A Comparative Study of Convolutional Neural Network Architectures for Detecting Prostate Cancer

K. M. Safin Kamal[1], Mysha Maliha Priyanka[1], Alfe Suny[1], Maimuna Akter Liza[1], Sanjeda Sara Jennifer[1], and *Ahmed Wasif Reza[1]

[1]Department of Computer Science and Engineering, East West University, Dhaka-1212, Bangladesh

Emails: kmsafinkamal@gmail.com, malihapriyanka81@gmail.com, alfeysuny@gmail.com, maimunaakterliza@gmail.com, jennifer.sanjedasara@gmail.com, wasif@ewubd.edu

* Corresponding Author

Abstract. One of the most prevalent common cancers among men is prostate cancer. Therefore, early detection is crucial for effective treatment. This study aims to detect prostate cancer using four Convolutional Neural Network (CNN) architectures. We evaluated our trained models and found a lower Root Mean Square Error (RMSE) of 2.9960 on the validation set indicating that our model can accurately detect prostate cancer in medical images. Our study suggests a promising prostate cancer detection model that could help improve patients' early diagnosis and treatment outcomes.

Keywords: Prostate cancer, CNN, EfficientNet-B1, MobileNet-V2, ResNet-18, DenseNet -121.

1 Introduction

Prostate cancer is a common and potentially fatal disease, especially in men. It is a significant contributor to global mortality rates caused by cancer [1]. In 2020, it was estimated that prostate cancer would claim the lives of 375,304 people globally. The risk increases with age and physical inactivity. The prostate gland plays a role in sperm production and is composed of four lobes. The more glandular tissue present, the greater the risk of prostate cancer [2]. While most prostate tumors are slow-growing, aggressive forms can be extremely dangerous. Treatment options include radiation therapy, surgery, and a combination of both [3]. Medical imaging, such as transrectal ultrasound and radiomic analysis, has significantly improved the screening, diagnosis, and evaluation process. In this paper, our primary focus is:

- Using the largest whole-slide image (WSI) dataset available for a prostate biopsy to grade the Gleason score and International Society of Urological Pathology (ISUP).
- We have proposed techniques using convolutional neural networks (CNN) to detect prostate cancer from biopsies at an early stage.
- We have compared CNN architectures like MobileNet-v2, DenseNet-121, and ResNet-18 with our proposed architecture, EfficientNet-B1 in order to find the most accurate architecture to detect prostate cancer.

The content of this paper is as described in the following: Section 2 discusses the related work of prostate cancer. After that, in Section 3, the research methodology

is discussed. The results and discussion are in the fourth section. Finally, the conclusion and future scope of the paper are written in Section 5.

2 Related work

In 2017 [4], Srdan Jovic proposed an Extreme Learning Machine (ELM) machine learning algorithm. The ELM model outperformed ELN, GP, and ANN in predicting prostate cancer probability, but the computational complexity of the machine learning algorithms was not reported.

Lal Hussain [5] proposed a combination of feature-extracting strategies to improve prostate cancer detection performance. It provides high accuracy in prostate cancer detection using the proposed methods in 2017. The study focused on utilizing the ROC curve to determine the accuracy of MRI in detecting prostate cancer but did not provide information on the computation of Positive Predictive Value (PPV) and Negative Predictive Value (NPV).

In 2020 [6], Richard Li trained a gradient-boosted tree model and applied SHapley Additive exPlanations (SHAP) values to model predictions. Retrospective analyses using SHAP values to interpret machine learning models' predictions revealed that patients at higher risk of prostate cancer tend to receive treatment more often than those at low or very low risk, potentially introducing a bias.

In 2021, a study by SAQIB IQBAL [7] on deep learning was compared to conventional methods for identifying prostate cancer. The study used deep learning, SVM, and RF, including transfer learning, to achieve a high accuracy rate on a dataset of 230 patient MRI scans. However, the study has limitations such as single institution data reliance, specific imaging modality restriction, lack of comparison to human expert results, and the need for further validation in clinical settings.

Olivier Regnier-Coudert in their study [8], compared the performance of several machine learning classifiers on a dataset of prostate cancer biopsy images. The study demonstrated the classifiers' ability to accurately predict the pathological stage of prostate cancer, emphasizing the significance of feature selection in machine learning analysis.

3 Research Methodology

Prostate cancer is a regular disease in males, with 1 million new cases and 350,000 deaths annually. Precise diagnoses are crucial, and Gleason grading is used for tissue evaluation. Preparing the dataset is an important part of the entire procedure, followed by the implementation phase. This section will elaborate on the dataset description, data preprocessing, research methodology, architecture selection, and implementation in detail.

3.1 Dataset Description

The dataset was collected from two medical centers and created by researchers with a vast dataset of 22,000 digitized biopsy images stained with H&E. This publicly

accessible collection represents a crucial resource in the field of digital pathology. The dataset was curated by Radboud University Medical Center and the Karolinska Institute for a specific challenge. The evaluation process involves identifying and classifying cancerous tissue based on Gleason patterns such as three, four, or five, which determine the tumor's growth patterns. The assigned Gleason score is then converted to an ISUP grade (ranging from 1 to 5), which influences treatment decisions. Your treatment strategy is aided by this score from your cancer treatment team. The Gleason grading system, known as the ISUP grade, is used to evaluate the aggressiveness of prostate cancer based on the condition's microscopic appearance. Based on the Gleason score, this new method assigns a cancer grade between 1 and 5. The less grade the cancer has, the less probable it is to spread. This score assists your cancer treatment team in planning your care. However, inconsistencies between pathologists pose a risk of missed tumors or overgrading, leading to unnecessary treatment or missed diagnoses. Table 1 shows the Gleason score and the ISUP grade.

Table 1. Gleason score and ISUP grade

Gleason score	ISUP grade
3+3	1
3+4	2
4+3	3
4+4	4
3+5	4
5+3	4
4+5	5
5+4	5
5+5	5

Digital pathology involves converting microscopy slides into digital images for examination using computational methods such as AI and machine learning, offering benefits like data analysis and high-resolution imaging.

3.1.1 Gleason patterns

To create digital slide images, WSI (Whole Slide Imaging) involves scanning thin tissue sections stained with hematoxylin and eosin (H&E). These images show glandular and connective tissues, with glands appearing as white openings. The Gleason grading system uses these visual features to assess prostate tissue health [9]. Higher grades indicate a diminishing healthy glandular structure, and the system categorizes levels into 3, 4, and 5. Detailed descriptions of the patterns are provided below and demonstrated in Figure 1:

a. Folded epithelium on benign prostate glands arranged in groups, characterized by tiny regular nuclei and pale cytoplasm.
b. Prostatic adenocarcinoma in Gleason Pattern 3: infiltration of dark cytoplasmic glands with enlarged nuclei, maintaining glandular differentiation.

c. Adenocarcinoma of the prostate in Gleason Pattern 4: irregular epithelial sheets, multiple lumina, fused glands, indicating partial loss of glandular differentiation.
d. Prostate adenocarcinoma in Gleason Pattern 5: a significant decrease in glandular differentiation, solid sheets or threads of cancer cells, distinct hematoxylin and eosin staining.

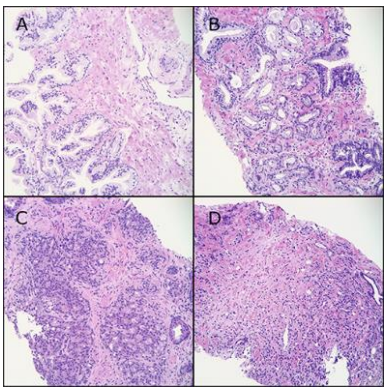

Fig. 1. Grading patterns are given as examples. (A) Glands that are in good health. Gleason patterns 3 to 5, respectively (B-D).

3.1.2 ISUP grade

A solitary biopsy is used to determine a Gleason score for prostate cancer. The most prevalent and second-most frequent patterns are identified by pathologists. It is considered if the second pattern accounts for at least 5% of the total area. The highest grade is always included. For instance, in Figure 2, a Gleason score of 4 + 3 becomes 4 + 4 if pattern 3 is less than 5%. A biopsy would have a score of 4 + 5 = 9 if it had a 3% Gleason 5, 37% Gleason 3, and 60% Gleason 4 rating. Based on the existing recommendations from the ISUP, Gleason scores are combined into an ISUP grade on a scale from 1 to 5, following Table 1 and Figure 2. In this competition, we utilize the designation ISUP grade 0 if there is no evidence of cancer in the sample.

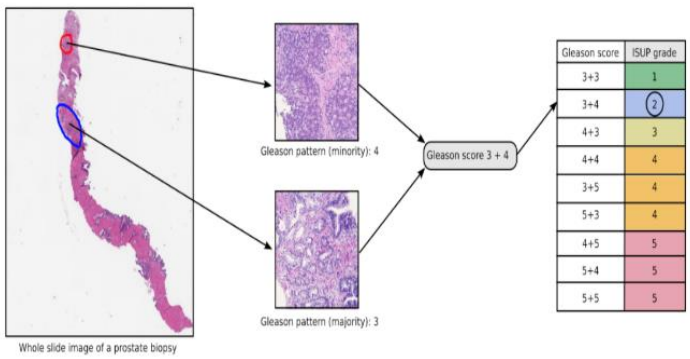

Fig. 2. Scoring Gleason and ISUP from a WSI of prostate biopsy

3.2 Dataset Preprocessing

To prepare the dataset, the data from biopsy slides is first loaded using OpenSlide. OpenSlide allows loading specific regions without having to load the entire image into memory. Prior to loading a biopsy slide, it is necessary to open the file. Once a file is open, data can be retrieved from it at any desired position and level. To provide basic information about a slide and display a thumbnail, a compact function is created [10]. Biopsies may have different rotations but no clinical significance. Color variations in biopsies are common due to diverse laboratory methods. Figure 3 shows the biopsy photo in detail after zooming in.

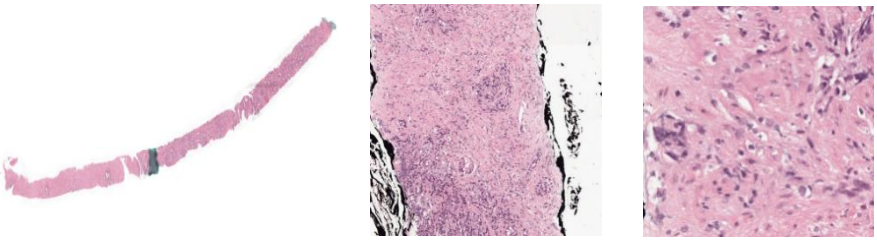

Fig. 3. Zooming the biopsy photo

OpenSlide allows the extraction of patches from slides at any position and offers a read region function for specific region loading, such as a 512x512 patch from level 0. Mask information in training slides varies between centers. Radboudumc individually labels prostate glands using semi-automatic deep learning algorithms, while Karolinska labels regions based on pathologists' remarks. Radboudumc's masks have noise and are loosely supervised. A helper function can display fundamental mask data by converting int labels to RGB colors. Matplotlib with a personalized color map can be used to efficiently visualize cancer regions as shown in Figure 4.

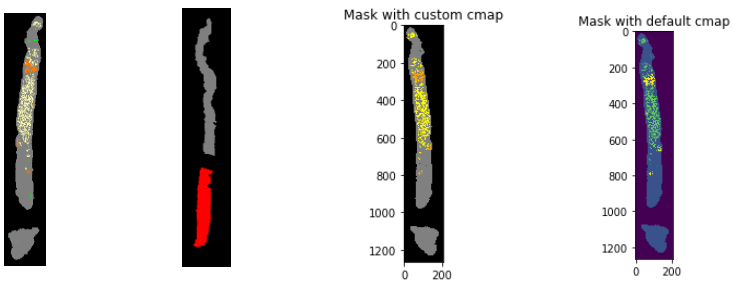

Fig. 4. Output of cmap masks

3.3 Model Selection

The ability of CNN designs to capture spatial hierarchies makes them suitable for use with image datasets. They employ pooling layers for downsampling and convolutional layers to identify characteristics like edges and textures. As a result, they can analyze pictures quickly, extract useful characteristics, and do tasks like object detection and classification with high accuracy. Choosing the right neural network architecture is vital as it affects accuracy, speed, and memory. Considering

the problem domain, data, and resources is crucial. An efficient network saves time, facilitates deployment, and reduces costs. For our research, we chose EfficientNet-B1, MobileNet-V2, DenseNet-121, and ResNet-18 architectures and compared these four neural networks.

3.3.1 EfficientNet-B1

Efficient Net is one of the CNN architectures that was developed by researchers at Google in 2019. Efficient Net models are named according to their scaling coefficient, which ranges from 0 to 8. A higher scaling coefficient indicates a larger and more complex model. For example, EfficientNet-B0 has the smallest number of parameters and computations, while EfficientNet-B7 has the largest. Figure 5 represents the structure of the EfficientNet architecture.

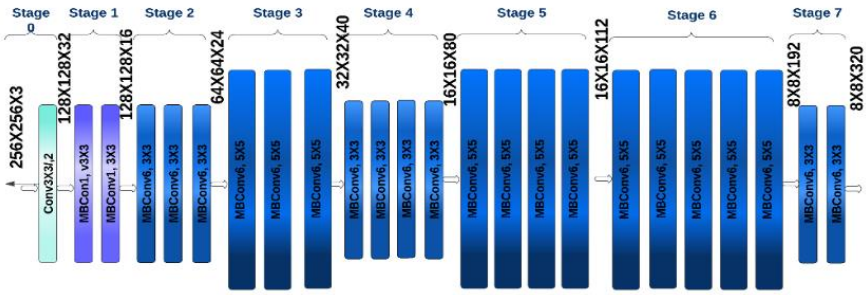

Fig. 5. Structure of Efficient Net

One of the main benefits of Efficient Net B1 is its scalability, which allows it to perform well on datasets with a large volume of images. EfficientNet-B1 uses a combination of depth-wise convolutions and feature fusion to extract meaningful features from images while minimizing the number of computations required. Feature selection helps a model enhance prediction accuracy and minimize computational complexity [11]. This allows it to process large volumes of images more efficiently than previous models while maintaining high accuracy.

3.3.2 MobileNetV2

MobileNet is a CNN architecture designed for efficient mobile and embedded vision applications [12]. Unlike traditional CNNs, it utilizes depth-wise separable convolutions, significantly reducing parameters and computational expenses [13]. MobileNet has different versions, including MobileNetV1, MobileNetV2, and MobileNetV3. Figure 6 shows the generic structure of this architecture.

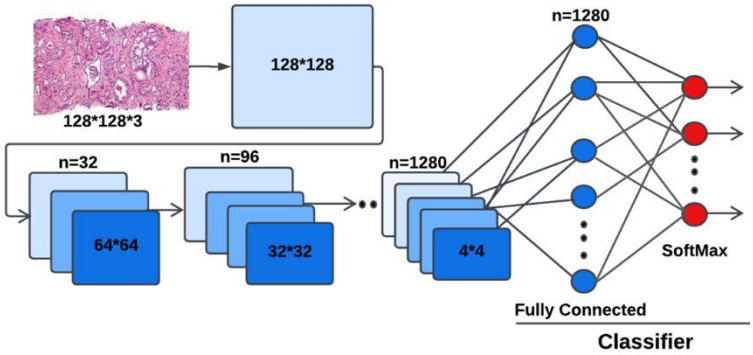

Fig. 6. MobileNet Network Architecture

3.3.3 DenseNet-121

DenseNet is a convolutional neural network used for computer vision tasks. It creates dense connections between layers, allowing the network to learn from more information. DenseNet-121 has four dense blocks and three transition layers. Each dense block consists of convolutional layers with batch normalization and ReLU activation. The output of each layer is added to the outputs of previous layers and given as input to the next layer. The first block generates 56x56 feature maps, which are halved in size for each subsequent block [14]. Figure 7 presents the overall architecture of DenseNet.

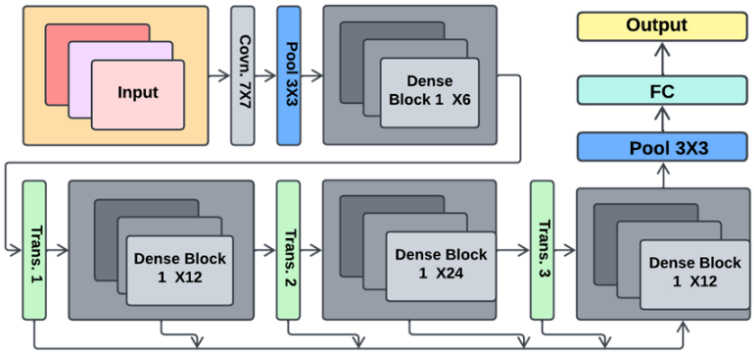

Fig. 7. DenseBlock and transition layer of DenseNet-121

3.3.4 ResNet-18

Residual Network, or in short, ResNet-18, is a convolutional neural network designed for tasks related to computer vision and image analysis such as classification, objection detection, semantic segmentation, and transfer learning. The architecture of ResNet-18 consists of 18 layers. As neural networks get deeper, they suffer from a problem of degradation that affects their ability to learn, and ResNet was built to handle this issue. Residual Connection allows information from earlier nodes to flow into deep layers by bypassing subsequent layers. [15]. The structure of the ResNet-18 architecture is presented in Figure 8.

Fig. 8. ResNet-18 Architecture

3.4 Implementation

After applying musk to each of the images, the model is trained using EfficientNetB1, MobileNetV2, DenseNet121, and ResNet18 CNN architectures on a given dataset using PyTorch. We tuned the hyperparameters and used the Adam Optimizer, setting the learning rate to 0.001, and ran 5 epochs. All four models follow the same specification, and the summary is presented in Table 2.

Table 2. Specifications of Model

Model Name	EfficientNetB1, MobileNetV2, DenseNet121, ResNet18
Optimizer	Adam Optimizer
Learning Rate	0.001
Evaluation Metrics	Root Mean Squared Error (RMSE)
Epochs	5

4 Results and Discussion

After the training phase using five epochs on all four models, the results showed that EfficientNet-B1 gave the lowest loss and RMSE, as shown in Table 3, which clearly means that it is better than the other models. The lower the loss and RMSE, the better the performance of a model. For more transparency, we also calculated the average loss and RMSE, which are given in Table 4.

Table 3. Performance Evaluation of Different CNN Architectures

	MobileNet-V2		**DenseNet-121**		**ResNet-18**		**EfficientNet-B1**	
Epoch	**Loss**	**RMSE**	**Loss**	**RMSE**	**Loss**	**RMSE**	**Loss**	**RMSE**
1	1.7374	3.0222	1.7266	3.0072	1.9106	3.9061	1.7241	2.9637
2	1.7208	2.9825	1.7182	3.0096	1.9090	3.9060	1.7105	3.0408
3	1.7201	3.0272	1.7190	3.2540	1.9090	3.9060	1.7072	2.9632
4	1.7164	3.0178	1.7174	3.9236	1.9090	3.9060	1.7052	2.9839
5	1.7153	3.0165	1.7155	3.1168	1.9090	3.9060	1.7047	3.0684

Table 4. Average Loss and RMSE of the Architectures

	MobileNet V2	DenseNet-121	ResNet-18	EfficientNet-B1
Average Loss	1.7220	1.7394	1.9093	1.7103
Average RMSE	3.0132	3.2622	3.9060	2.9960

Table 4 shows that even after the average of 5 epochs, EfficientNet-B1 is giving 1.7103 loss and 2.9960 RMSE, which is the lowest compared to the other models. Therefore, we can conclude that EfficientNet-B1 performs best for detecting prostate cancer.

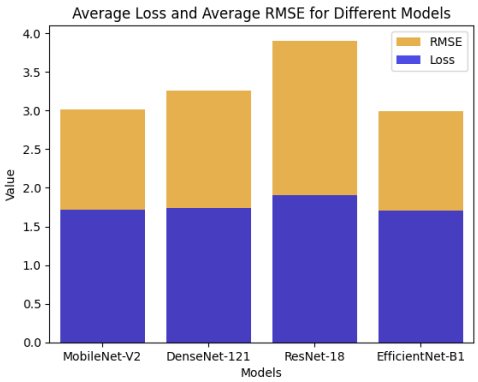

Fig. 9. Performance Comparison of the Architectures

Figure 9 presents a visual representation of the architecture in terms of loss and RMSE. We tried to compare the performances of our models with other state-of-the-art works. However, there is no existing work so far that has evaluated their model based on RMSE and loss; rather, they have different metrics for evaluation.

5 Conclusion

In conclusion, the prostate cancer dataset includes segmentation masks for images that can be used to locate cancerous regions accurately. The dataset contains Isup_grade and Gleason_score labels that indicate the severity of cancer on a scale of 0–5 and a more detailed rating system, respectively. We implemented MobileNet-V2, DenseNet-121, and ResNet-18 and acquired the root mean square error (RMSE) on the validation set to be 3.0132, 3.2622, 3.9060, and 2.9960, respectively. EfficeintNet-B1 performed well on our dataset among all the CNN architectures to accurately detect prostate cancer. Overall, the dataset and visualizations provided valuable insights into the assessment of prostate cancer grades and could aid in the development of more effective diagnostic and classification models. There are certain limitations to our paper as well. There were only two data providers in the dataset. Further, more data providers can be included. Our model comparison was limited to CNN architectures. We intend to extend our work by incorporating other deep-learning models and machine-learning algorithms. Furthermore, only one segmentation mask has been applied. In the future, other masks can be applied and compared. We also intend to perform more hyperparameter tuning in order to train

the model for more epochs instead of five, which might lead to more accurate results.

References

1. K. Sklinda, M. Frnczek, B. Mruk, and J. Walecki, ‘‘Normal 3T MR anatomy of the prostate gland and surrounding structures,’’ Adv. Med., vol. 2019, pp. 1–9, May 2019.
2. Mahumud RA, Alam K, Dunn J, Gow J. The burden of chronic diseases among Australian cancer patients: Evidence from a longitudinal exploration, 2007-2017. PLoS One. 2020 Feb 12;15(2):e0228744. doi: 10.1371/journal.pone.0228744. PMID: 32049978; PMCID: PMC7015395.
3. N. M. Schultz, K. O’Day, R. Sugarman, and K. Ramaswamy, ‘‘Budget impact of enzalutamide for nonmetastatic castration-resistant prostate cancer,’’ J. Managed Care Specialty Pharmacy, vol. 26, no. 4, pp. 538–549, Apr. 2020.
4. Jović S, Miljković M, Ivanović M, et al (2017) Prostate cancer probability prediction by machine learning technique. Cancer Investigation 35:647–651. doi: 10.1080/07357907.2017.1406496
5. Xya Hussain L, Ahmed A, Saeed S, et al (2018) Prostate cancer detection using machine learning techniques by employing a combination of features extracting strategies. Cancer Biomarkers 21:393–413. doi: 10.3233/cbm-170643
6. Li R, Shinde A, Liu A, et al (2020) Machine learning–based interpretation and visualization of nonlinear interactions in Prostate cancer survival. JCO Clinical Cancer Informatics 637–646. doi: 10.1200/cci.20.00002
7. Iqbal S, Siddiqui GF, Rehman A, et al (2021) Prostate cancer detection using deep learning and traditional techniques. IEEE Access 9:27085–27100. doi: 10.1109/access.2021.3057654
8. Regnier-Coudert O, McCall J, Lothian R, et al (2012) Machine learning for the improved pathological staging of prostate cancer: A performance comparison on a range of classifiers. Artificial Intelligence in Medicine 55:25–35. doi: 10.1016/j.artmed.2011.11.003
9. Partin A, Kattan M, Subong E, Walsh P, Wojno K, Oesterling J. Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer: a multi-institutional update. Journal of the American Medical Association 1997;277(18):1445–51.
10. P. Lakhani et al., “Hello world deep learning in medical imaging,” J. Digital Imaging 31(3),283–289 (2018).
11. Y. Fan, D. Shen, R.C. Gur, R.E. Gur and C. Davatzikos, Com-839 pare: Classification of morphological patterns using adaptive 840 regional elements, IEEE Trans Med Imaging 26 (2007), 93–841 105. doi: 10.1109/TMI.2006.886812.
12. Chiu Y-C, Tsai C-Y, Ruan M-D, et al (2020) Mobilenet-SSDv2: An improved object detection model for embedded systems. 2020 International Conference on System Science and Engineering (ICSSE). doi: 10.1109/icsse50014.2020.9219319
13. Lksdnc A. A. Abbasi, L. Hussain, I. A. Awan, I. Abbasi, A. Majid,M. S. A. Nadeem, and Q.-A. Chaudhary, ‘‘Detecting prostate cancer using deep learning convolution neural network with transfer learning approach,’’ Cognit. Neurodynamics, vol. 4, pp. 523–533, Apr. 2020.
14. G. Huang, Z. Liu and L. van der Maaten, “Densely Connected Convolutional Networks,” 2018.
15. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi: 10.1109/cvpr.2016.90