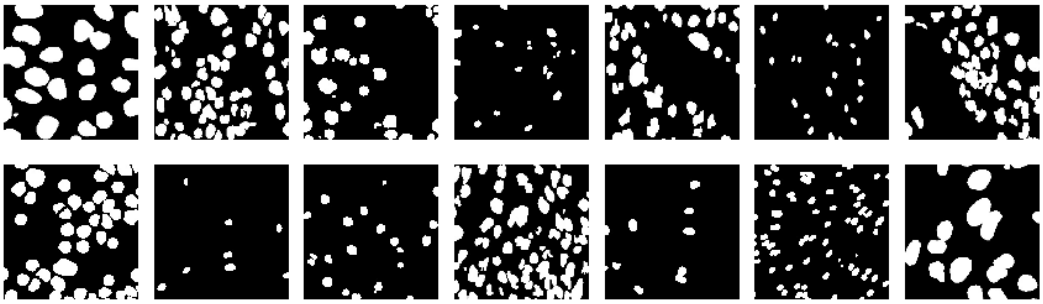

**Fig. 10.** Predicted Images.

## V. CONCLUSION

Medical image processing has been gaining a lot of attention recently due to the emergence of deeper and high-accuracy segmentation networks which can compete against humans and speed-up biomedical research to a great extent. Nuclei detection has always been a very a crucial step for cell analysis and recently many computer aided analysis approaches are being used for faster and more accurate medical analysis. With the inception of deep learning based intelligent analysis algorithms, medical industry and researchers are replacing classical computational image processing algorithms with sophisticated deep learning models. Unlike classic image processing algorithms deep learning models do not require manual pre-processing or feature engineering, nor do they require any manual parameter tweaking. In this paper the proposed model incorporates the latest advancements in the field of deep learning for accurate segmentation of nuclei from microscopy images of cells. It achieves an IOU of 0.9486 which is a significant improvement over the state-of-the-art U-Net++ network. The proposed model works effectively across a wide variety of types of nuclei and experimental systems. Robustness to cell types and experimental setups has been our main focus. Tackling the problem of automated nuclei detection can help to improve the rate of drug discovery and enable faster cures thus improve overall health and quality of life of the people.

## VI. FUTURE SCOPE

We plan to extend the method to include instance segmentation of the nuclei and nuclei counting. Instance segmentation helps in identifying each cell uniquely and cell counting is important for various medical diagnosis. This will further help in faster and better cell detection and experimentation. We also plan to extend this method to incorporate cell tracking.

**Conflict of Interest.** None of the authors have any conflict of interest of any kind.

## ACKNOWLEDGEMENT

We would like to thank Broad Institute for curating the BBBC038v1 dataset and encouraging us to use it for training and testing new algorithms for medical image analysis. The variety in the images helped us to train a robust model for nuclei-detection.

## REFERENCES


[1]. Broad Bioimage Benchmark Collection dataset Broad Bioimage Benchmark Collection dataset page https://data.broadinstitute.org/bbbc/BBBC038

[2]. Nurzynska, K. (2018). Optimal parameter search for colour normalization aiding cell nuclei segmentation. In *International Conference: Beyond Databases, Architectures and Structures*, 349-360.

[3]. Narotamo, H., Sanches, J. M., & Silveira, M. (2019). Segmentation of Cell Nuclei in Fluorescence Microscopy Images Using Deep Learning. In *Iberian Conference on Pattern Recognition and Image Analysis*, 53-64.

[4]. Chen, Y., Chen, G., Wang, Y., Dey, N., Sherratt, R. S., & Shi, F. (2019). A Distance Regularized Level-Set Evolution Model Based MRI Dataset Segmentation of Brain's Caudate Nucleus. *IEEE Access*, *7*, 124128-124140.

[5]. Pan, X., Li, L., Yang, D., He, Y., Liu, Z., & Yang, H. (2019). An Accurate Nuclei Segmentation Algorithm in Pathological Image Based on Deep Semantic Network. *IEEE Access*, *7*, 110674-110686.

[6]. Mahbod, A., Schaefer, G., Ellinger, I., Ecker, R., Smedby, Ö., & Wang, C. (2019). A Two-Stage U-Net Algorithm for Segmentation of Nuclei in H&E-Stained Tissues. In *European Congress on Digital Pathology*, 75-82.

[7]. Zeng, Z., Xie, W., Zhang, Y., & Lu, Y. (2019). RIC-Unet: An improved neural network based on Unet for nuclei segmentation in histology images. *IEEE Access*, *7*, 21420-21428.

[8]. Li, X., Wang, Y., Tang, Q., Fan, Z., & Yu, J. (2019). Dual U-Net for the Segmentation of Overlapping Glioma Nuclei. *IEEE Access, 7*, 84040-84052.

[9]. Zhou, Y., Onder, O. F., Dou, Q., Tsougenis, E., Chen, H., & Heng, P. A. (2019). CIA-net: Robust nuclei instance segmentation with contour-aware information aggregation. In *International Conference on Information Processing in Medical Imaging*, 682-693.

[10]. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN, in arXiv:1703.06870 [cs.CV].

[11]. Long, J., Shelhamer E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA,* 3431-3440.

[12]. 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*, 234-241.

[13]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and Pattern Recognition,* 770-778.

[14]. Ramachandran, P., Zoph, B., & Le, Q. V. (2017). Searching for activation functions. *arXiv preprint arXiv:1710.05941*, 1-13

[15]. Milletari, F., Navab, N., & Ahmadi S. A. (2016). V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *Fourth International Conference on 3D Vision (3DV)*, 565–571.

[16]. Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. Lecture Notes in Computer Science, vol 11211. Springer, Cham

[17]. Gourisaria, M. K., Das, S., Sharma, R., Rautaray, S. S., & Pandey, M. (2020). A Deep Learning Model for Malaria Disease Detection and Analysis using Deep Convolutional Neural Networks. *International Journal on emerging Technologies, 11*(2), 699-704.

[18]. Sharma, R., Das, S., Gourisaria, M. K., Rautaray, S. S., & Pandey, M. (2020). A Model for Prediction of Paddy Crop Disease using CNN. 2nd International Conferences on Computing analytics and Networking. In Progress in Computing, *Analytics and Networking,* 533-543.



**How to cite this article:** Rautaray, S. S., Dey, S., Pandey, M. and Gourisaria, M. K. (2020). Nuclei Segmentation in Cell Images using Fully Convolutional Neural Networks. *International Journal on Emerging Technologies*, *11*(3): 731–737.