Accurate nuclei segmentation in breast cancer tumour biopsies

SHENFIELD, Alex <http://orcid.org/0000-0002-2931-8077>, KASTURI, Surya and TRAN, William

Accurate Nuclei Segmentation in Breast Cancer Tumour Biopsies

Surya Kasturi
Dept. of Engineering and Mathematics
Sheffield Hallam University
Sheffield, United Kingdom
s.kasturi@shu.ac.uk

William T. Tran
Dept. of Radiation Oncology
Sunnybrook Health Sciences Center
Toronto, Canada
william.tran@sunnybrook.ca

Alex Shenfield*
Dept. of Engineering and Mathematics
Sheffield Hallam University
Sheffield, United Kingdom
a.shenfield@shu.ac.uk

***Abstract*—Breast cancer identification is a arduous process and diagnosing it using Haematoxylin and Eosin (H&E) stained pathology images is a significant challenge, with pathologists struggling to segment cancer nuclei accurately. This study will evaluate the efficacy of different methods utilising deep learning techniques for breast cancer nuclei segmentation, with a particular emphasis on U-Net architectures. The proposed methodology is divided into four stages: image enhancement, individual nuclei segmentation, feature extraction, and whole image binary segmentation. This work then conducts a rigorous comparison of different segmentation techniques.**

***Index Terms*—U-Net, Histopathology Images, Breast Cancer, ResNet, ResNext, DenseNet**

I. Introduction

Breast cancer is one of the leading causes of death among women. Globally in 2015, around 2.4 million women had breast cancer treatment, and 0.5 million died due to the disease [7]. Breast cancer screening comprises one or two radiologists examining scans for pre-symptomatic cancer symptoms with the goal of minimising sickness and death. Additionally, such screening has negative ramifications including over-diagnosis and over-treatment of tumours that would not have manifested throughout the woman's lifespan if she had not been properly examined. There is a dispute on the extent of over-diagnosis, ranging from 1% to 54% of cancers found with screening and the balance of benefits and risks associated with screening [36]. While detecting low grade ductal invasive cancer raises the risk of over-diagnosis [32], detecting grade 3 cancer results in fewer deaths [30]. Between 0.6% and 0.8% of women are diagnosed with cancer as a result of routine screening [3]. Additionally, breast screening programmes miss between 15% and 35% of cancers due to human error or because the tumour is not visible or detectable to the screening radiologist [13]. Studies suggest that the screening of mammography using AI has improved over the years, and AI systems have been shown to exhibit better performance while also reducing the effort of the second reader by 88% [24]. This comprehensive examination of the artificial intelligence system sets the path for clinical studies to improve the performance and reliability of cancer prediction.

* Corresponding author.

II. Literature Survey

In recent years, researchers have developed a number of useful and efficient deep learning methods. The majority of deep learning approaches are capable of automatically extracting features and optimising models to ensure robustness and high performance. Different deep learning models have been adopted to do tasks like segmentation, classification and cluster analysis. One such example for image analysis is Convolution Neural Networks (CNNs), which were originally proposed in 1980 [23], but have undergone significant improvements in recent years [12], [20], [27]. However, much of the focus of research in this area has been on image classification and it is unclear how much of this crosses over to the problem of biomedical image segmentation. A further drawback is the large number of images typically required to train a CNN effectively [10]. To deal with these challenges, in 2015, a team from the University of Freiburg in Germany, proposed a new CNN and encoder-decoder based architecture named U-Net, which plays a significant role in segmentation for biomedical images [26]. Due to the simplicity of the U-Net construction, it can easily be trained with a small number of images, allowing researchers to embed multiple encoders for improved results. For instance, a recent study showed that using an improved version of MobileNet as a backbone encoder for U-Net can provide good results to segment nuclei in the Triple-Negative Breast Cancer (TNBC) histopathology image dataset. This approach achieved 97% accuracy, a 98% Area Under the Curve (AUC), and a 59% Jaccard Similarity Coefficient [18].

Since the inception of U-Net there have been numerous adaptations and alterations to the architecture, with many different encoders used in published research papers; however, the backbone encoders for U-Net that stand out the most are VGG [27], ResNet [12], ResNext [33], DenseNet [15], MobileNet [14], and EfficientNet [31]. In particular, the ResNet architecture [12] has shown great promise for various image segmentation tasks in bio-medical applications and has been widely used [6], [37]. Several studies show that results can be improved both with a pre-trained encoder as a backbone architecture for nuclei segmentation and by using a multi-organ transfer learning approach for segmentation to make the model robust [5], [17]. In particular, using a pre-

978-1-6654-8462-6/22/$31.00

trained backbone based on the EfficientNet architecture with a multi-organ transfer learning approach achieved a recall score of 0.9272 and F1 score of 0.8008, outperforming many other techniques [5]. Simply combining pre-processing of the dataset with post-processing of results from a model can also significantly improve accuracy even with basic a U-Net, allowing it to outperform some of the advanced models with pre-trained encoders. For example, using stain normalisation on whole-slide images in histopathology helps the researchers to build a better model [2]; whilst a 2020 study showed that a basic U-Net trained on a H&E dataset utilizing pre-processing techniques (such as data augmentation) and further post-processing of the results with the watershed algorithm helps to achieve a Dice score of 0.84 [4]. Although the majority of prior research demonstrated good accuracy, there is still an opportunity for performance enhancement, and the majority of these techniques use post-processing techniques to improve results. Additionally, integrating different encoders with the basic U-Net architecture can provide better results in nuclei image segmentation tasks [22].

This research will design a deep learning model for nuclei segmentation by employing a variety of methodologies from the literature including:

- applying pre-processing techniques on openly available datasets
- using a multi-organ transfer learning method for segmentation
- modifying the basic U-Net architecture to employ Batch-Norm [16]
- comparing the performance of different encoders with and without pre-training

All the models will be trained using a combination of the TNBC and MoNuSeg datasets but will be evaluated using a held-out subset of the TNBC dataset for testing.

III. Methodology

Recent developments in convolutional neural networks, particularly with respect to their optimisation, have elevated them to the state-of-the-art in object recognition and classification problems. The research discussed in this manuscript will evaluate a range of sophisticated state-of-the-art encoder-decoder based architectural U-Net models to segment nuclei in breast cancer H&E images. There are four essential steps to apply deep learning to breast cancer nuclei segmentation:

1) Data preparation - including resizing images from the dataset and augmenting existing datasets with new data.
2) Training the candidate models - in this study we used a mixture of the MoNuSeg dataset and a subset of the TNBC dataset.
3) Using the trained models to predict nuclei segmentation masks on the test set - in this case we used a held-out subset of the TNBC dataset.
4) Evaluating the results using a range of performance metrics to determine the best performing model architecture.

TABLE I: Candidate model architectures

Model	Pre-trained
Modified U-Net	ImageNet
ResNet50 Encoder	ImageNet
ResNet101 Encoder	ImageNet
ResNext50-32 4d Encoder	ImageNet
ResNext101-32 8d Encoder	ImageNet
Dense-Net 121 Encoder	ImageNet
Dense-Net 201 Encoder	ImageNet
Basic U-Net	No
ResNet50 Encoder	No
ResNet101 Encoder	No
ResNext50-32 4d Encoder	No
ResNext101-32 8d Encoder	No
Dense-Net 121 Encoder	No
Dense-Net 201 Encoder	No

This study focuses on distinguishing tumour nuclei from normal tissue in Whole Slide Images (WSIs) by using U-Net-based segmentation techniques with a variety of encoder models and different pre-training strategies. Table I provides a full list of the candidate model architectures.

A. Dataset and Pre-Processing

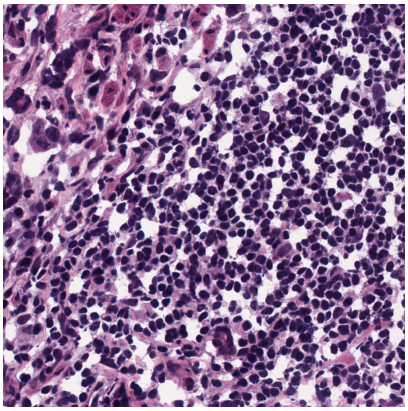
The current study uses images from the Triple-Negative Breast Cancer dataset (TNBC) made available by the International Cancer Genome Consortium (ICGC) [25] and the Multi-Organ Cancer dataset (known as MoNuSeg) published by The Cancer Genome Atlas (TCGA) [21].

The TNBC dataset consists of 50 H&E stained breast pathology images at 40x magnification and their corresponding masks (i.e. the annotated cell nuclei) with a resolution of 512 x 512 pixels each. In this dataset there are a total of 4022 annotated cell nuclei (with an average of 80 cell nuclei per sample). The MoNuSeg dataset consists of 30 images with corresponding masks. Unlike the TNBC dataset, MoNuSeg images and masks are at a higher resolution of 1024 x 1024 pixels. Again, images are captured with 40X magnification. The MoNuSeg collection contains 22,000 nuclear border annotated images; however, these cover a range of different organ types, unlike the TNBC dataset which is solely focused on breast tissue. An example H&E stained digital pathology image and corresponding mask (i.e. cell nuclei annotations) is shown in Fig.1.

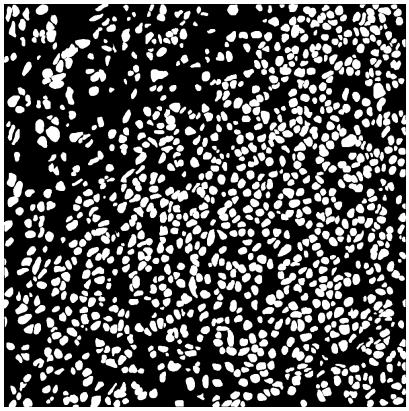
In this study, the training set comprises of all 30 images from the MoNuSeg training dataset and 50% of the TNBC images (i.e. 25 images). In the data pre-processing step all the images were converted using down-sampling into 512 x 512 image resolutions, and then a data augmentation process (changing image orientation) was applied to increase the size of the dataset. Algorithm 1 represents the data augmentation process for both images and masks in the training dataset only. After data augmentation and pre-processing, the training set contains 165 images and corresponding masks with a total of 26,022 annotated nuclei.

B. Model Architecture

The research described in this paper employs U-Net model architecture with different backbone encoders to improve the



(a) Digitized Whole Slide Image



(b) Nuclei Mask Annotation

Fig. 1: A sample H&E stained image and corresponding nuclei mask

Algorithm 1 Data Augmentation algorithm

Input: Dataset, D, consisting of N image and mask pairs (I,M) with a mix of 512 x 512 and 1024 x 1024 resolutions.
Output: Augmented Dataset, D_aug, containing 3 x N image and mask pairs at 512 x 512 pixel resolution.
Algorithm:
for each image and mask pair, (I,M) = 1 to N **do**

1) $(I, M)_1$ = $Flip_H(I, M)$
2) $(I, M)_2$ = $Flip_V(I, M)$
3) $(I, M)_3$ = $Rotate_{45}(I, M)$
4) $(I, M)_{aug}$ = $(I, M)_1$, $(I, M)_2$, $(I, M)_3$

for each image and mask pair, $(I, M)\_aug$ **do**

a) Down-sample image and mask to 512 x 512 resolution
b) Write images into the new path D_aug

accuracy of the segmentation of nuclei in breast histopathology WSIs. Table I provides an overview of the different encoders used in this study.

U-Net: The U-Net model [26] is a kind of autoencoder that is exceptionally good at semantic segmentation. It was specifically designed for biomedical image segmentation, though has subsequently been used for a range of semantic segmentation tasks. As can be seen from Fig.2, it is a U-shaped model with two paths comprising a contracting or downsampling path (on the left side) and an expansive or upsampling path (on the right side). The contracting path is really a simple convolutional network consisting of blocks of two 3x3 unpadded convolutions that are repeatedly applied, each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for downsampling. This architecture has double the amount of feature channels at each downsampling step. After each stage of the downsampling process, context information about the image is gathered and transmitted toward the path of the upsampling process. Every step in the expansive path begins with an upsampling of the feature map, followed by a 2x2 convolution ("up-convolution") that halves the number of feature channels. Cropping between the downsampling and upsampling stages is required because of the loss of boundary pixels in each convolution. A concatenation will be applied with the corresponding cropped feature map from the contracting path and two 3x3 convolutions, followed by a ReLU. Finally, a 1x1 convolution is employed at the final layer to convert each 64-component feature vector to the appropriate number of classes [26].

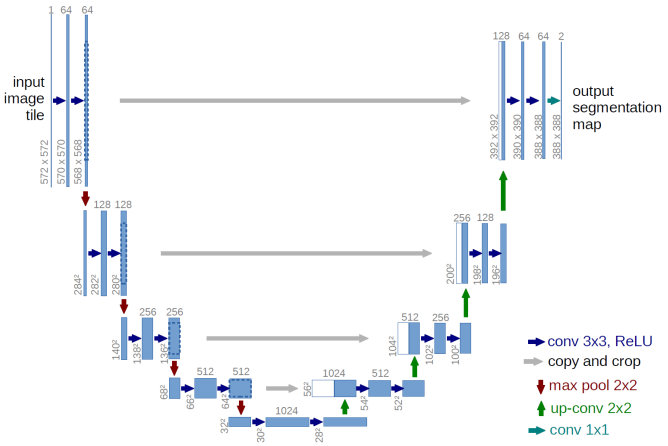


Fig. 2: U-Net Architecture [26]

In the current research work, a modified version of the U-Net model was introduced incorporating batch normalisation between each convolutional block (see Fig.3). A recent study shows that implementing batch normalisation improves the training process [8]. So, after each 3x3 convolution block, batch normalisation is introduced with the rest of the network following the same structure as U-Net.

Encoders:

1) ResNet: It is well known that deep convolutional neural networks are great at identifying low-level, mid-level, and high-level features from images, and stacking more layers generally gives us better accuracy but it raises the problem of exploding gradients [9]. Originally these problems were

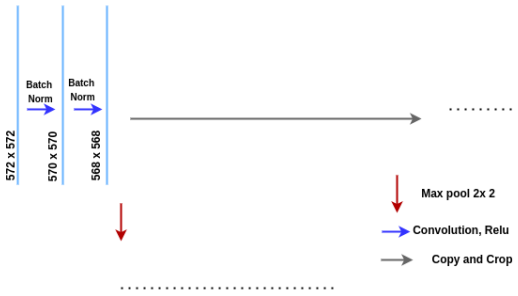


Fig. 3: Modified U-Net Architecture

addressed through normalised initialisation and intermediate initialisation, which enable networks with more layers to converge. However, it has been shown that when deep networks start to converge accuracy gets saturated and then degrades [11]. This issue was addressed by taking the layers from the deeper model and adding identity layers to it [12] (Fig.4 shows an example residual block using these identity connections). The resulting deep model should then not generate a more significant training error than its counterpart, since the extra layers were merely the identity layers. Using a deep residual learning framework, a significant improvement on the benchmark ImageNet challenge was achieved [12].

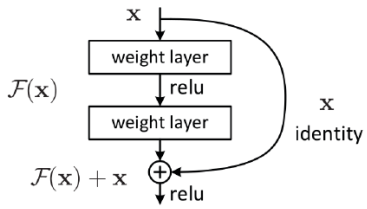


Fig. 4: A residual learning building block [12]

2) ResNext: As the number of hyper-parameters rises, so does the complexity and difficulty of developing architectures, which is particularly the case when there are several neural network layers involved. The VGG-nets [27] use a simple yet effective strategy based on stacking blocks of the same shape in order to construct intense networks. ResNets [12] use the same technique, stacking modules with the same structure on top of one another. This fundamental rule causes hyper-parameter possibilities to become progressively limited and the relevance of depth in neural networks to become more apparent. Because of its simplicity, this strategy may also reduce the likelihood of over-adapting the hyper-parameters to a given dataset in the future.

Unlike VGG-net, the Inception family of models [16] used Split-Transform-Merge topologies that are capable of achieving high accuracy with little theoretical complexity [29]. Using split-transform-merge techniques, Inception modules are predicted to be able to match the representational capability of densely packed layers but at a far lower computational power. However, various characteristics and hyper-parameters must be considered when adapting Inception architectures for new datasets and tasks, and this may be challenging. There have been several issues with implementing Inception models, including the fact that each transformation has its own set of filter numbers and widths. Despite this, careful pairings of these components provide outstanding neural network designs [33]. Similarly to VGG and ResNets, ResNext uses a repeating layers strategy while exploiting the split-transform-merge mechanism in an easy, extensible way. A module in the ResNext network performs a set of transformations, each on a low-dimensional embedding, whose outputs are aggregated by summation. The transformations to be aggregated are all of the identical topologies (see Fig.5). Because of this architecture, a large number of transformations may be implemented without a specific design.

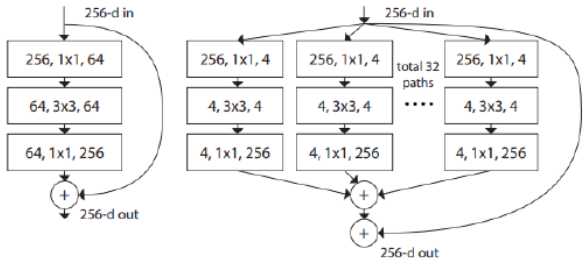


Fig. 5: Left: ResNet block; Right: ResNext block with cardinality = 32 [33]

3) DenseNet: Convolutional Neural Networks are now the dominant model for computer vision tasks. Although CNNs were initially developed two decades ago, recent improvements and iterations make the neural networks deeper, and even exceed 100 layers (e.g. ResNet, Highway Networks). Due to the more extensive networks, the information about the input can vanish and wash out by the time it reaches the end. Many new publications addressed these problems, including models such as ResNets [12] and Highway Networks [28] that pass signals from one layer to another via identity layer. Despite the diverse network topologies and training procedures, these methods provide short routes from the previous layers to the following layers.

DenseNet is a novel design that employs a straightforward connection pattern: to enable maximum information flow between network levels all DenseNet layers are physically connected with matching feature-map sizes (see Fig.6). To maintain the feed-forward nature, each layer receives new inputs from previous levels and transmits its own feature maps to succeeding layers. Unlike ResNets, DenseNet layers never combine features using summation before passing them into the next layer; instead, they combine features via concatenation.

IV. Implementation

A. Training

All the networks described in Table I were designed and developed using the Pytorch framework with the help of the

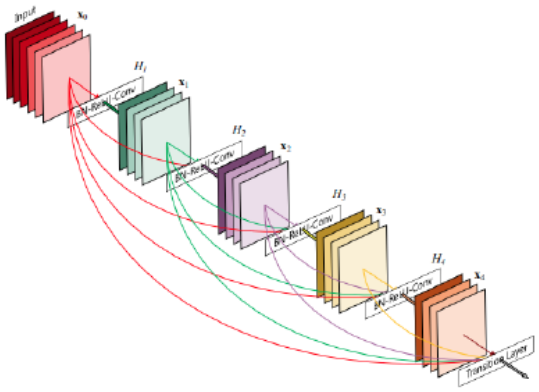


Fig. 6: A 5 layer dense block with growth rate k = 4 [15]

Segmentation Models package [34] for various encoders. The training of the models was done on an Ubuntu 21.04 LTS Linux system with 12GB of GPU (nVidia GeForce RTX 3060), Ryzen 7 processor, and 32GB of RAM. The training and testing dataset images were pre-processed to increase the dataset size and convert them to a standard set of dimensions (as described in Section III-A). The data augmentation strategy used changes the orientation of the images to increase the number of samples. The current research didn't use any other data augmentation techniques so as to prevent image distortion.

Initially, we trained each model architecture on the training dataset and evaluated each epoch of the model on the validation set. The hyperparameters for training the model are a learning rate of 1e-4 and a batch size of 3 images (due to GPU memory limitations). Each model was trained for 20 epochs using the Adam optimiser [19] to adjust the network weights and a combination of Dice Loss and the Binary Cross-Entropy loss function to check the loss of the network while training (see Equation 1).

$$\text{Our loss function} = \text{BCE loss} + \text{Dice loss} \tag{1}$$

$$\text{BCE loss}(y, \hat{y}) = -(y \log(\hat{y}) + (1 - y) \log(1 - \hat{y}) \tag{2}$$

$$\text{Dice loss}(y, \hat{y}) = 1 - \frac{2(y * \hat{y}) + 1}{y + \hat{y} + 1} \tag{3}$$

where, y is the ground truth pixel class and $\hat{y}$ is the pixel class predicted by the model.

Note in the Dice loss equation (Equation 3), one is added in both the numerator and denominator to ensure that the function is not undefined in edge case scenarios such as when $y = \hat{y} = 0$.

B. Evaluation

Models are split for training purposes into two types - those with no pre-trained weights and those with pre-trained weights. Except for the modified U-Net, the ImageNet weights were utilised as initial weights in the second set of models. To validate the results of the trained model, this study uses a test dataset containing a subset of the TNBC dataset all with dimensions of 512 x 512 pixels.

TABLE II: Final training and validation losses for all the models after 20 epochs

Encoder	Final Train Loss	Final Valid Loss
Modified U-Net	0.448	0.486
ResNet-50	0.465	0.604
ResNet-101	0.493	0.622
ResNext-50-32 4d	0.480	0.604
ResNext-101-32 8d	0.472	0.629
DenseNet-121	0.450	0.537
DenseNet-201	0.453	0.590
Pre-trained ResNet-50	0.322	0.436
Pre-trained ResNet-101	0.337	0.455
Pre-trained ResNext-50-32 4d	0.338	0.436
Pre-trained ResNext-101-32 8d	0.317	0.441
Pre-trained DenseNet-121	0.359	0.431
Pre-trained DenseNet-201	0.349	0.433

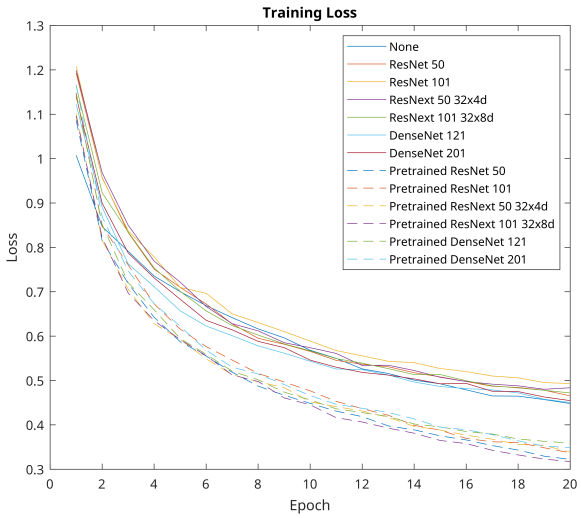


Fig. 7: Model training loss per epoch

Fig.7 illustrates the training loss of all the models over time. The results presented are the median of 10 runs of the training process (however, the performance of the models was extremely consistent - with a maximum of 2% difference in the validation loss between different runs). Table II shows that the model with the best training loss is the U-Net with pre-trained ResNext 101 encoder, but the model with the best validation loss is the U-Net with the pre-trained DenseNet 121 encoder. Also worth noting is that fully training complex encoders from scratch requires more data and more epochs (as can be seen in the difference in training and test performance between the basic modified U-Net and the U-Nets using more complex backbones). Fig.7 and Fig.8 show that after 20 epochs models with a pre-trained encoder have almost converged, whereas the non-pre-trained encoder models are still improving.

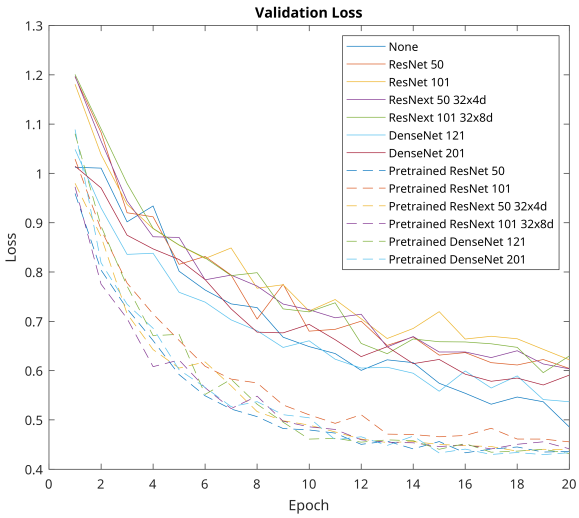


Fig. 8: Model validation loss per epoch

V. Results and Discussion

The model performance was evaluated using a range of metrics including the Jaccard Index (also known as Intersection over Union - a measure of segmentation performance), F1 score, Precision, and Recall. Table III shows that the majority of pre-trained encoder-based models outperformed the non-pre-trained encoders. The pre-trained DenseNet-121 model delivers the best Jaccard Index of 0.669, the highest F1 score of 0.800, and the highest accuracy of 0.939; however, the modified U-Net model provides a superior Recall score and runs significantly faster than other models. Additionally, the pre-trained ResNet-50 model provides greater Precision than the other models.

Fig.9 shows and example of how the modified U-Net performs in the segmentation process for a sample image from the test dataset. Detailed observation shows that the model struggles to separate overlapping nuclei (shown as red boxes in Fig.9). Fig.10 and Fig.11 show the performance of pre-trained ResNet-50 and the pre-trained DenseNet-121 on the segmentation task using the sample sample image. You can see from these figures that even the most sophisticated models have difficulties in distinguishing between overlapping nuclei (again, highlighted by red boxes in Fig.10 and Fig.11). The key reason for DenseNet-121's superior performance is due to its architectural characteristics, which include features passing through to all of the layers allowing the model to deliver better-segmented results than other models. In addition, as compared to other models, the DenseNet-121 model is a bit slower due to extra computational load introduced by this methodology. The modified U-Net model is the quickest of the models presented because of its straightforward design, which allows it to process 129 images per second. On the basis of Fig.8, we can also conclude that the pre-trained models converged more quickly than the non-pre-trained models. Despite the performance differences generated by the different models shown in Table III, all of the models had issues with the segmentation of overlapping nuclei. In order to make more enhancements and developments to the same project, we may use post-processing approaches (such as watershed transformations) to reduce the noise and improve some of the overlapping issues that have been identified.

VI. Conclusion

This research study demonstrates the utility of the transfer learning strategy in the segmentation of cell nuclei in Triple-Negative Breast Cancer patients. The encoders ResNet, ResNext, and DenseNet were chosen to be used in the research, which was conducted on various datasets. These experiments demonstrate that pre-trained models could be trained more quickly and subsequently surpass non-pre-trained models in terms of performance. When comparing pre-trained models to non-pre-trained models, the pre-trained models achieve better outcomes in terms of Precision, Recall, and F1 score. Results show that Deep Learning models trained on various organ cell images may yield excellent performance for breast cancer tumour cells on nuclei segmentation tasks even when trained on different organ cell images. When compared to the most modern encoder-based models, the improved U-Net produces similar results while being much quicker both to train and to use in prediction.

A key limitation of this study is the small number of unique WSIs used in training and testing the segmentation algorithms. Additional training data would allow us to increase the accuracy of all the networks tested in this work, as well as improving model generalisation and reducing overfitting. Recently, researchers adopted two key techniques to enhance the datasets: scalable crowdsourcing of the annotation process [1], and creating additional annotated images using Generative Adversarial Networks (GANs) [35]. The outcomes of the existing models can potentially be improved in the future using these two methodologies.

References

[1] Mohamed Amgad, Lamees A Atteya, Hagar Hussein, Kareem Hosny Mohammed, Ehab Hafiz, Maha AT Elsebaie, Ahmed M Alhusseiny, Mohamed Atef AlMoslemany, Abdelmagid M Elmatboly, Philip A Pappalardo, et al. Nucls: A scalable crowdsourcing, deep learning approach and dataset for nucleus classification, localization and segmentation. *arXiv preprint arXiv:2102.09099*, 2021.

[2] Andreea Anghel, Milos Stanisavljevic, Sonali Andani, Nikolaos Papandreou, Jan Hendrick Rüschoff, Peter Wild, Maria Gabrani, and Haralampos Pozidis. A high-performance system for robust stain normalization of whole-slide images in histopathology. *Frontiers in medicine*, page 193, 2019.

[3] Cornelia J Baines, AB Miller, C Wall, DV McFarlane, IS Simor, R Jong, BJ Shapiro, L Audet, M Petitclerc, and D Ouimet-Oliva. Sensitivity and specificity of first screen mammography in the canadian national breast screening study: a preliminary report from five centers. *Radiology*, 160(2):295–298, 1986.

[4] Sercan Çayır, Serim Hande Tarcan, Samet Ayaltı, Salar Razavi, Fariba Khameneh, Şükrü Burak Çetin, Gizem Solmaz, Gülşah Özsoy, Çisem Yazıcı, Cavit Kerem Kayhan, et al. Nuclei segmentation in hematoxylin and eosin (h&e)-stained histopathological images using a deep neural network. In *2020 28th Signal Processing and Communications Applications Conference (SIU)*, pages 1–5. IEEE, 2020.

TABLE III: Evaluation metrics of each model

Encoder	Jaccard	F1	Recall	Precision	Accuracy	FPS
Modified U-Net	0.662	0.795	0.821	0.788	0.936	129.525
ResNet-50	0.588	0.737	0.743	0.757	0.922	71.619
ResNet-101	0.574	0.725	0.756	0.732	0.916	40.562
ResNext50-32-4d	0.581	0.731	0.793	0.708	0.914	75.894
ResNext101-32-8d	0.571	0.722	0.818	0.676	0.907	36.545
DenseNet-121	0.594	0.742	0.812	0.707	0.912	30.929
DenseNet-201	0.610	0.756	0.791	0.747	0.921	19.537
Pre-trained ResNet-50	0.647	0.783	0.750	0.833	0.938	72.825
Pre-trained ResNet-101	0.648	0.784	0.815	0.773	0.932	42.558
Pre-trained ResNext50-32-4d	0.649	0.782	0.799	0.795	0.932	76.976
Pre-trained ResNext101-32-8d	0.660	0.792	0.793	0.810	0.938	36.140
Pre-trained DenseNet-121	0.669	0.800	0.806	0.806	0.939	32.241
Pre-trained DenseNet-201	0.662	0.794	0.785	0.819	0.938	19.608

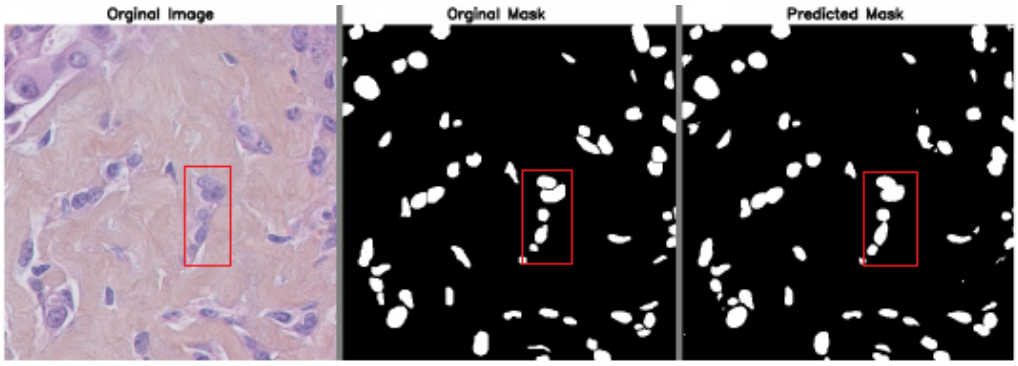


Fig. 9: Modified U-Net segmentation results

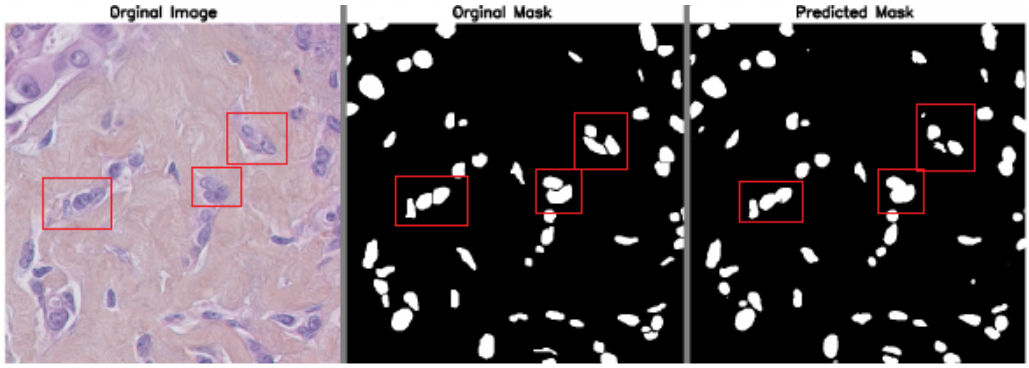


Fig. 10: Pre-trained ResNet-50 segmentation results

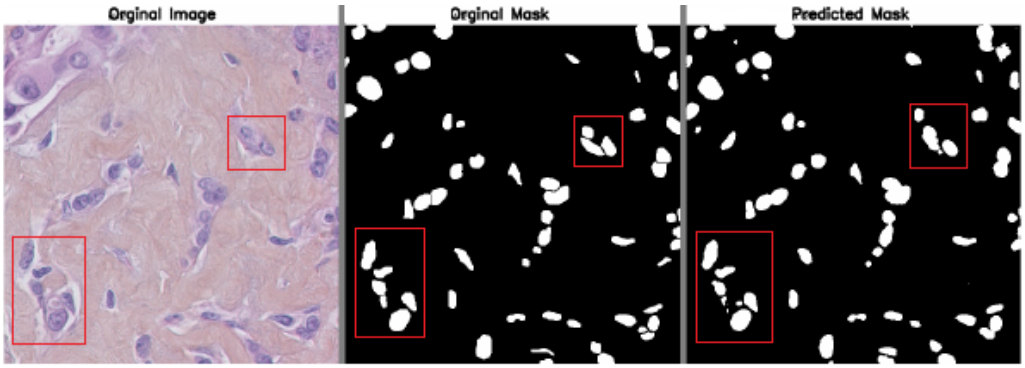


Fig. 11: Pre-trained DenseNet-121 segmentation results

[5] Tuan Le Dinh, Seong-Geun Kwon, Suk-Hwan Lee, and Ki-Ryong Kwon. Breast tumor cell nuclei segmentation in histopathology images using efficientunet++ and multi-organ transfer learning. *Journal of Korea Multimedia Society*, 24(8):1000–1011, 2021.

[6] Théo Estienne, Maria Vakalopoulou, Stergios Christodoulidis, Enzo Battistela, Marvin Lerousseau, Alexandre Carre, Guillaume Klausner, Roger Sun, Charlotte Robert, Stavroula Mougiakakou, et al. U-resnet: Ultimate coupling of registration and segmentation with deep nets. In *International conference on medical image computing and computer-assisted intervention*, pages 310–319. Springer, 2019.

[7] Christina Fitzmaurice, Christine Allen, Ryan M Barber, Lars Barregard, Zulfiqar A Bhutta, Hermann Brenner, Daniel J Dicker, Odgerel Chimed-Orchir, Rakhi Dandona, Lalit Dandona, et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study. *JAMA oncology*, 3(4):524–548, 2017.

[8] Shang-Hua Gao, Qi Han, Duo Li, Ming-Ming Cheng, and Pai Peng. Representative batch normalization with feature calibration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8669–8679, 2021.

[9] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256. JMLR Workshop and Conference Proceedings, 2010.

[10] Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, et al. Recent advances in convolutional neural networks. *Pattern Recognition*, 77:354–377, 2018.

[11] Kaiming He and Jian Sun. Convolutional neural networks at constrained time cost. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5353–5360, 2015.

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[13] Nehmat Houssami and Kylie Hunter. The epidemiology, radiology and biological characteristics of interval breast cancers in population mammography screening. *NPJ Breast Cancer*, 3(1):1–13, 2017.

[14] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications, 2017.

[15] G Huang, Z Liu, L Van Der Maaten, and KQ Weinberger. Densely connected convolutional networks. paper presented at: Proceedings of the ieee conference on computer vision and pattern recognition (cvpr). 2017.

[16] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR, 2015.

[17] Alexandr A Kalinin, Vladimir I Iglovikov, Alexander Rakhlin, and Alexey A Shvets. Medical image segmentation using deep neural networks with pre-trained encoders. In *Deep learning applications*, pages 39–52. Springer, 2020.

[18] Anusree Kanadath, J Angel Arul Jothi, and Siddhaling Urolagin. Histopathology image segmentation using mobilenetv2 based u-net model. In *2021 International Conference on Intelligent Technologies (CONIT)*, pages 1–8. IEEE, 2021.

[19] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[20] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012.

[21] Neeraj Kumar, Ruchika Verma, Deepak Anand, Yanning Zhou, Omer Fahri Onder, Efstratios Tsougenis, Hao Chen, Pheng-Ann Heng, Jiahui Li, Zhiqiang Hu, et al. A multi-organ nucleus segmentation challenge. *IEEE transactions on medical imaging*, 39(5):1380–1391, 2019.

[22] Andrew Lagree, Majidreza Mohebpour, Nicholas Meti, Khadijeh Saednia, Fang-I Lu, Elzbieta Slodkowska, Sonal Gandhi, Eileen Rakovitch, Alex Shenfield, Ali Sadeghi-Naini, et al. A review and comparison of breast tumor cell nuclei segmentation performances using deep convolutional neural networks. *Scientific Reports*, 11(1):1–11, 2021.

[23] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[24] Scott Mayer McKinney, Marcin Sieniek, Varun Godbole, Jonathan Godwin, Natasha Antropova, Hutan Ashrafian, Trevor Back, Mary Chesus, Greg S Corrado, Ara Darzi, et al. International evaluation of an ai system for breast cancer screening. *Nature*, 577(7788):89–94, 2020.

[25] Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE transactions on medical imaging*, 38(2):448–459, 2018.

[26] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[27] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[28] Rupesh K Srivastava, Klaus Greff, and Jürgen Schmidhuber. Training very deep networks. *Advances in neural information processing systems*, 28, 2015.

[29] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.

[30] Laszlo Tabar, Tony Hsiu-Hsi Chen, Amy Ming-Fang Yen, Sam Li-Sheng Chen, Jean Ching-Yuan Fann, Sherry Yueh-Hsia Chiu, May MS Ku, Wendy Yi-Ying Wu, Chen-Yang Hsu, Yu-Ying Chen, et al. Effect of mammography screening on mortality by histological grade. *Cancer Epidemiology and Prevention Biomarkers*, 27(2):154–157, 2018.

[31] Mingxing Tan and Quoc V. Le. Efficientnet: Rethinking model scaling for convolutional neural networks, 2020.

[32] PA Van Luijt, EAM Heijnsdijk, Jacques Fracheboud, LIH Overbeek, MJM Broeders, Jelle Wesseling, GJ den Heeten, and HJ de Koning. The distribution of ductal carcinoma in situ (dcis) grade in 4232 women and its impact on overdiagnosis in breast cancer screening. *Breast Cancer Research*, 18(1):1–10, 2016.

[33] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1492–1500, 2017.

[34] P Yakubovskiy. Segmentation models pytorch. github repository, 2020.

[35] Xin Yi, Ekta Walia, and Paul Babyn. Generative adversarial network in medical imaging: A review. *Medical image analysis*, 58:101552, 2019.

[36] Danny R Youlden, Susanna M Cramb, Nathan AM Dunn, Jennifer M Muller, Christopher M Pyke, and Peter D Baade. The descriptive epidemiology of female breast cancer: an international comparison of screening, incidence, survival and mortality. *Cancer epidemiology*, 36(3):237–248, 2012.

[37] Qiao Zhang, Zhipeng Cui, Xiaoguang Niu, Shijie Geng, and Yu Qiao. Image segmentation with pyramid dilated convolution based on resnet and u-net. In *International conference on neural information processing*, pages 364–372. Springer, 2017.