SEGMENTATION AND QUANTIFICATION OF NUCLEI IN FLUORESCENT IMAGES USING DEEP LEARNING TECHNIQUE

by

Aishwarya Shrestha

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in Computer Science

at

The University of Wisconsin-Milwaukee

December 2023

ABSTRACT

SEGMENTATION AND QUANTIFICATION OF NUCLEI IN FLUORESCENT IMAGES USING DEEP LEARNING TECHNIQUE

by

Aishwarya Shrestha

The University of Wisconsin-Milwaukee, 2023
Under the Supervision of Professor Susan McRoy

In the study of bioimage analysis, the segmentation of cell nuclei is the pivotal stage. The identification of the nucleus aids researchers in comprehending the underlying processes for drug discovery and detection of cancerous cells. Nucleus segmentation is a challenging problem due to the presence of overlapping nuclei, image intensity heterogeneities, and image noise. Furthermore, performing manual segmentation at the pixel level poses a major obstacle due to its time-consuming nature, the need for expert professionals, and its susceptibility to errors. To overcome these shortcomings, we provide a pipeline to perform automatic nucleus segmentation and quantification for fluorescent images using deep learning. In this study, to perform nucleus segmentation, we modify the U-Net model by substituting its encoder with the EfficientNet model while retaining the U-shaped structure unaltered. This adaptation in feature extraction encoder path has shown U-Net to perform better in creating fine-grained segmentation map. The proposed pipeline delivers notable results, with an F1-score of 87% and an Intersection over Union (IoU) of 80%. A post-processing step was employed to enable morphological quantification for each segmented nucleus. Additionally, we present novel datasets with annotated boundaries for over 3000 nuclei, tailored for binary segmentation tasks. Therefore, the work presented in this study demonstrates the potential for automatic nuclei extraction from fluorescent images.

To my parents,

for your constant love and support.

© Copyright by Aishwarya Shrestha, 2023
All Rights Reserved

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2D	Two dimensional
3D	Three dimensional
AI	Artificial Intelligence
BBBC	Broad Bioimage Benchmark Collection
CNN	Convolution Neural Network
DCNN	Deep Convolution Neural Network DCNN
DL	Deep Learning
FCN	Fully Convolutional Network
FL	Focal Loss
FPN	Feature Pyramid Network
GPU	Graphics processing unit
IoU	Intersection over Union
KDSB18	Kaggle data science bowl 2018
MBConv	Mobile Inverted Convolution
PSPNet	Pyramid Scene Parsing Network
SENet	Squeeze-and-Excitation Networks

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Prof. Susan McRoy, who supported me throughout my program. I am thankful for her time and valuable feedback. In addition to my advisor, I am grateful for the guidance of Prof. Qingsu Cheng who assisted me as a mentor. I appreciate his consistent dedication, patience, and motivational support throughout my research journey. Without his encouragement, I would not have been able to complete my research. I am also thankful to the Department of Biomedical Engineering at UWM (University of Wisconsin Milwaukee) for providing the experimental data used in this study. I would also like to thank Prof. Rohit Kate for agreeing to be my thesis committee member. Furthermore, I acknowledge the funding support from NASA, grant #80NSSC23K0989, which made this research possible.

Finally, I would like to thank my parents and brother for their unwavering love and support throughout my academic journey. Thank you for your guidance, trust, and everything you have done for me.

Chapter 1

1 Introduction

1.1 Problem Statement

In the study of microscopic images, segmenting cell nuclei is one of the most important tasks in biomedical image analysis. For various image analysis workflow, nuclei segmentation is the first step for extracting quantitative biological signals [1]. Accurate cell nuclei identification helps researchers comprehend the underlying process. Research studies based on nuclei segmentation such as cell counting, cell tracking, morphological analysis, and 3D reconstruction have been performed for disease diagnosis, assessment of cellular health, drug discovery, and evaluation of therapeutic efficacy. However, pixel-by-pixel manual segmentation of these images presents significant challenges due to its labor-intensive nature, need for expert knowledge, and being prone to errors. Additionally, computer vision-based semantic segmentation of cell nuclei poses a challenge due to large intensity inhomogeneities in the foreground and the background, unconventional morphology of nuclei instances, the presence of noise in the image, and instances of touching nuclei that require separation. To overcome these shortcomings, this research concentrates on developing a pipeline to perform automatic nucleus segmentation and quantification.

Deep learning techniques, particularly convolutional neural networks (CNNs), have gained significant attention in recent years, particularly in the field of image analysis [2, 3]. Advancement in deep learning techniques has enabled automated segmentation. These automation have revolutionized biological image analysis workflows, which have been widely applied in various domains such as medical image processing for segmentation of retinal blood vessels, phenotype

extraction, microscopy-based high-content screening, and anatomical/cell structure detection [2, 4–7]. Deep learning algorithms can provide accurate mappings and optimized performance due to the capability of dealing with diverse images. Therefore, this study explores the potential use of a deep learning technique to identify cell nuclei in fluorescent images for streamlining the segmentation process and saving researchers' time. The deep learning method requires a significant number of annotated labels, which is costly and time-consuming to generate. To address this challenge, we leverage the encoder-decoder-based U-Net [8] architecture, which yields favorable outcomes even with a limited training dataset through the utilization of data augmentation techniques [9]. The approach in this study involves connecting the compound scaled EfficientNet [10] as the encoder for the feature extraction stage with the U-Net decoder for reconstructing the fine-grained segmentation map. This modification results in a more effective U-Net model with enhanced encoder and decoder capabilities.

Furthermore, to address the issue associated with touching nuclei regions, we implement a post-processing step using the watershed algorithm [11]. The objective of this post-processing step is to effectively separate touching instances, assigning a unique label to each identified cell, and providing morphological measurements for each nucleus. Additionally, in this study, we have generated an annotated dataset that contains fluorescence microscopy images of 2D breast mammary gland cells with DAPI stain, along with corresponding ground truth images. This novel dataset can provide valuable training data for segmentation algorithms, eliminating the need for laborious manual tracing.

1.2 Literature review

The segmentation of nuclei offers valuable insights into the morphology of nuclei, the content of DNA, and the condensation of chromatin. Multiple biomedical analyses, including cell quantification [12], and cell phenotype analysis [13] can be performed based on accurate cell nuclei segmentation. The labor-intensive nature and reliance on the subjective judgment of human operators make manual segmentation impractical when accurate detection of a large number of nuclei is essential [1]. To overcome this issue, there are research efforts focused on developing segmentation techniques with deep learning and computer vision approaches [3].

The classical computer vision-based approaches employed for performing nucleus segmentation include Otsu's thresholding followed by watershed algorithm [14], K-means clustering and region growing [15], graph-cuts based methods, active contours, level-set algorithm [16, 17]. The dominant approaches for nuclei segmentation like Fiji/Image [18], CellProfiler [19], and Icy [20]are based on the tradition approaches [2]. However, these approaches are sensitive to noise, display high specificity for particular image types, and commonly demand the fine-tuning of manual settings [21].

In the recent past, algorithms based on deep learning and convolutional neural networks (CNN) have demonstrated state-of-the-art performance in the segmentation of medical imaging, as they exhibit capability to autonomously learn parameters, and demonstrate strong generalization capacity. Various international competitions [1, 22–24] that have been held to find potential solutions based on deep learning. However, a genuinely general solution is still awaited [1]. U-Net emerged as the winner of the ISBI bioimage segmentation challenge stands out as one of the most utilized architectures in biomedical image segmentation, particularly in the context of cell nuclei. U-Net is recognized for achieving effective segmentation results with a reduced number of training

dataset, primarily relying on extensive data augmentation of annotated samples. U-Net performs semantic segmentation by classifying individual pixels; however, it lacks the capability to classify entire objects (sets of pixels). Considering the necessity for distinct identification of each nucleus in nuclei segmentation, our study introduces a pipeline comprising an improved U-Net model and a subsequent postprocessing step to effectively address this challenge.

Table 1: Summary of nuclei segmentation works based on variation of U-Net.

Reference	Paper details	
Zeng, Xie [1][25]	Model	Residual Inception Channel Attention-UNet It has a greater ability to differentiate between slight differences in colors within deeper or more complex backgrounds
	Datasets	Images of cells from seven different organs (breast, kidney, liver, prostate, bladder, colon, and stomach) were captured from The Cancer Genome Atlas (TCGA) dataset [9].
	Limitations	Challenging to delineate when the nuclei and boundaries are not clearly discernible.
Mahbod, Schaefer [2]	Model	Two-Stage U-Net
	Datasets	TCGA dataset [9]
	Limitations	The suggested model is inadequate for accurately segmenting complex shapes.
Kong, Genchev [3]	Model	Two-Stage Stacked U-Nets (SUNets) SUNets combine four parallel backbone networks through attention generation mechanisms.
	Datasets	The Cancer Genome Atlas (TCGA) [9] and Triple-Negative Breast Cancer (TNBC) Dataset [26]
	Limitations	Huge number of training parameters
[27]	Model	Pyramid Scene Parsing with SegNet (PSPSegNet)

		U-Net and SegNet share a similar architecture, SegNet reuse pooling indices.
	Datasets	Multi-Organ Nuclei Segmentation (MoNuSeg) [24]
	Limitations	Model relies on synthetic training data, as a result it may not achieve highly precise representations of cell shapes.
Jha et al. [28]	Model	Double U-Net Utilizes U-Net at the network's base to effectively gather additional semantic information
	Datasets	Multiple dataset
	Limitations	It has greater number of parameters compared to U-Net, leading to an increase in training time.

Table 1 summarizes the results of my review of the literature related to nuclei segmentation approaches using variations of the U-Net model. The table comprises different models used, datasets utilized and limitations of current research. Based on this review, it is evident that only a limited number of articles have explored nuclei segmentation in fluorescent images, and a restricted number of researchers have employed an automatic end-to-end deep learning-based method for segmentation and quantification. To address this gap, our research focuses specifically on this area, deliberately avoiding reliance on publicly available artificially generated training data to ensure a more focused and contextually relevant approach to attain precise representations of cell shapes.

Furthermore, segmentation tools in the field of biology often leads to notable segmentation errors. This happens because the underlying algorithms are based on certain assumptions about how things should be, and these assumptions embedded in the computational design of existing algorithms, which might not always match what happens in real situations. The dominant approaches to nucleus segmentation were initially developed when the biological systems and

studied phenotypes were generally less complex. The ideal method for nuclei segmentation would be generic to any binary segmentation samples and provide a fully automated solution. Therefore, for our study we concentrate on fluorescent images to deliver an automated pipeline for nuclei segmentation.

1.3 Thesis outline

The remainder of this thesis is divided into four chapters. Chapter two reviews the background of tools and techniques employed in this study. Chapter three describes the procedure for generating predicted output for the nuclei segmentation task using the proposed method. In the fourth chapter, we present experimental results and discussions. The concluding part of the thesis, followed by future work, is covered in the final section.

Chapter 2

2 Background

2.1 Data Collection

The dataset used in this research was obtained from the Department of Biomedical Engineering at UWM (University of Wisconsin-Milwaukee). The 3D data were collected using Zeiss LSM 710 imaging system equipped with a Zeiss Apochromat 40X/1.1 (0.8mm working distance) water immersion objective lens originating from Cheng et al.'s research [29]. Advanced imaging system utilized in this research helps to accurately assess mammary gland cultures under varying environmental stress conditions within a controlled laboratory environment. The data were generated for investigating the effects of radiation exposure in the context of breast cancer research and treatment. The resulting data formed the basis for the dataset used in this study. We present a new dataset called the 'Breast Mammary Gland Dataset (BMGD)' which includes manually annotated ground truth (labelled image). Therefore, we provide a comprehensive description of the methods used for construction and the annotation process in Chapter 3 Methods. The dataset consists of 547 images with 3000 over cell nuclei. Each image includes a corresponding manually annotated ground truth mask, which provides precise delineation of the objects of interest as shown in **Figure 1**.

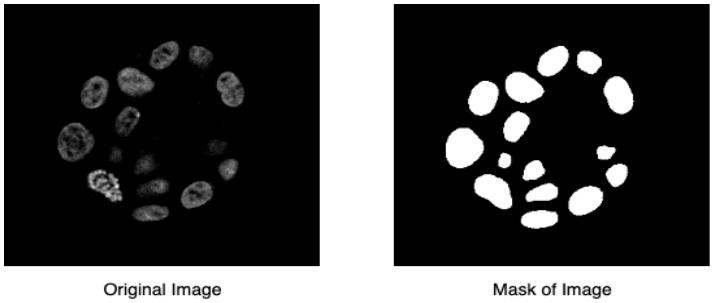


Figure 1. Demonstration of original image and ground truth mask

To assess the efficacy of model trained in our dataset, we predict and discuss the comparative analysis with two other open-source data, namely Kaggle Data Science Bowl 2018 (KDSB18) [1], and Broad Bioimage Benchmark Collection (BBBC022) [30] dataset. These datasets are considered as the benchmark for studies on nuclei segmentation [31]. KDSB 2018 consists of data collected from different research labs and includes annotated histopathological images that display nuclei. BBBC022 consists of U20S cells captured with fluorescence microscopy.

2.2 Deep Learning Techniques

2.2.1 Binary Semantic Segmentation

In binary semantic segmentation, the objective is to categorize individual pixels within an image as either belonging to a specific class or not. It involves dividing the image into two distinct classes. In this research, the primary goal is to process an image and divide it into two the foreground, and the background. Foreground represents the nuclei cells of interest, and the background encompasses the rest of the image. Binary segmentation techniques help to automate the process of diagnosis and analysis, making it easier for researchers and healthcare professionals to extract meaningful information from medical images [32].

2.2.2 Neural Network

A neural network is a computer-based system that learns to perform a task by analyzing training examples. Typically, these examples are manually labeled. This network consists of input layer, one or more hidden layer, and output layer. Each node connects to another and has an associated weight and threshold. In the network, inputs are introduced through the initial layer and subsequently traverses hidden layers, where it undergoes complex multiplication and addition

operations to acquire the ability to make predictions. The number of hidden layers depends on the complexity of the network. Throughout the training process, weights and thresholds are continuously adjusted to minimize the error until it reaches the output layer.

2.2.3 Deep Learning

Deep learning algorithms aim to leverage the hidden patterns within the input data distribution to discover meaningful representations, at multiple hierarchical levels, where higher-level learned features are expressed in relation to lower-level features as stated by Bengio [33]. Deep-learning algorithm consist of neural network with more than three layers, including the input and output. Traditional machine learning methods perform decision making based on feature extraction. Consequently, to determine segmentation mask, algorithm needs to guess which nuclei are important and design complex algorithm to capture them. In contrast, deep learning algorithm automatically learn and extracts relevant features from the data during the training phase. This helps to streamline the process and improves accuracy.

2.2.4 Convolution Neural Network

Deep learning techniques, such as Convolutional Neural Networks (CNNs) improve the prediction performance by using big data and abundant computational resources. CNN architecture consists of a convolutional layer, a pooling layer, and fully connected layers, making them well-suited for processing grid-like data, such as images. Convolution is a mathematical operation involving two functions, resulting in a third function that describes how the shape of one function is changed by the other [34]. Similarly, in CNN we slide filter/kernel (3×3 matrix) and input image to extract local features.

EfficientNet Model

EfficientNet incorporates CNNs as a fundamental building block. EfficientNet employs a compound scaling approach where it uniformly adjusts the ConvNet model's width, height, and image resolution while maintaining a constant ratio. The architecture consists of multiple blocks, and each block is composed of several sub-blocks known as MBConv blocks, which were initially introduced by the creators of MobileNetV2 [35]. An MBConv block comprises a sequence of Expand, DepthwiseConv, Squeeze, Reshape, Expand, and Excite layers. A visual representation of the MBConv concept is depicted in **Figure 2**. The family of EfficientNet networks has a different number of these MBConv blocks. They first developed a base architecture, EfficientNetB0, from which they derived eight variants, labeled as EfficientNetB0 through EfficientNetB7, using compound scaling. As we go from EfficientNetB0 to EfficientNetB7, depth, width, resolution, and model parameter goes on increasing [36]. Furthermore, the model architecture uses the Squeeze-and-Excitation (SE) optimization to further enhance the model's performance.

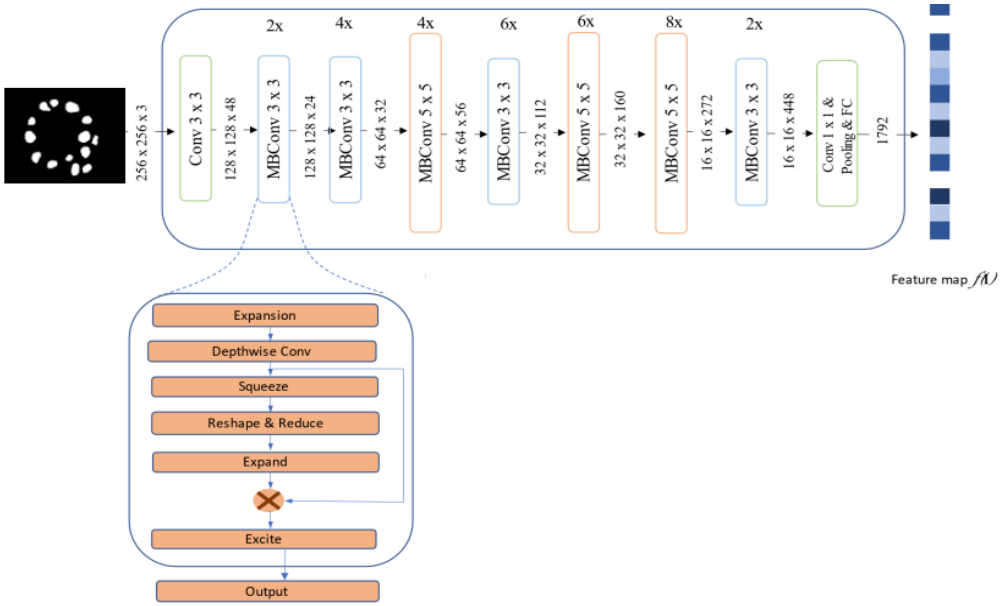


Figure 2. General architecture diagram of EfficientNet model [36]

2.2.5 Encoder-decoder architecture

In encoder-decoder based architecture [37], the process involves encoding the input data into a compact lower-dimensional representation and subsequently decoding it to reconstruct the original data. This model comprises three main components: 1. Encoder, 2. Bottleneck, and 3. Decoder. Encoder handles the input data by progressively reducing spatial dimensions while extracting high-level features. This reduction is accomplished by using a sequence of convolutional layers, pooling layers, and dropout layers. Following the encoder there is a bottleneck (connecting point) that combines the extracted features for creating lower-dimensional representation of the input image while preserving essential information. **Figure *3*** illustrates the architectural diagram of the encoder-decoder model.

Conversely, the decoder component takes this compact representation to enlarge it back to its original dimensions through the help of up-sampling layers. Decoder is symmetric to encoder with the layers gradually expanding the dimensionality of the data until it matches the original input's dimensionality.

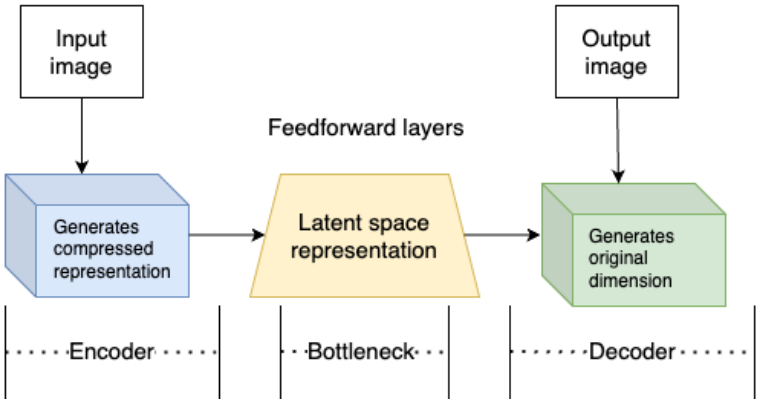


Figure 3. Architecture of encoder-decoder model with bottleneck layer.

2.2.6 Encoder-Decoder Based Architectures for Deep Convolutional Neural Networks

In this section, we present some state-of-art encoder-decoder-based architectures employed for our comparative analysis.

U-Net

The U-Net architecture[8] resembles an encoder-decoder architecture which contains convolutional layers on both encoder and decoder. It is developed by Olaf Ronneberger et al. for Biomedical Image Segmentation at the University of Freiburg, Germany. This optimized semantic segmentation network is designed to learn from fewer training samples. It is an improvement over the existing fully convolutional networks [38] for semantic segmentation.

U-Net is a U-shaped convolutional network that uses skip connections to preserve features at different resolutions. The architecture consists of a contracting path (encoder) to capture context and a symmetric expanding path (decoder). The encoder part consists of a stack of convolution, activation and pooling layers that progressively reduce the spatial dimensions of the input image while extracting features. The decoder path mirrors the encoder path in reverse, gradually up sampling the feature maps and combining them with skip connections to capture fine-grained details and reconstruct the segmented output. The skip connections between layers of the encoding branch and decoding branch provide the decoder with access to spatial information to reason about up sampling the segmentation.

Feature Pyramid Network (FPN)

Feature Pyramid Network [39] is a deep CNN architecture proposed in 2017 from Facebook AI Research team FNP presents simple framework to build feature pyramids inside the convolutional

neural network (CNN). Unlike U-Net, it makes prediction at each layer, therefore, it reuses the multi-scale feature maps from different layers computed in the forward pass. Yet, it fails to take advantage of reusing the higher-resolution maps within the feature hierarchy, leading to a missed opportunity for detecting small objects.

LinkNet

LinkNet [40] is another deep CNN architecture with a pretrained encoder in the network. The LinkNet architecture employs a residual layer in each of the encoder and decoder blocks to dissect the image, reconstruct it, and then route it through a set of concluding convolutional layers. It is suitable to perform segmentation in real time.

Pyramid Scene Parsing Network (PSPNet)

Pyramid Scene Parsing Network (PSPNet) [41] was proposed to address challenges related to context information in semantic segmentation. It incorporates a pyramid pooling module, which captures contextual information at different scales. This module allows the network to consider both local and global contextual information.

2.2.7 Transfer Learning

The typical approach for training network involves starting from randomly initialized weights. This approach is mathematically expensive and requires a large amount of data. Conversely, transfer learning is a widely adopted technique in computer vision where network weights are initialized with models exhibiting superior performance. One of the approaches to perform transfer learning is to use pre-trained models which are trained on extensive benchmark datasets and serve

as a valuable starting point for addressing new, dissimilar tasks. Therefore, we pretrained with ImageNet dataset as the backbone for our deep convolutional neural network (CNN) models

2.3 Annotation tool

To train our model with the dataset, we used Labkit plugin within the FIJI [18] software to meticulously label all the data (both for training and testing). Labkit provides user-friendly manual and automated image segmentation procedures. Each image in the dataset is accompanied by manually annotated ground truth mask, offering an exact demarcation of the cells we are interested in. Notably, this dataset contains annotations for over 3,000 nuclear boundaries. A cumulative time exceeding 1000 hours has been dedicated to the data annotation step.

2.4 Performance Metrics

We evaluate our method by using two types of metrics, mean Intersection Over Union and F1-score.

2.4.1 Intersection Over Union (IoU)

IOU is object-level metric that uses a measure of area coverage to identify correctly segmented nuclei. This metric measures the overlap between the predicted (estimated) and ground truth regions of objects in an image. We calculated IoU with threshold of 0.5 as the ratio of the intersection of the predicted and ground truth regions to their union. This is often expressed as a value between 0 and 1. Let's consider that we have two boxes (images) A and B where, A belongs to predicted box and B belongs to ground truth box. The calculation involves determining the ratio of intersection area ($A \cap B$) and the union area ($A \cup B$), as illustrated in the **Figure 4**. Hence, the IoU is computed according to **Equation 1** as follows:

$$IoU = \frac{|A \cap B|}{|A \cup B|}$$

Equation 1. Formula for calculating intersection-of-union.

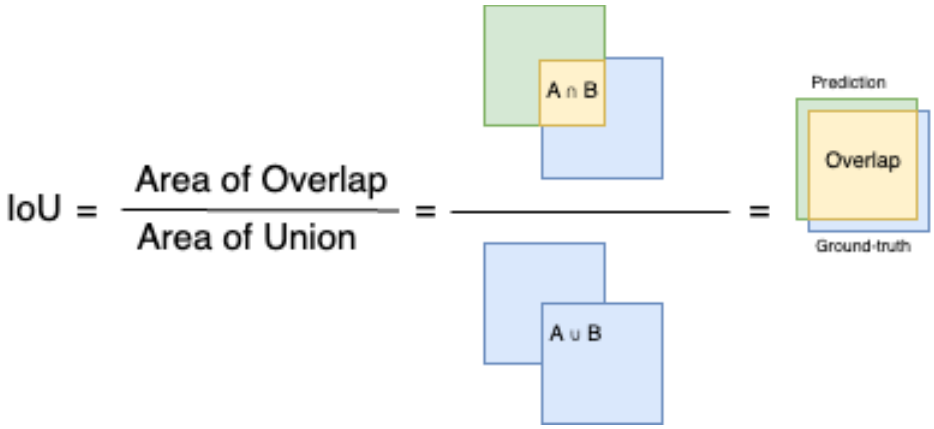


Figure 4. Diagram to depict concept of IOU metric.

2.4.2 F1-Score

F1 Score aims to measure the balance between precision and recall as a pixel-level evaluation metrics. Precision is the ratio of the true positive results to the sum of true positive and false positives. On the other hand, recall is the ratio of true positive results to the sum of true positives and true negatives in the test set. It is computed by the harmonic mean of precision and recall as in **Equation 2**:

$$\text{F1 Score} = \frac{2 \text{ x Precision x Recall}}{\text{Precision} + \text{Recall}}$$

Equation 2. Formula for calculating F1-Score.

Precision and Recall are calculated as, **Equation 3**:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Equation 3. Equation for Precision and Recall where FP, TP and FN denote false positive, true positive and false negative.

2.4.3. Loss Function

The loss function is used by models to learn the trainable parameters, such as weights and biases. Due to the sizable imbalance of overlapping and background regions, we trained the model using combined of dice loss and binary focal loss. The purpose of combining these loss functions is to create a composite loss that takes into account both the similarity between the predicted and ground truth and the difficulty of classifying the pixels. Dice loss uses the dice coefficient to measure the overlapping of the pixels of the predicted labels with the ground truth label. Binary focal loss generalizes binary cross-entropy by introducing a hyperparameter γ (gamma), called the focusing parameter, that allows hard-to-classify examples to be penalized more heavily relative to easy-to-classify examples [3] .The formulas for Dice loss, Focal loss, and Total loss are illustrated in **Equation 4**, **Equation 5**, **Equation 6**.

$$\text{Dice} = \frac{2 * |\,\text{X} \cap \text{Y}\,|}{|\,\text{X}\,| + |\,\text{Y}\,|}$$

Equation 4. Equation for calculating dice coefficient.

$$\text{FL}(p_t) = -\alpha_t(1 - p_t)^{\gamma} \log(p_t).$$

Equation 5. Equation for calculating focal loss.

Total loss = Dice loss + (1 * Binary Focal loss)

Equation 6. Equation for total loss.

Chapter 3

3 Methods

In this section, we describe the methods we used with the architectures of the deep learning models for nuclei segmentation. The comprehensive approach encompasses various stages, each playing a crucial role in achieving accurate and robust segmentation results. The following subsections outline the specific procedures involved in data preparation, model development, and the post-processing methods, including tasks such as separating boundaries of touching nuclei, labeling individual nuclei, and calculating various metrics for analysis.

3.1 Data Preprocessing

As an initial pre-processing procedure, we employed a Python script to extract 2D image slices from the 3D dataset by utilizing a threshold intensity level set above 2000. We chose 2D images to expedite analysis since the features of interest are concentrated in only a few slices. Subsequently, we undertook manual removal of noisy data to ensure accurate data curation. The raw images collected are of various sizes and cannot be fed into our deep learning model directly since our model requires fixed-size input images. To standardize image sizes across our dataset and for the consideration of GPU memory limitation, we resized them to 256 x 256 pixels.

3.2 Data Annotation

The annotation process involves several steps. Initially, we perform manual pixel-based selection of nuclear boundaries, marking them as foreground by highlighting them in red color. We also provide examples of background selection, which are colored in blue. In this way, we label a few

pixels per class (foreground and background) and use a random forest classifier inbuilt in the Labkit to generate the mask. We manually verify the segmented mask comparing with original image. If noise is present, we apply Gaussian blur and/or Laplacian of Gaussian to detect the edges. Once the segmentation is finalized, we rescale the image values to a 0-1 range dimension and change the mask type to an 8-bit image. **Figure 5** offers an overview of the data annotation process employed in each step.

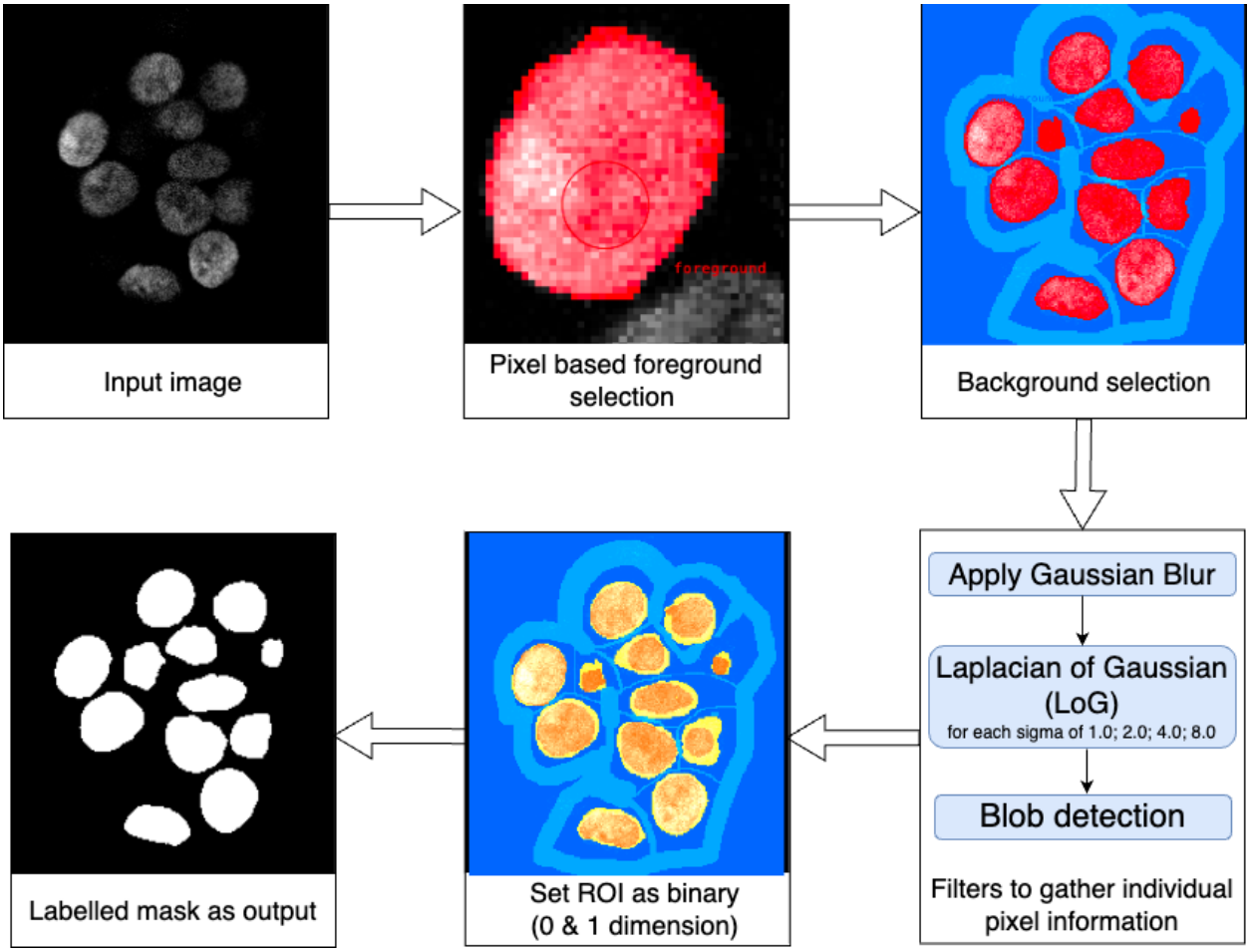


Figure 5. Demonstration of data annotation pipeline

3.3 Data Augmentation

To minimize this issue of overfitting, we applied augmentation techniques to the images. We utilize the albumentation library developed by Buslaev et al. (2020) to achieve our goal [42]. A variety of conventional 2D image processing techniques such as transformation of horizontal flip, random crop, elastic transformation, shift scale rotation and random brightness contrast in different order, are applied in our method, examples in **Figure 6**.

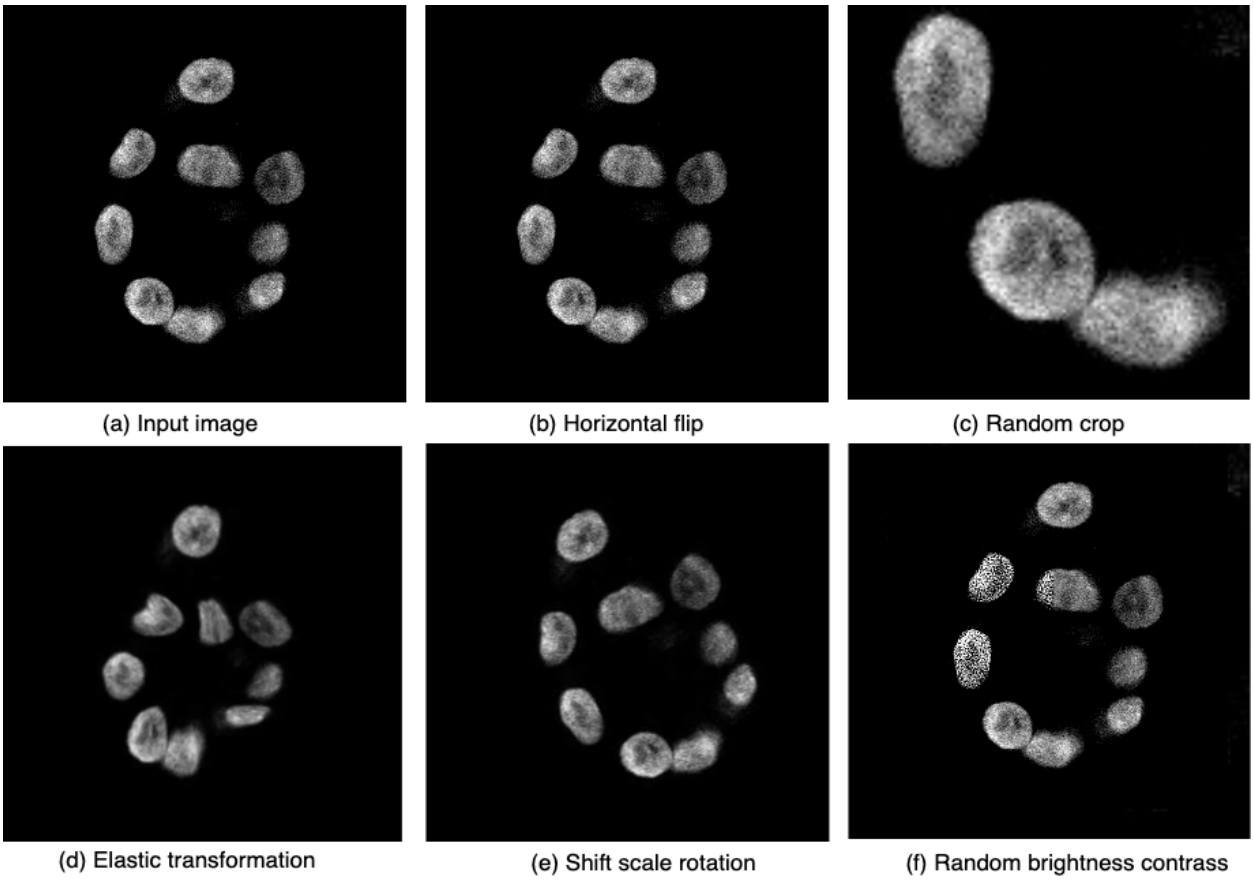


Figure 6. Example of transformations applied for data augmentation.

3.4 Test Bed and Experimental Setup

In this section, we elaborate on the test bed and hyperparameter settings selected for the target task to obtain optimal results. The deep learning model was implemented using Keras with Tensorflow backend and data was processed using Python 3.10 and libraries in Python such as Keras, NumPy, pandas, and scikit-learn. We used Google Colab with the high-performance NVIDIA Tesla T4 GPU with CUDA 9.0. We divided the data using the train_test_split function from the scikit-learn library. In this split, 80% of the dataset is allocated for training the proposed model, while the remaining 20% is reserved for evaluation, as indicated in

Table *3*. The training was performed for 100 epochs, with each model taking approximately 2-4 hours to complete. A batch size of 4 yielded superior results in comparison to both 2 and 8, therefore it was set to 4. We employed Adam as the optimizer with a learning rate (LR) of $10e^{-4}$ for training, aiming to maximize loss reduction [8]. This choice of optimizer, along with the specified LR, was particularly effective in managing sparse gradients associated with the nuclei pixel in the foreground. The entire model was trained on a server with the specifications listed in **Table 2**.

Table 2. Hyperparameter settings for training process.

No	Hyperparameters	Settings
1	Optimizer	Adam
2	Callbacks	ReduceLROnPlateau
3	Learning rate	$10\ e^{-4}$
4	Batch size	4
5	Epochs	100
6	Loss function	Combined Loss

Table 3. Subdivision of dataset for training the model.

Set	Number of Images	Percentage
Train	437	80 %
Validation	55	10 %
Test	55	10 %
Total	547	100 %

Backbone Model

In our study, we conducted several experiments with different configurations by substituting the typical set of convolution layers found in the encoder of U-Net [8] framework with different pretrained models. We leveraged sophisticated models like Resnet50 [43], Inception ResNet V2 [10], Vgg19 [44], Densenet121 [45], MobileNet [46] and EfficientNet B0 – B7 [36] as the encoder part for the U-Net for comparative analysis in this study. In the proposed semantic segmentation task, EfficientNet-B5 used as a backbone provides the best result among all our experiment. As a result, we've decided to elaborate on this architecture.

U-Net with EfficientNet Encoder

We used the state-of-the art EfficientNet-B5 model as encoder, instead of using traditional convolutional layer. The proposed network is constructed using the first four layers in the contracting path, and the last four layers in the expansion path, with the fifth layer bringing it all together by connecting the two. As we progress through the encoder's depth, the image size gradually decreases. Starting with images of size 256 x 256 x 3, we downscale them to 8 x 8 x 256

by effectively capturing the image context. The characteristics of each feature map, including the number of channels, resolution, and bandwidth, are illustrated in **Figure 7**.

To ensure efficient model training, we utilize input images with a resolution of 256 x 256 x 3. In the decoder section, these images are upscaled to a size of 256 x 256 x 1. Skip connection concatenate the encoder section of network with the upscaling layer to capture spatial features of image. The output passes through a final convolutional layer featuring a 1×1 convolution kernel and a sigmoid activation function. The incorporation of a powerful convolutional neural network like EfficientNet in the U-Net significantly enhances the overall algorithm performance.

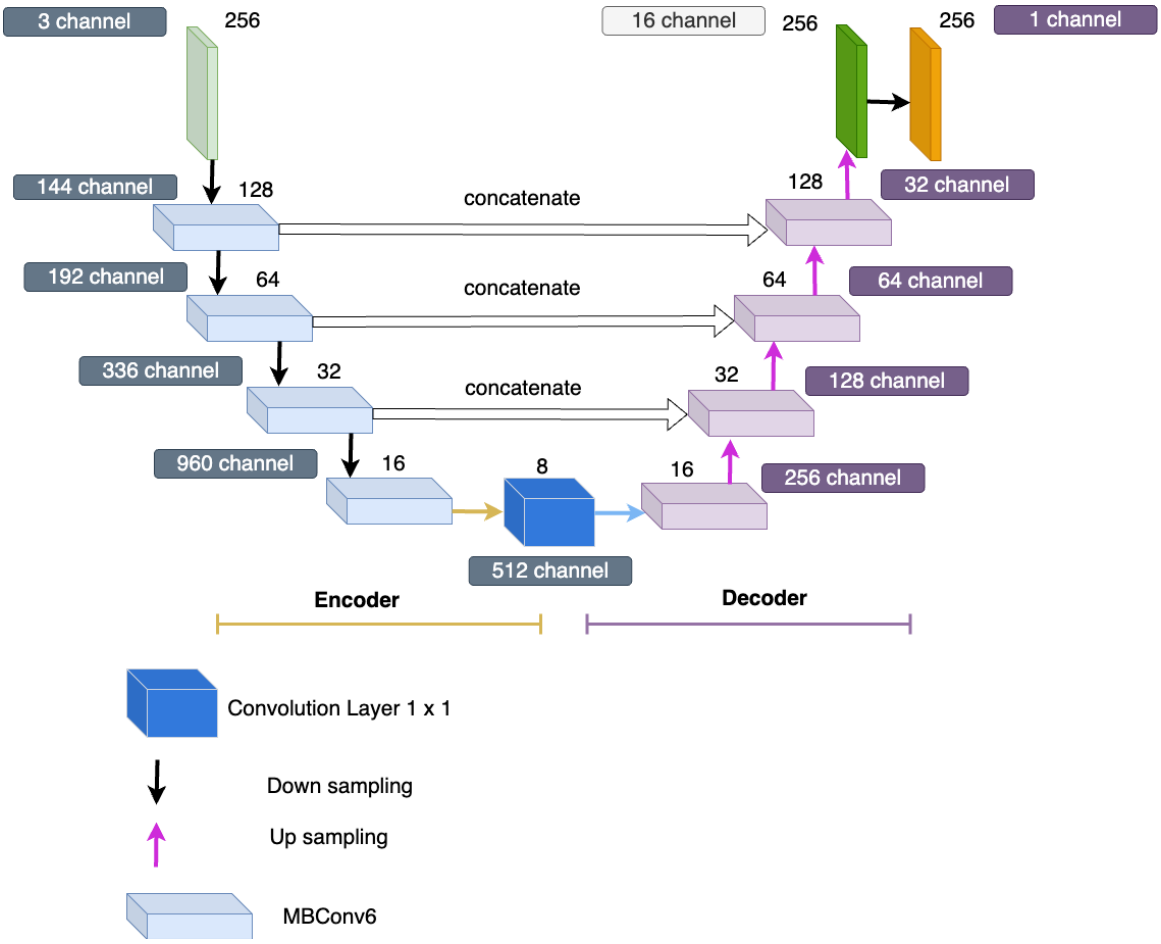


Figure 7. EfficientNet as encoder in U-Net architecture

3.5 Post processing

As a post processing step, we implemented histogram-based segmentation including marker-based watersheds to separate the nuclei that are close to each other. **Figure 8** shows the workflow carried out in this step. For the quantification of nuclei, we labeled each cell with numbering and calculated various metrics by labeling each nucleus, such as area, equivalent diameter, mean intensity, perimeter, small diameter, and large diameter. All measurements are provided in millimeters, and the area has also been expressed in square micrometers in **Figure 11**.

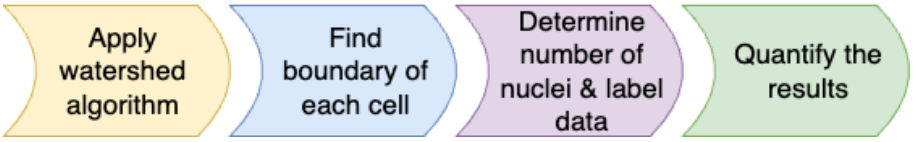


Figure 8. Pipeline to demonstrate the steps in post-processing method.

Chapter 4

4 Results

4.1 Evaluating the performance of Encoder-Decoder Architectures

Our investigation begins with comparison of state-of-art encoder-decoder based model for nuclei segmentation. According to result obtained in **Figure 9**, U-Net demonstrates superior performance compared to FPN, LinkNet, and PSPNet in terms of F1-score. Therefore, we selected the U-Net framework to carry out further experiments.

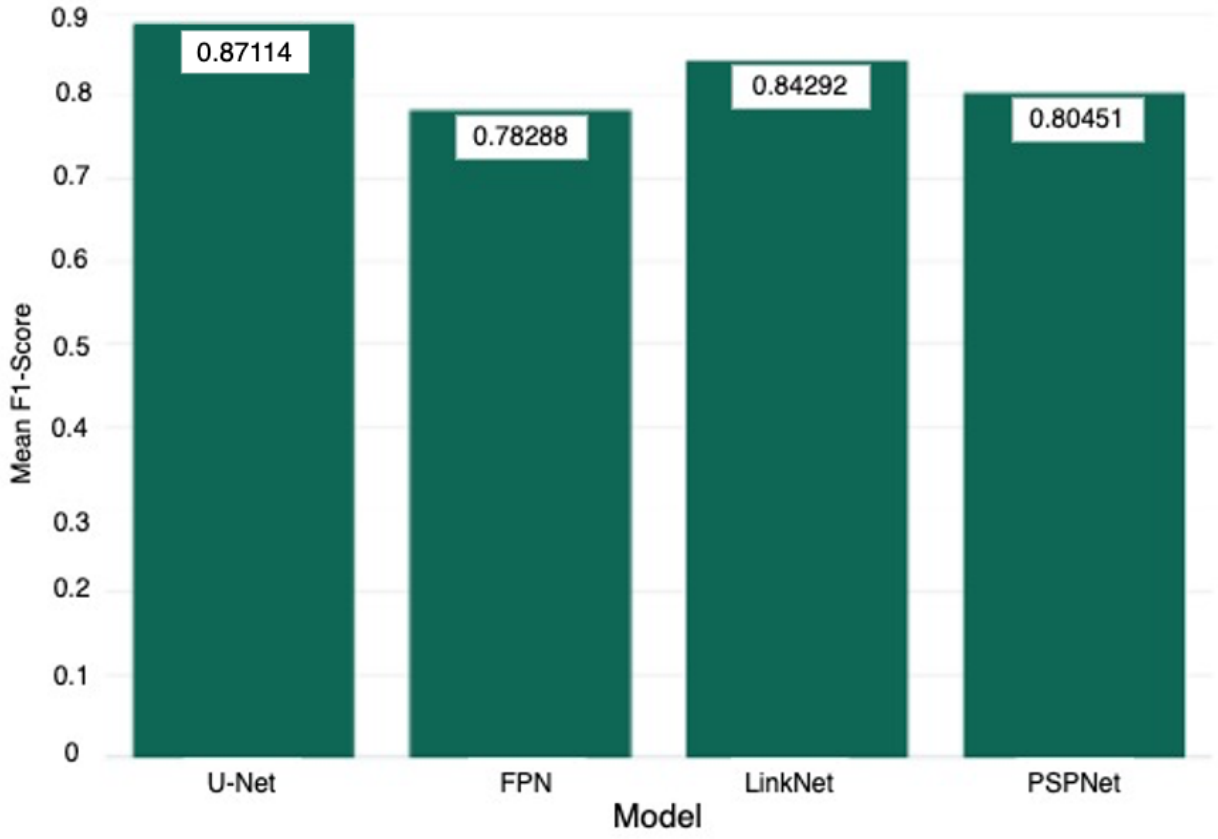


Figure 9. Comparison of F1-Score among different encoder and decoder-based architecture.

4.2 Evaluating the performance of various Encoders in the U-Net

Table 4 represents the quantitative comparison based on performance of different methods on BMGD dataset. **Figure 10** illustrates the loss per epoch graph for purposed model (Efficient-B5 with U-Net model). **Table 5**, on the other hand, presents a comparison of results for the same predicted image across various models.

Table 4. Quantitative comparison of different methods on BMGD dataset.

Model	Encoder	F1-Score	IoU Score	Validation Loss
U-Net	**EfficientNetB5**	**0.87114**	**0.80898**	**0.02456**
U-Net	Resnet50	0.77501	0.65102	0.16675
U-Net	Inception ResNet V2	0.84259	0.74521	0.13852
U-Net	Vgg19	0.45499	0.37303	1.09673
U-Net	Densenet121	0.82262	0.71279	0.14649
U-Net	MobileNet	0.77661	0.64324	0.15554

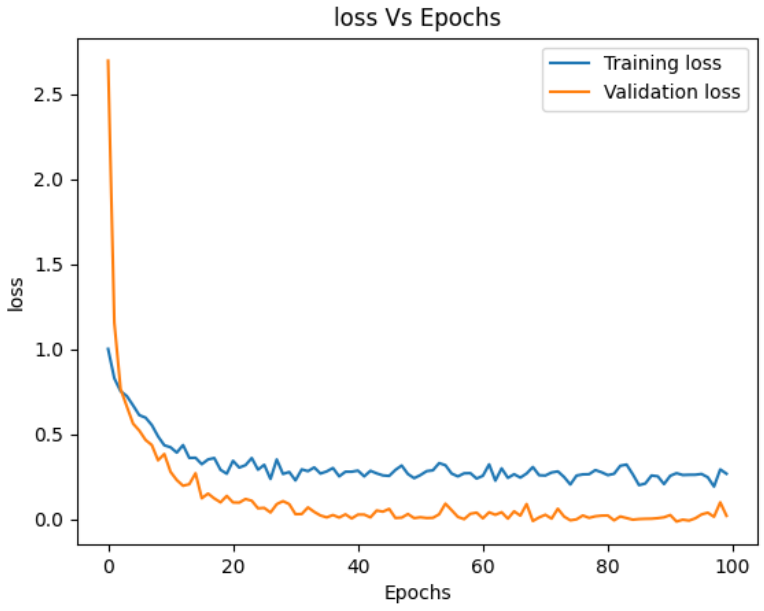
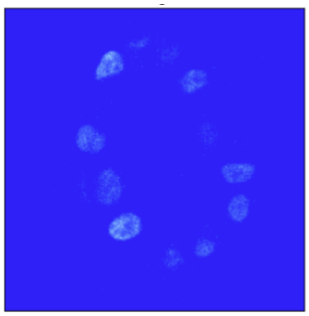
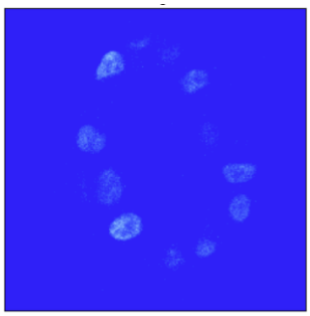
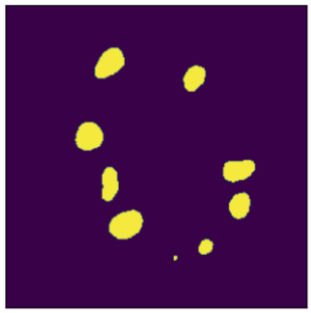
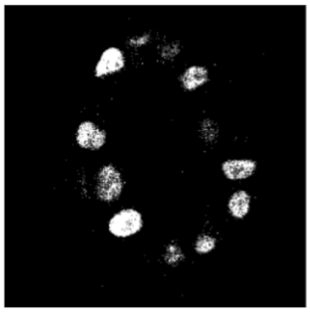
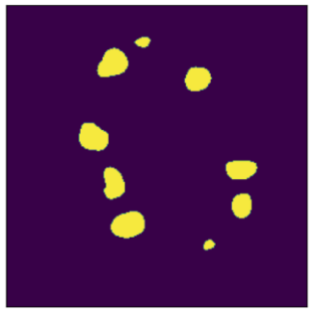
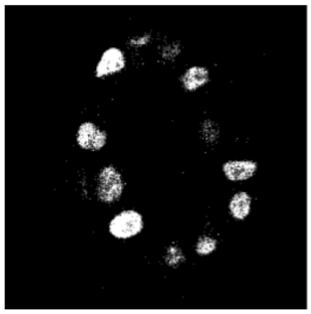


Figure 10. Illustration of loss per epoch graph for EfficientNet-B5

Table 5. Illustration of predicted image in different model.

Model	Original Image	Ground Truth	Predicted Image
EfficientNet-B5 (best)			
ResNet-50			
Inception-ResNet-v2			
Vgg19			

4.3 Evaluating the performance of EfficientNet Family (B0-B7)

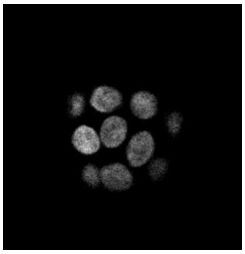
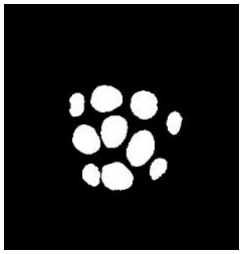
In the context of our study, we conducted a comprehensive analysis of the EfficientNet framework family. Our findings indicate that, when applied to our dataset, Efficient-B5 demonstrated superior generalization performance, presented in **Table 6**.

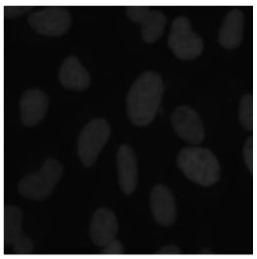
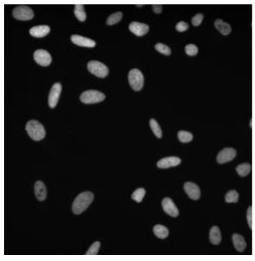
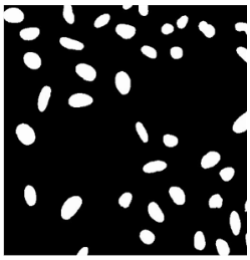
Table 6. Quantitative comparison of EfficientNet Family based on evaluation metrics.

Encoder	F1-Score	IoU Score	Total trainable params
EfficientNet B0	0.84532	0.79948	10,071,501
EfficientNet B1	0.8499	0.7896	14,225,747
EfficientNet B2	0.85753	0.79099	12,577,137
EfficientNet B3	0.86522	0.81435	17,867,833
EfficientNet B4	0.86953	0.81842	25,607,833
EfficientNet B5	**0.87114**	**0.80898**	37,293,953
EfficientNet B6	0.85056	0.7598	50,679,593
EfficientNet B7	0.85334	0.81177	74,735,393

Furthermore, **Table 7** provides a visual presents the predicted outcomes of the purposed model (Efficient-B5 with U-Net) with KDSB18 and BBBC022 dataset, showcasing the effectiveness of the proposed model. These datasets consist of previously unseen test images for our model.

Table 7. Visual comparison of predicted result in different dataset based on proposed model.

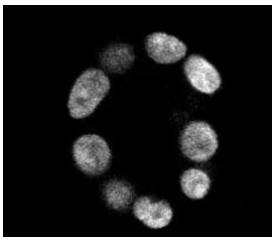
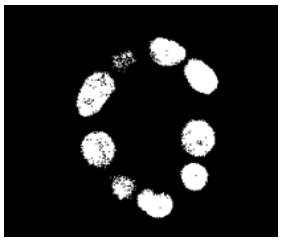
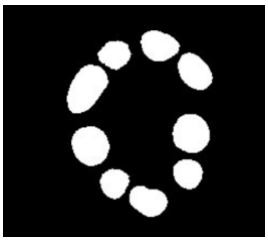
Dataset	# of Images	Stain	Image type	Size	Original Image	Predicted result
BMGD	500	DAPI	.tif	Varies		

KDSB18	546	DAPI	.tif	512 × 512		
BBBC022	50	DAPI	.png	690 x 512		

4.4 Semantic Segmentation Comparison: Classical approach vs. Proposed method

The findings showcased in **Error! Reference source not found.** of this section represents semantic segmentation result comparison of classical computer vision approach using Otsu thresholding segmentation method and the method proposed in this study.

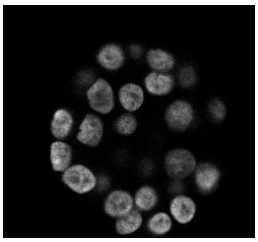
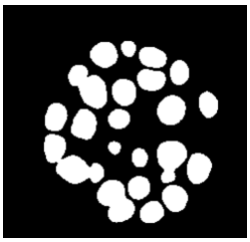
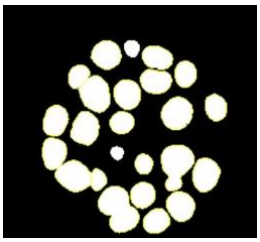
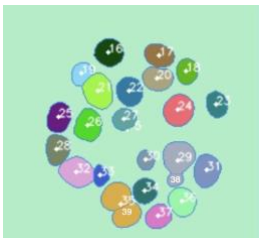
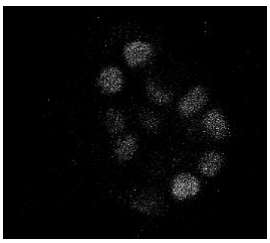
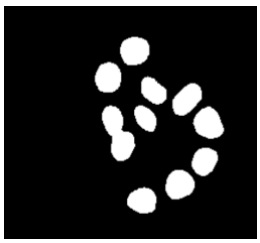
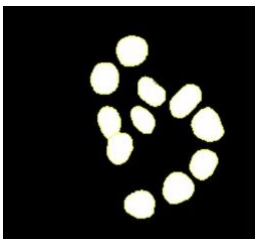
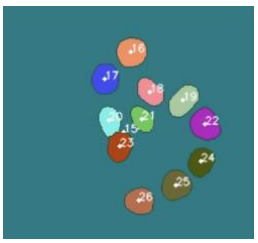
Table 8. Nuclei segmentation using classical approach vs proposed method.

Dataset	Original image	Classical approach	Predicted result
Result			

4.5 Post-Processing and Quantification Metrics for Nuclei Analysis

In **Table 9** and **Figure 11** present the outcomes of the quantified result from the post-processing step, along with metrics for the analysis of nuclei segmentation.

Table 9. Illustration of the original image, the predicted outcome, and the labeled nuclei.

Dataset	Original image	Predicted image	Separated borders	**Result** Labelled nuclei
Image 1				
Image 2				

		label	area	equivalent_diameter	mean_intensity	perimeter	axis_major_length	axis_minor_length	area_sq_microns
2	1	16	548	26.41467907	255	85.84062043	27.06183191	25.82253665	112.951568
3	2	17	476	24.61832698	250.1785714	80.76955262	28.80730527	21.10270548	98.111216
4	3	18	372	21.76338923	255	69.69848481	24.28611484	19.50537881	76.675152

Figure 11. Quantified result of each nucleus is presented in the Excel file.

4.6 Discussion

For extraction of features, we experimented with several encoders such as EfficientNet, ResNet-50, InceptionResNetV2, Vgg19, Densenet121 and MobileNet with U-Net decoder presented in **Table 4**. The result concludes that EfficientNet-B5 as the backbone of the nuclei segmentation network can extract texture of nuclei with the best results of F1 score of 87.1% with a mean IOU score of 80%. We observed a noticeable enhancement in the F1-score, which increased from 0.85485 to 0.871 as a result of changing the data split from 70-30% to 80-20%. Furthermore, the

proposed network demonstrated effective generalization, as evidenced by a validation loss of 0.02456. In **Figure 10**, we presented loss per epoch graph. As we have not applied augmentation in the validation dataset, therefore it results in lower validation loss. This gives the impression of validation being less challenging compared to the training dataset. Low validation loss indicates that the model's predictions on unseen data closely aligned with the actual target values, reflecting its robust performance in capturing underlying patterns in the dataset. **Table 5** displays a comparison of results for the same predicted image across different models, indicating that the combination of Efficient-B5 with U-Net yields the most favorable outcome.

The experimental findings presented in **Table 6** demonstrate that EfficientNet-B5, despite having a lower parameter count compared to EfficientNet-B6 and EfficientNet-B7, outperforms them. Based on this experimental finding, we can state that the use of multiple blocks with different receptive field sizes in EfficientNet-B5 for multi-scale feature extraction is particularly advantageous in segmentation tasks where objects of different sizes need to be accurately delineated. Since our dataset is small, EfficientNet-B7's greater capacity might not be fully exploited, which would result in declining results.

Furthermore, **Table 8** demonstrates that our deep learning-based model outperforms a classical computer vision-based segmentation approach. This superiority is particularly notable because in the result, the Otsu algorithm struggles to identify an optimal threshold in cases where there is no clear bimodal distribution in the intensity histogram. This illustrates that the shortcomings of traditional computer vision methods become apparent when confronted with images that lack clear distinctions between background and foreground values. In such situations, the proposed model excels, delivering more precise and nuanced segmentation results.

Chapter 5

5 Conclusion

5.1 Summary

In this study, we establish a pipeline for the automatic segmentation of nuclei in DAPI-stained fluorescent images. The aim is to assist medical experts in their diagnosis by automating the entire process. Even experts can occasionally make mistakes, hence the need for full pipeline automation. In conclusion, this study has successfully addressed the challenges associated with semantic segmentation of cell nuclei, including unconventional morphology, noise, and overlapping instances. We implement a U-Net encoder-decoder based approach utilizing a pre-trained EfficientNet-B5 as the network backbone for pixel-level nuclei segmentation. The experimental results demonstrated the effectiveness of our proposed method for extracting nuclei cells, with F1-score of 87% and an IoU of 80%. The model demonstrates highly accurate segmentation compared to state-of-the-art networks like ResNet-50, InceptionResNetV2, Vgg19, Densenet121, and MobileNet. Moreover, this investigation provides quantified outcomes for the analysis of individual cell nuclei. The manual segmentation conducted on this dataset proves as a valuable asset for training segmentation algorithms, alleviating the labor-intensive manual tracing required by other researchers.

5.2 Limitations and Future work

In addition to the successful implementation and promising outcomes of our automated nuclei segmentation pipeline, there are avenues for future. Firstly, the dataset meticulously curated for this research holds significant potential for further investigations using alternative models and

methodologies. Recognizing the value of sharing this resource with the scientific community, we intend to publish a data paper, thereby facilitating its utilization in diverse research endeavors. Moreover, it's imperative to acknowledge that our current research framework operates within the realm of supervised learning. As part of our future endeavors, we envision exploring the domain of self-supervised learning. Embracing cutting-edge algorithms like Barlow's Twin, we aim to enhance the robustness of our model by delving into the realm of unsupervised or self-supervised learning approaches. These endeavors underscore our commitment to continual improvement and innovation in the field of automated nuclei segmentation for medical diagnosis applications.

References

1. Caicedo, J.C., Goodman, A., Karhohs, K.W., Cimini, B.A., Ackerman, J., Haghighi, M., Heng, C.K., Becker, T., Doan, M., McQuin, C., Rohban, M., Singh, S., Carpenter, A.E.: Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. Nat Methods. 16, (2019). https://doi.org/10.1038/s41592-019-0612-7
2. Basu, A., Senapati, P., Deb, M., Rai, R., Dhal, K.G.: A survey on recent trends in deep learning for nucleus segmentation from histopathology images, (2023)
3. Shen, D., Wu, G., Suk, H. Il: Deep Learning in Medical Image Analysis. Annu Rev Biomed Eng. 19, (2017). https://doi.org/10.1146/annurev-bioeng-071516-044442
4. Chen, G., Huang, S., Cao, L., Chen, H., Wang, X., Lu, Y.: Application of Plant Phenotype Extraction Using Virtual Data with Deep Learning. In: Journal of Physics: Conference Series (2022)
5. Cao, Y., Liu, Z., Zhang, P., Zheng, Y., Song, Y., Cui, L.: Deep Learning Methods for Cardiovascular Image. Journal of Artificial Intelligence and Systems. 1, (2019). https://doi.org/10.33969/ais.2019.11006
6. Boutros, M., Heigwer, F., Laufer, C.: Microscopy-Based High-Content Screening, (2015)
7. Adeyinka, A.A., Adebiyi, M.O., Akande, N.O., Ogundokun, R.O., Kayode, A.A., Oladele, T.O.: A Deep Convolutional Encoder-Decoder Architecture for Retinal Blood Vessels Segmentation. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2019)
8. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W.M., and Frangi, A.F. (eds.)

Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. pp. 234–241. Springer International Publishing, Cham (2015)

9. Tomczak, K., Czerwińska, P., Wiznerowicz, M.: The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge, (2015)
10. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the Inception Architecture for Computer Vision. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2016)
11. Ng, H.P., Ong, S.H., Foong, K.W.C., Goh, P.S., Nowinski, W.L.: Medical image segmentation using k-means clustering and improved watershed algorithm. In: Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation (2006)
12. Alahmari, S.S., Goldgof, D., Hall, L.O., Mouton, P.R.: A Review of Nuclei Detection and Segmentation on Microscopy Images Using Deep Learning With Applications to Unbiased Stereology Counting. IEEE Trans Neural Netw Learn Syst. (2022). https://doi.org/10.1109/TNNLS.2022.3213407
13. Bougen-Zhukov, N., Loh, S.Y., Lee, H.K., Loo, L.H.: Large-scale image-based screening and profiling of cellular phenotypes, (2017)
14. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern. 9, (1996)
15. Zhang, C., Xiao, X., Li, X., Chen, Y.J., Zhen, W., Chang, J., Zheng, C., Liu, Z.: White blood cell segmentation by color-space-based k-means clustering. Sensors (Switzerland). 14, (2014). https://doi.org/10.3390/s140916128

16. Irshad, H., Veillard, A., Roux, L., Racoceanu, D.: Methods for nuclei detection, segmentation, and classification in digital histopathology: A review-current status and future potential. IEEE Rev Biomed Eng. 7, (2014). https://doi.org/10.1109/RBME.2013.2295804
17. Xing, F., Yang, L.: Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: A comprehensive review, (2016)
18. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A.: Fiji: An open-source platform for biological-image analysis, (2012)
19. McQuin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, B.A., Karhohs, K.W., Doan, M., Ding, L., Rafelski, S.M., Thirstrup, D., Wiegraebe, W., Singh, S., Becker, T., Caicedo, J.C., Carpenter, A.E.: CellProfiler 3.0: Next-generation image processing for biology. PLoS Biol. 16, (2018). https://doi.org/10.1371/journal.pbio.2005970
20. De Chaumont, F., Dallongeville, S., Olivo-Marin, J.C.: ICY: A new open-source community image processing software. In: Proceedings - International Symposium on Biomedical Imaging (2011)
21. Narotamo, H., Sanches, J.M., Silveira, M.: Segmentation of Cell Nuclei in Fluorescence Microscopy Images Using Deep Learning. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2019)
22. International Symposium on Biomedical Imaging (ISBI) Cell Tracking Challenge. 2019, https://biomedicalimaging.org/2019/challenges/

23. Simon Graham, Mostafa Jahanifar, Quoc Dang Vu, Giorgos Hadjigeorghiou: Conic: Colon nuclei identification and counting challenge 2022. Presented at the

24. Verma, R., Kumar, N., Patil, A., Kurian, N.C., Rane, S., Graham, S., Vu, Q.D., Zwager, M., Raza, S.E.A., Rajpoot, N., Wu, X., Chen, H., Huang, Y., Wang, L., Jung, H., Thomas Brown, G., Liu, Y., Liu, S., Jahromi, S.A.F., Khani, A.A., Montahaei, E., Baghshah, M.S., Behroozi, H., Semkin, P., Rassadin, A., Dutande, P., Lodaya, R., Baid, U., Baheti, B., Talbar, S., Mahbod, A., Ecker, R., Ellinger, I., Luo, Z., Dong, B., Xu, Z., Yao, Y., Lv, S., Feng, M., Xu, K., Zunair, H., Hamza, A. Ben, Smiley, S., Yin, T.K., Fang, Q.R., Srivastava, S., Mahapatra, D., Trnavska, L., Zhang, H., Narayanan, P.L., Law, J., Yuan, Y., Tejomay, A., Mitkari, A., Koka, D., Ramachandra, V., Kini, L., Sethi, A.: MoNuSAC2020: A Multi-Organ Nuclei Segmentation and Classification Challenge. IEEE Trans Med Imaging. 40, (2021). https://doi.org/10.1109/TMI.2021.3085712

25. Zeng, Z., Xie, W., Zhang, Y., Lu, Y.: RIC-Unet: An Improved Neural Network Based on Unet for Nuclei Segmentation in Histology Images. IEEE Access. 7, (2019). https://doi.org/10.1109/ACCESS.2019.2896920

26. Raju, R., Paul, M.M., Asokachandran, V., George, B., Radhamony, L., Vinaykumar, M., Girijadevi, R., Pillai, R.R.: The Triple-Negative Breast Cancer Database: An omics platform for reference, integration and analysis of triple-negative breast cancer data, (2014)

27. Hassan, L., Saleh, A., Abdel-Nasser, M., Omer, O.A., Puig, D.: Efficient Multi-Organ Multi-Center Cell Nuclei Segmentation Method Based on Deep Learnable Aggregation Network. Traitement du Signal. 38, (2021). https://doi.org/10.18280/ts.380312

28. Jha, D., Riegler, M.A., Johansen, D., Halvorsen, P., Johansen, H.D.: DoubleU-Net: A deep convolutional neural network for medical image segmentation. In: Proceedings - IEEE Symposium on Computer-Based Medical Systems (2020)
29. Cheng, Q., Parvin, B.: Organoid model of mammographic density displays a higher frequency of aberrant colony formations with radiation exposure. Bioinformatics. 36, 1989–1993 (2019). https://doi.org/10.1093/bioinformatics/btz888
30. Carpenter et al.: Broad Bioimage Benchmark Collection, https://bbbc.broadinstitute.org/BBBC002
31. Meijering, E.: Cell segmentation: 50 Years down the road [life Sciences]. IEEE Signal Process Mag. 29, (2012). https://doi.org/10.1109/MSP.2012.2204190
32. Song, Y., Cai, W., Huang, H., Wang, Y., Feng, D.D., Chen, M.: Region-based progressive localization of cell nuclei in microscopic images with data adaptive modeling. BMC Bioinformatics. 14, (2013). https://doi.org/10.1186/1471-2105-14-173
33. Bengio, Y.: Deep Learning of Representations for Unsupervised and Transfer Learning. JMLR Workshop Conf Proc. 7, (2011)
34. Yousef, R., Gupta, G., Yousef, N., Khari, M.: A holistic overview of deep learning approach in medical imaging. Multimed Syst. 28, (2022). https://doi.org/10.1007/s00530-021-00884-5
35. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: MobileNetV2: Inverted Residuals and Linear Bottlenecks. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2018)
36. Tan, M., Le, Q. V.: EfficientNet: Rethinking model scaling for convolutional neural networks. In: 36th International Conference on Machine Learning, ICML 2019 (2019)

37. Xing, Y., Zhong, L., Zhong, X.: An Encoder-Decoder Network Based FCN Architecture for Semantic Segmentation. Wirel Commun Mob Comput. 2020, (2020). https://doi.org/10.1155/2020/8861886

38. Ji, J., Lu, X., Luo, M., Yin, M., Miao, Q., Liu, X.: Parallel Fully Convolutional Network for Semantic Segmentation. IEEE Access. 9, (2021). https://doi.org/10.1109/ACCESS.2020.3042254

39. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: Feature pyramid networks for object detection. In: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 (2017)

40. Chaurasia, A., Culurciello, E.: LinkNet: Exploiting encoder representations for efficient semantic segmentation. In: 2017 IEEE Visual Communications and Image Processing, VCIP 2017 (2017)

41. Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J.: Pyramid scene parsing network. In: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 (2017)

42. Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A.A.: Albumentations: fast and flexible image augmentations. Information. 11, 125 (2020)

43. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2016)

44. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings (2015)

45. Singh, D., Kumar, V., Kaur, M.: Densely connected convolutional networks-based COVID-19 screening model. Applied Intelligence. 51, (2021). https://doi.org/10.1007/s10489-020-02149-6

46. A. Howard, M.Z.B.C.D.K.W.W.T.W.M.A. and H.Adam.: MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. BibSonomy. (2017)