

Article

Automatic Tumor Identification from Scans of Histopathological Tissues

Mantas Kundrotas, Edita Mažonienė and Dmitrij Šešok *

Department of Information Technologies, Vilnius Gediminas Technical University, Sauletekio Al. 11,
LT-10223 Vilnius, Lithuania

* Correspondence: dmitrij.sesok@vilniustech.lt

Abstract: Latest progress in development of artificial intelligence (AI), especially machine learning (ML), allows to develop automated technologies that can eliminate or at least reduce human errors in analyzing health data. Due to the ethics of usage of AI in pathology and laboratory medicine, to the present day, pathologists analyze slides of histopathologic tissues that are stained with hematoxylin and eosin under the microscope; by law it cannot be substituted and must go under visual observation, as pathologists are fully accountable for the result. However, a profuse number of automated systems could solve complex problems that require an extremely fast response, accuracy, or take place on tasks that require both a fast and accurate response at the same time. Such systems that are based on ML algorithms can be adapted to work with medical imaging data, for instance whole slide images (WSIs) that allow clinicians to review a much larger number of health cases in a shorter time and give the ability to identify the preliminary stages of cancer or other diseases improving health monitoring strategies. Moreover, the increased opportunity to forecast and take control of the spread of global diseases could help to create a preliminary analysis and viable solutions. Accurate identification of a tumor, especially at an early stage, requires extensive expert knowledge, so often the cancerous tissue is identified only after experiencing its side effects. The main goal of our study was to expand the ability to find more accurate ML methods and techniques that can lead to detecting tumor damaged tissues in histopathological WSIs. According to the experiments that we conducted, there was a 1% AUC difference between the training and test datasets. Over several training iterations, the U-Net model was able to reduce the model size by almost twice while also improving accuracy from 0.95491 to 0.95515 AUC. Convolutional models worked well on groups of different sizes when properly trained. With the TTA (test time augmentation) method the result improved to 0.96870, and with the addition of the multi-model ensemble, it improved to 0.96977. We found out that flaws in the models can be found and fixed by using specialized analysis techniques. A correction of the image processing parameters was sufficient to raise the AUC by almost 0.3%. The result of the individual model increased to 0.96664 AUC (a more than 1% better result than the previous best model) after additional training data preparation. This is an arduous task due to certain factors: using such systems' applications globally needs to achieve maximum accuracy and improvement in the ethics of AI usage in medicine; furthermore if hospitals could give scientific inquiry validation, while retaining patient data anonymity with clinical information that could be systemically analyzed and improved by scientists, thereby proving AI benefits.

Keywords: machine learning; automatization; histopathology; medical imaging



Citation: Kundrotas, M.; Mažonienė, E.; Šešok, D. Automatic Tumor Identification from Scans of Histopathological Tissues. *Appl. Sci.* **2023**, *13*, 4333. <https://doi.org/10.3390/app13074333>

Academic Editors: Levente Adalbert Kovács and László Szilágyi

Received: 6 March 2023

Revised: 27 March 2023

Accepted: 28 March 2023

Published: 29 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Machine learning integration in the medical field, especially image analysis for histopathology and cancer research could make a huge impact due to the possibility of rapid and more accurate results. In recent years, advancements in technology have revolutionized the health system enabling it to use digitized imaging in order review patient data through computer systems and applications. Digital content can be simply stored without losing its quality and reshared between health specialists, especially because

of the increasing number of image analysis applications and large image compression tools [1–3]. Unfortunately, switching from glass to digital analysis requires new expensive hardware, extensive memory size and especially trained technicians [4]. To this day a legitimate comparison research method using accepted AI techniques that could be used for histopathology does not exist, especially because of expensive hardware and eligible access to patient data. Additionally, the main challenge is that different WSI's scanners lower the results, and the most significant barrier to the widespread use of AI approaches in the clinic is undoubtedly the limited generalizability of algorithms [5]. Despite this, medical professionals might help develop new techniques that could be applied in clinical practice as the start of implementing AI in histology [4].

A cancer, also called tumor, is a formation of altered, unregulated, unlimitedly proliferating but not clearly defined clusters of abnormal cells that lead to terminal illnesses. There are three main types of tumors: benign, premalignant, and malignant. Benign tumors are defined as less dangerous and harmful due to the fact of being noninvasive, however, certain cases declare that they can become malignant. Malignant tumors usually grow rapidly, penetrating and destroying healthy tissues, spreading to distant organs and finally, metastasize [6,7]. There is a variety tumor identification methods: magnetic resonance imaging (MRI), computed tomography (CT), Single-Photon-Emission Computed Tomography (SPECT), and other medical imaging technologies are used to determine the exact type, location, and level of threat that cancer-damaged tissue has [8,9]. International research for cancer statistics in 2020 showed over 19.3 million new cancer cases from all over the world. Even though the medical field is rapidly growing and progressing, the World Health Organization (WHO) declared that cancer is the main leading cause of death, unfortunately reflecting the size of population [10].

In histopathology analysis, specialists must focus on objective and accurate identification of diagnosis due to the complexity of diseases. Digital imaging allows to analyze histopathological specimens with a quantification of slides technique using machine learning methods as deep learning. Such promising technology of multi-layered artificial neural networks allows to perform a quantity of tasks adapting consistent tasks. Other studies have suggested that using digital histopathological images makes it possible to identify cancer cells [11–13]. To increase the efficiency of tumor identification research with the reduction in errors, automatization using mathematical methods could be a solution. Taking into consideration this relevant case, AI experts [14–17] have already developed a variety of fully exploitable technologies that take place in the medical field. For instance, to detect or identify a brain, spine and chest tumor, the following methods are applicable: K-means, SVM, Level Set, Adaboost, Naive Bayes classifier, ANN classifier, convolutional neural networks, multilayer perceptron neural networks for analyzing magnetic resonance, and computed tomography type data analysis [18–22].

In this article, we use machine learning for recognizing cancer cells that are applied over histopathological images consisting immense number of pixels. We are aware of deep learning methods that can benefit the identification and recognition of tumor cells. The machine learning algorithm goes over the given data, or as in our cases an image, or an image that shows cancer affected tissue, following an algorithm mode to ascertain from the training data and assign to make a prediction for further work. If this algorithm advances and increases the number of executions of improved and correct diagnosis—then we consider it as a learned task [23]. To explain machine learning, we use certain terms: supervised, unsupervised, semi-supervised, and reinforcement learning [24] algorithms [25]. Furthermore, deep learning gives us opportunities to explore wide-ranging data [26]. To gain the best possible accuracy, we must use an artificial neural network (ANN) as one of the machine learning techniques that allows to remove the artifacts (errors) that naturally appear in different types of data. Such errors, especially in the medical field, are one of the most common issues for misinterpretation. An ANN can be explained as assembly of neurons that are arranged in a sequence of multiple layers. The activation of an ANN begins from an input layer whose main aim is to correctly choose the format and transfer

data to different layers until it reaches the last or the output layer. Fascinatingly, all other layers have different numbers of neurons and are identified as hidden because they play an important role allowing to learn data structure and give the ability to classify its type. In general, the operation of neural networks is very similar to linear regression, which means that each individual neuron is just like a linear regression model consisting of input data, its weights, bias and finally the output [27,28]. Medical image analysis empowered by deep learning, especially using a convolutional neural network (CNN), and its significant results for different types of cancer detection in histopathology scans (two-dimensional data), has drawn attention from scientists all over the world [29–31]. The main reason is that it gives the ability to clinicians to make correct health diagnoses and receive precise analysis of illnesses that can be compared with previous samples. Training deep learning models, including CNN, demand significant training size and computing resources due to the profuse number of pixels in an image [1,32]. One of the subtypes of CNN, called the residual convolutional neural network (ResNet), can work with large datasets (the training and test material) even if the neural network increases in depth (extend in number of stacked layers), benefiting in a reduction in error rate [33,34]. Another subtype of CNN is the DenseNet, that reduces the vanishing gradient effect in DNN. This network is formed of dense blocks that are detached by a transition layer that minimizes the size of the delivered feature maps that will be transferred to the following layers. To compare ResNet and DenseNet we must understand how data are being directed between the layers. A transition layer, which aims to minimize the size of the generated feature maps that will be sent to the following layers, separates each dense block in a DenseNet network [35]. A residual learning unit is a feature of ResNet, which was created to counteract deep neural network degeneration. This system is built as a feed-forward network with a shortcut link that enables the addition of new inputs and outputs. The main benefit is that it increases the classification accuracy without complicating the model's actual design [36].

Following reviews of different tumor identification systems, it became obvious that selecting the machine learning system that is best suited to produce the desired result based on the current data comes only after the data has been properly analyzed and processed [37–39]. There is a lot of space for improvement because most of the systems that have been presented only process very small amounts of data using straightforward neural networks [40,41]. Applying more recent advancements in data collection and processing, which have been successfully used in other scientific fields, along with a new generation of convolutional neural networks with a more complex structure, guarantees the most accurate outcome for tumor detection.

In this research, we applied machine learning to histopathological images that consisted of immense numbers of pixels to recognize cancer cells. Automatic systems can solve complex problems that require an extremely fast response, accuracy, or take place on tasks that require both a fast and accurate response at the same time. Such systems and image pre-processing that are based on machine learning algorithms can be adapted to work with medical images, especially whole slide images (WSIs), that will allow clinicians to review a much larger number of health cases in a shorter time and give the ability to identify the preliminary stages of cancer or other diseases, thereby improving health monitoring strategies.

The main contributions of this paper are as follows:

- Modern augmentation and image preprocessing methods to analyze WSIs,
- Creating an adaptive U-Net model architecture,
- Adding different optimizers for best outcoming result in AUC.

The paper is organized as follows:

- In Section 2, we did a review of machine learning models, architectures, algorithms, and other techniques that can be used for histopathological WSIs,
- Section 3 outlines the methodology, that step by step describes the machine learning model, dataset, and accuracy requirements for further experiments,
- Section 4 consists of the design of the experiments, the main values, graphical and statistical results,

- In Section 5, we list the major accomplishments and talk about the outcomes,
- In Section 6, we conclude our work and identify potential work directions.

2. Related Work

2.1. Medical Imaging

The field of computer-aided medical image analysis, particularly the recognition of cancer in histopathological images, has received extensive attention from researchers [42–44]. It has been heavily focused on the area of computer-aided medical image analysis, new and improved applications in bioinformatics [45], particularly the identification of cancer in histopathological images as it enhances the benefits of quality diagnostics for patients [46]. One of the medical disciplines that has seen a sizable number of deep learning applications is digital pathology and microscopy, especially as advancements in technology have given us an opportunity in using histopathological scans-digital images [47–49].

Bright prospects have been demonstrated for the detection, segmentation, and classification of diseases, especially finding cancer cells using newly developed techniques that provide pathologists with decision support and most likely are guiding to more accurate results. A new cross-domain field of graph-based deep learning, which seeks to learn informative representations of medical images in an end-to-end manner [50], has emerged because of the increased attention given to the adaptation of deep learning from images. Convolutional neural networks, a deep learning method that excels at tasks like histology segmentation, are being used more frequently in pathology [51,52].

Thus, artificial intelligence tools based on machine learning techniques could fundamentally alter pathologists' and medical professionals' workflow in the future, whether through computer-aided diagnosis or simply by accelerating laborious manual tasks. The current rigid classification and analysis system that enables difficult visual recognition tasks, like separating tumors from normal tissue in slides and classifying different types of tumors [53] may be replaced by a more analytical and flexible model that incorporates radiological, biological, and clinical variables based on deep learning [54,55].

2.2. Machine Learning Models

The study of algorithms and statistical models that computer systems use to accomplish a task without being explicitly programmed is known as ML and its main purpose is to learn from data. Algorithms and statistical models are divided into groups according to the main principle of operation. For instance, a function that maps an input to an output that is learned through supervised learning using sample input-output pairs. Algorithms that require outside assistance are those that fall under the category of supervised machine learning. They use labelled training data made up of a collection of training examples to infer a function [56]. Unsupervised learning undoubtedly is the biggest goal and struggle for medical diagnostic systems [57,58]; unfortunately, most of these models simply cannot be applied as they are capable of working only with a limited amount of imaging data [59].

Our attention was drawn to U-Net model, that is one of the most widely used CNN architectures and differs from the others for two reasons. First, it is made of two parts: first—descends resolution, second, after each block, a maximum signal selection operation is used, which doubles the width and height. After this operation, the number of channels used is doubled. The second part uses the same convolutional blocks, but the local resolution reduction operation uses an up-sample operation that doubles the height and width. Later, this layer is combined with high-frequency information from the dimensionally appropriate first part of the network. By passing information from the left part of the network to the right in this way, the network can gain more detailed and accurate information [60]. The U-Net architecture can surely be implemented for histopathological data for nuclei segmentation and, according to Ibtihaj et al. [61], it significantly improves segmentation performance.

Learning Rate and Planning Algorithms

Another major challenge is the learning rate (LR). While using huge amounts of data such as in histopathological medical images, the training time obviously takes longer and to achieve the most optimal performance we need to choose a learning rate as one of the hyper-parameters [62]. The LR describes any positive number that is less than the integer one; it is also a parameter that controls the weight adjustment [63]. If the learning step is specified as one or very close to one, it means that when the backpropagation calculation is performed, the original weight matrix of the calculated layer will be adjusted using the full error matrix. This means that adjusted model weights for the current data input will return zero error. This is acceptable for one part of the data set from which the gradient is calculated, but this tightening to one set at a time means that if the next set will not likely to be like the current one, output will show an error and subsequently the gradient itself will be very large. Such a process will cause the model to diverge. The opposite of this process would slow the convergence, or the model will turn to stagnation. This happens when too small a learning step is chosen. During this process, the calculated model weight correction matrix is multiplied by a number very close to zero, thus performing an extremely minimal weight correction. Compared to too large a learning step, a small constant close to zero does not cause an impact. The model is learning, but the process is slightly slower than it could be [64].

From a practical perspective, a deep artificial neural model occasionally performs remarkably well on training data but dreadfully on testing data. This behavior happens when the model overlearns; it remembers the numerical values of the model's inputs and classes, nevertheless it fails to learn the necessary features from the training data. Regularization techniques are used to address this issue because a model that has been overtrained cannot be used for real medical cases. Comparing to other machine learning algorithms, regularization is a relatively simple operation. Its purpose is to prevent the model from memorizing incoming data by adjusting the values of the weight matrix throughout the model. Scientists use regularization even as a consistent strategy while analyzing histopathological imaging [61,65].

Another useful idea for using DNN is transfer learning, and the use of this technique to solve such issues as extracting certain features or tuning [66]. Basically, when there is a small amount of labeled data in the target dataset, it is demonstrated to be a powerful tool to combat overfitting. The authors of [67] observed that transfer learning performed successfully full training, achieving this in a shorter period also lowering the labor required for manual data labeling.

Resources are the most common issue when using deep neural networks to solve complex problems. As we said above, it is necessary to wisely choose and use a sufficiently large and complex deep neural network in situations where the data are extremely complex and extensive.

3. Materials and Methods

3.1. Proposed Model

To achieve objective results, it is important to maintain a standardized training process. In the field of machine learning, even small changes can dramatically determine or change the result. For equal initialization of weights, we used Xavier's algorithm [67–69]. It led us to fix a random seed so we could use the same model with the exact same initialization weights for different experiments. We did not limit or fix how long learning can last because the model might have different architectures. The only limitation we set was the verification algorithm that counts the duration of learning; the main reason was that it halts learning if the model has not improved in the previous ten iterations. Besides that, we used the SGD optimizer [70] with acceleration, and a cyclic learning step with warm-up was used for training purposes. For the first 1000 iterations, the model learns with a fixed learning step of 1×10^{-4} and then gradually increases to 5×10^{-3} over 2400 iterations and descends to 5×10^{-5} over 3600 iterations,

shown in Figure 1. Warm-up applies only to the first epoch. The total number of iterations per epoch that we used was 6000, while the batch size was 64.

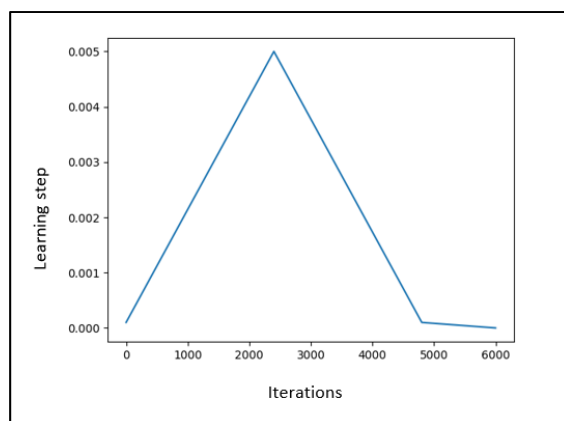


Figure 1. An example model that learns with fixed learning rate.

In the interim, testing and validation data sets were not augmented or otherwise adjusted. Additionally, 5 million artificially augmented images were created from the training dataset for training. In addition to that, for experiments that compare models, but not individual model accuracy by adjusting data or hyper parameters, we used Binary Cross-Entropy objective function as well as AUS metrics. Before creating a new neural network architecture, changing the objective function, or training data, it is necessary to find out whether the data set is suitable. The fastest way to do this is to use trained models. Using this process, it is possible to draw conclusions about how each architecture affects the final accuracy. For this work, the ResNet, DenseNet, MobileNet, EfficientNet, Inception architectures [71–73] were selected from their derived versions from the Tensorflow library. All these models were already trained on the ImageNet dataset [74], which means that each model already had 1000 output classes.

Proper data processing is one of the most crucial components of training deep neural networks. Considering a number of more than 200,000 WSIs seems to be enough for training data, however as we previously stated, larger convolutional models easily remember the statistics of incoming data and tend to retrain. As a result, using augmentation techniques such as cropping, vertical and horizontal conversion, random rotation by 0–360 degrees, and contrast adjustment shown in Figure 2, let us create synthetic data with the same class value but a different representation. The original information was tightly compressed in *.h5 format. For each set type—training, validation, or testing—were divided into two files: actual images, and the values for the classes according to the corresponding index are in the other. These assemblies were scanned using the h5py library. Afterwards, a generator function was created, which simultaneously reads from two files and combines the read values into a pair. Next, these data were appropriately augmented and converted to 32-bit floating point format depending on the type. For final training, the data are typically compressed into some deep learning-friendly data format. Because the Tensorflow library was used in this work, all data were stored in *.tfrecord format. Before the images were sent to the input layer of the model, the data were normalized. Each pixel in the image was given a value between $[-1, 1]$ during normalization; the Equation (1) is shown below.

$$x = x/127.5 - 1 \quad (1)$$

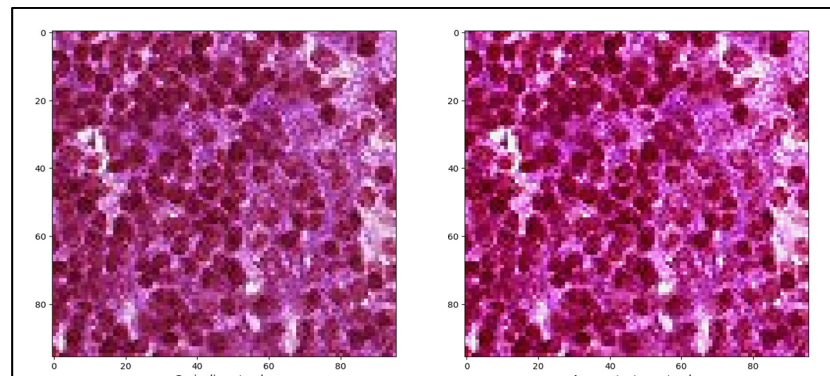


Figure 2. Contrast adjustment applied on image.

3.2. Dataset

We used the PatchCamelyon (<https://github.com/basveeling/pcam>), 2020 (accessed on 2 January 2022) image classification dataset. It is made up of 327,680 color images (96×96 pixels each) that were taken from a histopathologic image of a lymph node section as shown in Figure 3. Each image has a binary label that indicates whether metastatic tissue is present. The entire data set was divided into two parts. The first one was for training only; it consisted of 262,144 photos. The second one was for testing; it contained 32,768 photos. The dataset authors assured that all the splits have an even distribution of positive and negative examples, and there was no overlap in the whole slides images (WSI) between any of the splits. When it comes to labeling, a patch with a positive label has at least one pixel of tumor tissue in its central region. During the experiments, it was observed that excessive sharpening, darkening or contrast enhancement of the image erases all the information identifying the cancer from the histopathological image. Furthermore, it was observed that the influence of heavy data on the total error of the training data set during learning was extremely small.

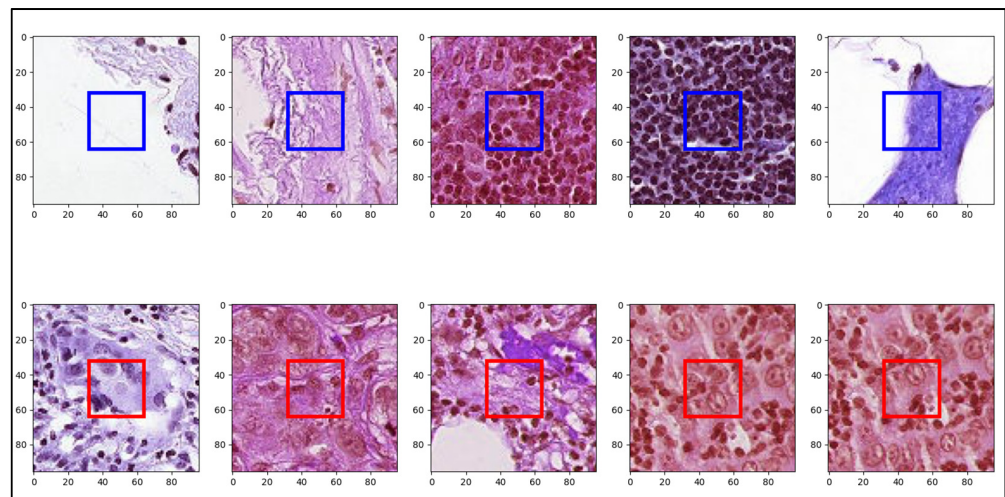


Figure 3. Random samples of the PatchCamelyon dataset that are stained with hematoxylin-eosin (H&E), blue (healthy tissue) and red squares (cancerous tissue) showing central regions.

3.3. Accuracy Calculation

In practice, the performance of a neural network is evaluated by several criteria: calculation speed, overall accuracy, adaptability when applied to a new data set, etc. As it is very important to consider that final accuracy and reliability are the most important factors for medical image analysis, calculation speed and memory space were not evaluated in our work. We considered four estimates, wherein TP stands for true positives, FP for false

positives when the model mixes negatives with positive class, and FN for false negatives when the model mixes positive with negative classes (equations listed sequentially):

- Precision using Equation (2),

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

- Recall using Equation (3),

$$\text{Recall} = \frac{TP}{FP + FN} \tag{3}$$

- F1-score using Equation (4),

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

- AUC. Measures the quality of the model in terms of sensitivity and accuracy over the entire set of limits.

4. Experiments and Results

4.1. Experimental Setup

As we stated previously, the U-Net type architecture is currently one of the most popular templates to search for a suitable architecture. It allows you to easily check the signal information of different levels individually and by combining high and low frequency signals in separate levels; in our work, we called it the M-model, shown in Figure 4.

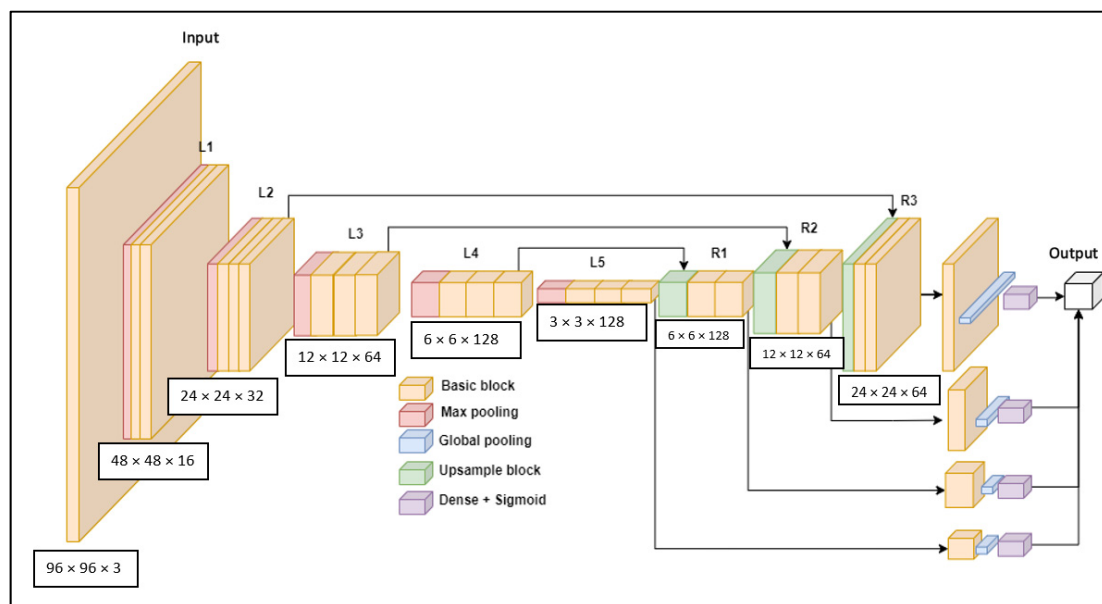


Figure 4. The M-model based on the U-Net network architecture.

By default, U-Net networks form a complete U-shape—they transmit the entire input signal to the output of the model, which is applicable in this work due to efficiency and the type of task itself. Higher resolution output is usually used for image reproduction or segmentation tasks. In this research, classification was performed from low-resolution images, so the high-resolution output will not be superior. The following changes were made to the intermediate modules of the network structure, that we will call the E-Module, shown in Figure 5. Compared to the commonly used standard ResNet module, the E module was more efficient and faster—two 3×3 convolutions were changed to one compressed 3×3 convolution and one 1×1 convolution, with the same number of

channels as the number of input layer channels. Furthermore, the typical ReLU (rectified linear unit) activation was changed to PReLU (parametric rectified linear unit) [75]. In this way, the model will be able to apply the best activation for each layer according to the direction of the gradient. Together with this activation function, it was possible to see the behavior of the model by analyzing the values of the activation coefficients. Finally, a Dropout layer was added at the end of the module, which performs a dual regularization: it reduces the signal bandwidth and allows the dropout factor to be increased to 1 at the end of training to discard supposedly unnecessary parts of the model.

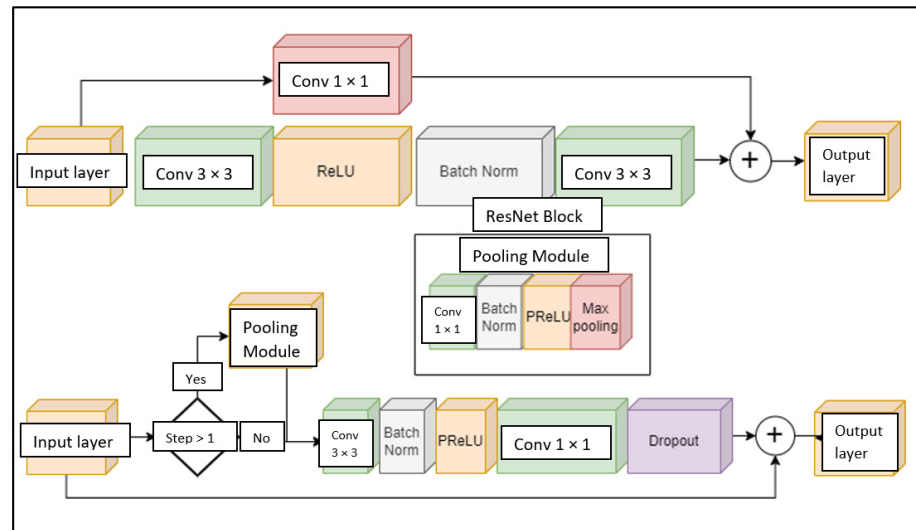


Figure 5. E-module based on ResNet block design.

Additionally, the entire network structure was reworked into a dynamic format. Based on the principle of the EfficientNet architecture [76], the rules for growing the width and length of the network were established and adaptive regularization were applied with Equations (5)–(9) shown below:

- Number of filters (5):

$$\text{Number of filters} = \max\left(16, \frac{f \cdot s \cdot F}{h \cdot 2}\right) \cdot g \tag{5}$$

- Number of blocks (6):

$$\text{Number of blocks} = \max\left(2, \frac{\text{Number of filters}}{64}\right) \tag{6}$$

- Exclusion factor (7):

$$\text{Exclusion factor} = \max\left(0.05, \frac{\text{Number of filters}}{m/2}\right) \tag{7}$$

- L2 regularization (8):

$$\text{L2 regularization} = 1 \times e - 5 + \frac{f}{8} \times 0.0001 \tag{8}$$

where f is the base number of filters, F is the filter multiplier, s is the filter multiplier, h is the height of the network in pixels, g is the number of convolutional groups of the network. The number of blocks indicates how many internal blocks will make up the mesh module after each decrease or increase in height and width of the mesh. Exclusion factor—indicates how many neurons will be turned off in percentage in each block, where m indicates the

maximum possible number of filters. The L2 regularization [77,78] specifies the value of the L2 regularization constant for all network convolutions.

To understand whether all the changes made to the model gave advantages compared to ResNet50 or DenseNet121 models, the following experiments were performed. According to the already presented formula for changing the structure of the model, three sizes for f were selected: 8, 16, 32. Sizing allows us to see the areas where the network performs too poorly, overlearns, and where it performs optimally. It is additionally important to find out from which part of the network the best result can be obtained. Therefore, five output layers were added to the already existing network. Accordingly, we noticed a trend while comparing the results of all outputs. All models gave their best performance from the outputs of L5 and R1. This suggests that using a model with such a structure for this task, a higher resolution not only does not provide additional benefits, but also spoils the result. We can be assured that the most suitable model size for this task is $f = 16$, as in the first test, models where $f = 8$ and $f = 32$ performed worse. Among the model outputs, it can also be seen that higher resolution did not provide enough benefits.

Overall, compared to the first experiment, most of the results have improved, especially the output AUC of L5 has increased from 0.95279 to 0.95341. According to these results, it can be confidently stated that the U-Net type architecture did not provide enough benefits for this task. Moreover, comparing the obtained results with ResNet50 and Densenet121, the new model was already superior in terms of accuracy and learning speed—the new model reached the maximum result in eight epochs, which none of the previously described models was able to do when learning with newly initialized weights. According to the results of the first tests, the U-Net model was modified. We removed the right part of U-Net layers and all intermediate connection, also fixed size multiplier $f = 16$. Such pruning not only provided speed, but also allowed to gain greater accuracy due to less information reuse, this improved model will be called the MS-model.

4.2. Results

First, we made additional training validation performed on the artificially augmented and non-augmented datasets. The ResNet and DenseNet networks with ImageNet weights were applied for the test. Both models performed with more than 1% greater accuracy with augmented data than with non-augmented data (Table 1). It can be assumed that these models were too large for such a task and most probably overlearned as a result, however the AUC value on the training data set shown in Table 2 confirmed this. Without augmentation, the AUC was almost at unity at 0.993, while with augmented data it was only 0.975 with the ResNet50 model. Although retraining has occurred, the results show that the data generated for training was fine, and the selected augmentations were useful.

Table 1. Learning from zero with testing data set.

Model	AUC (Area under the Curve)	
	Using Augmentation	Not Using Augmentation
ResNet50	0.95001	0.93988
DenseNet121	0.95511	0.93780

Table 2. Learning from zero with learning data set.

Model	AUC (Area under the Curve)	
	Using Augmentation	Not Using Augmentation
ResNet50	0.96501	0.99297
DenseNet121	0.98891	0.99971

After making sure that the training environment and data were correct and everything was working as it should, further analysis was performed. Each selected model was trained

twice. First, using ImageNet weights and then applying Xavier weights with initialization according to the training protocol, results are shown in Table 3.

Table 3. Model architectures analysis results.

Model	AUC (Area under the Curve)	
	ImageNet Weights	Xavier Initialization Weights
DenseNet121	0.95672	0.94560
ResNet50	0.95078	0.94380
ResNet50 V2	0.95078	0.94380
MobileNetV1	0.94954	0.93855
MobileNetV2	0.95065	0.95395
Inception	0.94697	0.94608
EfficientNetB0	0.95121	0.94608
EfficientNetB1	0.93876	0.94608
EfficientNetB0 V2	0.94570	0.75981
EfficientNetB1 V2	0.94287	0.79871

Comparing with ImageNet weights, DenseNet achieved the best result. According to the learning graph, it exceeded the AUC value of 0.95 after only two epochs. The next best was ResNet50, although it did not show the second result according to the table, but it also exceeded the AUC value of 0.95 after two epochs. Collectively, this means that these two architectures were the most suitable for this task. Furthermore, the ResNetV1 and ResNetV2 models excelled the most. Although their results were not the best, they exceeded 0.94 AUC in just five epochs. This shows that the ResNet-type blocks and persistent connections were well suited for this task due to their fine gradient feedback.

A new training session was performed with the trimmed model called the MS-model. Two tests were selected. The first was by extending the training using the best weights, as shown in Table 4 and gave better results after reusing weights. As the original structure of the network remained the same, weights could easily be transferred from one network to another. The second was to train the same network with ever new weights initialization.

Table 4. MS-model results. New initialization indicates that the model weights were generated by Xavier initialization from a newly chosen random point.

Learning Iteration	AUC
Reusing weights	0.95501
New initialization 1	0.95498
New initialization 2	0.95508
New initialization 3	0.95505

A popular way to get a better result is to change the optimization algorithm. Although all training was done with the SGD optimizer, which should potentially be the best for this problem, other methods could potentially find a better solution simply because of the difference in the optimization Hyper parameters. As a result, shown in Table 5, we learned that of several optimizers, the best result was achieved with the AdamW optimizer.

Table 5. Comparison of optimization methods.

Learning Iteration	AUC
SGD	0.95510
Adam	0.95475
AdamW	0.95515
Ranger	0.95500

Based on the results of the last test that adding more regularization is better, changes were made to the model and training code. The L2 regularization of the convolutional weights used so far was increased to 1×10^{-4} , and the learning frequency cycle used in the SGD optimizer was converted to cosine. As expected, this teaching principle worked very well. After the fifth iteration, we managed to achieve an AUC of 0.95911, which was almost 0.4% better than the last best model trained with the AdamW optimizer.

In addition to that, we added the TTA [79,80] method, as the training data were adjusted very flexibly—we selected certain featuring: color channel change, vertical flip, horizontal flip, rotate -90 degrees, rotate $+90$ degrees, lighten the image by 5%, darken the image by 25%; unfortunately, the individually processed images results gained only 0.9590 AUC.

Another approach that did not require model retraining, reengineering, or other changes was model ensemble. This can be done in several ways. We proposed two methods for it. The first one, which combines the outputs of the different models according to the arithmetic mean and returns a single result, and the next one, which combines the weights of the different models into one common list of weight matrices.

5. Discussion

After all strategies and methods, we observed that different training optimizers, more heavily augmented data, learning step graph, or even a well-chosen learning starting point of the model, all influence the overall result when using them combined in an ensemble; it can be argued that the models or weights will converge if their individual performance does not overlap. This means that the models must make different errors among themselves. Therefore, it is obvious that by combining the ensemble methods of the last two experiments, it will be possible to achieve even higher accuracy.

It can be seen in Table 6 that all used methods are models that lead our MS-model to be improved from 0.95918 to 0.96675 AUC.

Table 6. Summary of results. The difference column represents the AUC difference between the type and the first starting point in DenseNet121.

Ensemble Type	AUC (Area under the Curve)	
	AUC	Difference
DenseNet121	0.95672	-
M-model training 5 outputs together	0.95405	-0.267%
M-model training 5 outputs separately	0.95491	-0.1891%
MS-model	0.95508	-0.164%
MS-model with AdamW	0.95515	-0.157%
MS-model with repeated training	0.95911	0.239%
MS-model TTA	0.96870	1.198%
MS-model ensemble	0.96592	0.920%
MS-model connecting weights	0.96240	0.568%
TTA + weights and models ensemble	0.96922	1.250%
MS-model after corrections	0.96147	0.475%
MS-model after corrections with repeated training	0.96675	1.003%
Group of ensembles from all experiments	0.96977	1.305%
Optimized ensemble based on the best model	0.97673	2.001%

To sum up, it can be stated that artificial neural networks are able to distinguish tissue areas affected by cancer quite well. The developed MS-model is more accurate and faster than most of the models presented in the “Patch Chameleon” standings as shown in Figure 6. It can be said that such accuracy was most influenced by the effective architecture of the neural network and research on combining and assembling models. As shown in Figure 7, we have reached a maximum **F1 score of 0.924** with a threshold of **0.393**; this score measured on our last best model’s accuracy (that is evaluated from precision and recall), as long as the confusion matrix is that shown in Figure 8, giving

us a result of meaning how many times our best model gave us correct predictions: true positives—15,096, true negatives—15,189, false negatives—1295, and false positives—1188. According to experiments, even with less-than-ideally prepared training data, the last ensemble method managed to exceed **0.9691 AUC**.

#	User (Team)	Created	AUC	Comment
1st	rezazad68 (LSTM-U-net)	25 June 2019	1.0000	First submission by Reza Azad
1st	JasonHuang	24 May 2019	1.0000	test submission
3rd	Debu (Test123)	19 Sept. 2019	0.9971	
4th	moncco	10 May 2019	0.9968	The first trial
5th	mantaskundrotas2	21 May 2022	0.9767	TTA + ResNet Assembly
6th	zhongqiaoyong (Dr. Strange)	20 May 2019	0.9737	
7th	GeertLitjens	15 June 2019	0.9691	DenseNet with TTA

Figure 6. “Patch Chameleon” standings. The result of this work is in the fifth place—marked 5th.

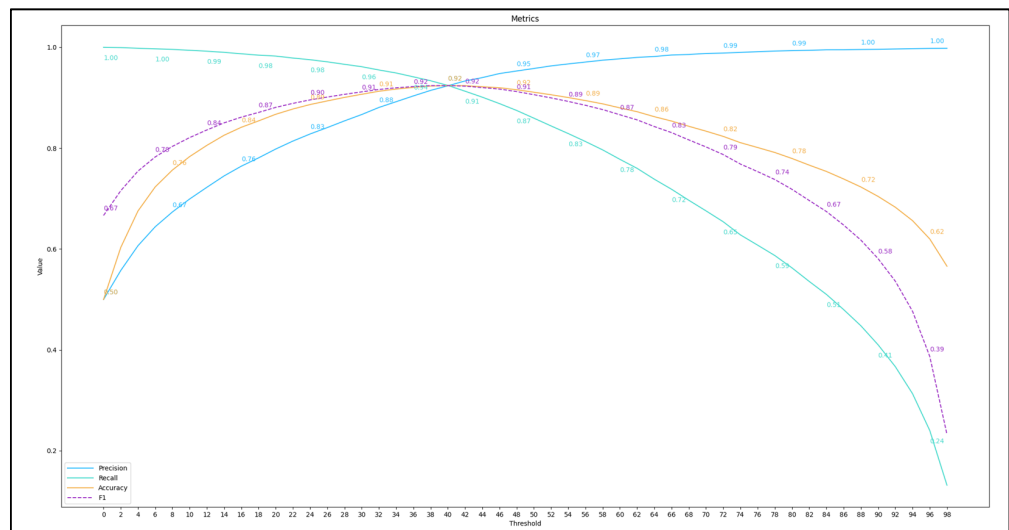


Figure 7. Maximum F1 score reached on an optimized ensemble with the best model (threshold = 0.393).

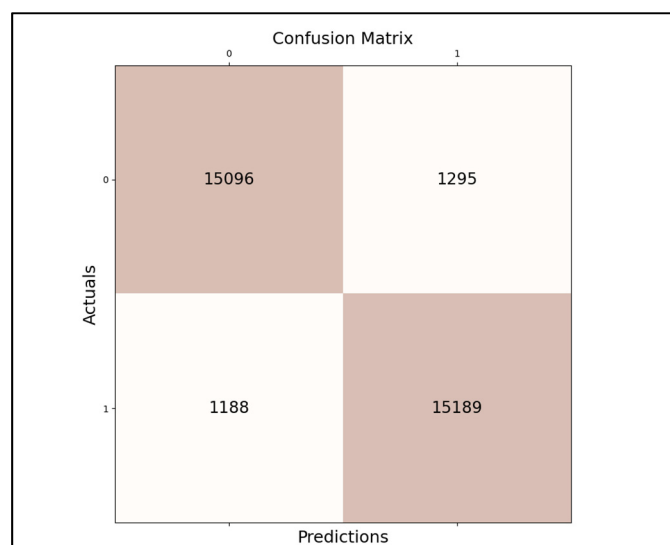


Figure 8. Confusion matrix reached with the best model.

6. Conclusions

In this work, we proposed to use ML and different neural network techniques to find a solution to WSIs histopathological data analysis. Our extensive experiments showed that the application of artificial deep neural networks for the classification of medical images, compared to other classical methods, are superior in almost all criteria. First, the CNN generalize well and perform similarly well on unseen data, even with additional constraints. From the obtained experiments, the AUC difference between the training and testing datasets was about one percent. Second, these models were highly flexible, allowing the size and architecture of the model to be experimentally tailored to the type of task. Using the M-model in several training iterations, we managed to reduce the model size by almost twice and increase the accuracy from 0.95491 to 0.95515 AUC. Third, when properly trained, convolutional models perform well on groups of various sizes. The result increased to 0.96870 with the TTA method, and 0.96977 with the addition of the multi-model ensemble. Fourth, by applying special analysis methods, it was possible to identify the shortcomings of the models and correct them. After finding that excessive and inappropriate image augmentation was detrimental to learning, a correction of the image processing parameters was sufficient to increase the AUC by almost 0.3%. Moreover, after additional training data preparation, the result of the individual model increased to 0.96664 AUC.

Due to the complexity of histopathological images, current image classification methods still lack accuracy and stability, and even the final model ensemble result of 0.97673 was not sufficient for the system to work autonomously. The results are too unpredictable, so this type of system can only be used as a guide as an image analysis tool for the physician.

Nevertheless, even such an achievement is important—it is a step closer to the ideal accuracy that would exceed human resolution.

In future work, not only another optimizing, and DNN techniques and architectures, but also imaging improving unsupervised methods could be used to gain best accurate results that further diagnostics tools for faster and more accurate cancer detection in histopathology imaging.

Author Contributions: Conceptualization, M.K. and D.Š.; methodology, M.K. and E.M.; software, M.K.; writing—original draft preparation, M.K.; visualization, M.K.; investigation, M.K. and E.M.; editing, E.M.; writing—review, supervision, project administration, funding acquisition, D.Š. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ahmed, S.; Shaikh, A.; Alshahrani, H.; Alghamdi, A.; Alrizq, M.; Baber, J.; Bakhtyar, M. Transfer learning approach for classification of histopathology whole slide images. *Sensors* **2021**, *21*, 5361. [CrossRef]
2. Pantanowitz, L.; Dickinson, K.; Evans, A.J.; Hassell, L.A.; Henricks, W.H.; Lennerz, J.K.; Lowe, A.; Parwani, A.V.; Riben, M.; Smith, C.D.; et al. American Telemedicine Association clinical guidelines for telepathology. *J. Pathol. Inform.* **2014**, *5*, 39. [CrossRef]
3. Helin, H.; Tolonen, T.; Ylinen, O.; Tolonen, P.; Napankangas, J.; Isola, J. Optimized JPEG 2000 compression for efficient storage of histopathological whole-slide images. *J. Pathol. Inform.* **2018**, *9*, 20. [CrossRef]
4. Kim, I.; Kang, K.; Song, Y.; Kim, T.J. Application of Artificial Intelligence in Pathology: Trends and Challenges. *Diagnostics* **2022**, *12*, 2794. [CrossRef]
5. Van der Laak, J.; Litjens, G.; Ciompi, F. Deep learning in histopathology: The path to the clinic. *Nat. Med.* **2021**, *27*, 775–784. [CrossRef]
6. Saba, T. Journal of Infection and Public Health. 13 Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons, and challenges. *J. Infect. Public Health* **2020**, *13*, 1274–1289. [CrossRef]
7. Lansdowne, L.E. Tumor Biology. Cancer Research from Technology Networks. Accessed August 16, 2022. Available online: <http://www.technologynetworks.com/cancer-research/infographics/tumor-biology-359548> (accessed on 1 October 2021).
8. Azhari, E.E.M.; Hatta, M.M.M.; Htike, Z.Z.; Win, S.L. Tumor detection in medical imaging: A survey. *Int. J. Adv. Inf. Technol.* **2014**, *4*, 21.
9. Díaz-Pernas, F.J.; Martínez-Zarzuela, M.; Antón-Rodríguez, M.; González-Ortega, D. A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network. *Healthcare* **2021**, *9*, 153. [CrossRef]
10. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer J. Clin.* **2021**, *71*, 209–249. [CrossRef]
11. Litjens, G.; Sánchez, C.I.; Timofeeva, N.; Hermsen, M.; Nagtegaal, I.; Kovacs, I.; Hulsbergen-van de Kaa, C.; Bult, P.; Van Ginneken, B.; Van Der Laak, J. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci. Rep.* **2016**, *6*, 26286. [CrossRef]
12. He, L.; Long, L.R.; Antani, S.; Thoma, G.R. Histology image analysis for carcinoma detection and grading. *Comput. Methods Programs Biomed.* **2012**, *107*, 538–556. [CrossRef] [PubMed]
13. Pandian, A.P. Identification and classification of cancer cells using capsule network with pathological images. *J. Artif. Intell.* **2019**, *1*, 37–44.
14. Ehteshami, B. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *J. Am. Med. Assoc.* **2017**, *318*, 2199–2210. [CrossRef] [PubMed]
15. Maseer, Z.K.; Yusof, R.; Bahaman, N.; Mostafa, S.A.; Foozy, C.F.M. Benchmarking of machine learning for anomaly based intrusion detection systems in the CICIDS2017 dataset. *IEEE Access* **2021**, *9*, 22351–22370. [CrossRef]
16. Senkamalavalli, R.; Bhuvaneswari, T. Improved classification of breast cancer data using hybrid techniques. *Int. J. Adv. Eng. Res. Sci.* **2017**, *5*, 237467. [CrossRef]
17. Jiang, P.; Peng, J.; Zhang, G.; Cheng, E.; Megalooikonomou, V.; Ling, H. Learning-based automatic breast tumor detection and segmentation in ultrasound images. In Proceedings of the 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), Barcelona, Spain, 2–5 May 2012; pp. 1587–1590.
18. Parmar, S.; Gondaliya, N. A Survey on Detection and Classification of Brain Tumor from MRI Brain Images using Image Processing Techniques. *Int. Res. J. Eng. Technol. IRJET* **2018**, *5*, 162–166.
19. Jayade, S.; Ingole, D.T.; Ingole, M.D. MRI brain tumor classification using hybrid classifier. In Proceedings of the 2019 International Conference on Innovative Trends and Advances in Engineering and Technology (ICITAET), Shegoan, India, 27–28 December 2019; pp. 201–205.
20. Peter, R.; Korfiatis, P.; Blezek, D.; Oscar Beitia, A.; Stepan-Buksakowska, I.; Horinek, D.; Flemming, K.D.; Erickson, B.J. A quantitative symmetry-based analysis of hyperacute ischemic stroke lesions in noncontrast computed tomography. *Med. Phys.* **2017**, *44*, 192–199. [CrossRef] [PubMed]
21. Seetha, J.; Raja, S.S. Brain tumor classification using convolutional neural networks. *Biomed. Pharmacol. J.* **2018**, *11*, 1457. [CrossRef]
22. Yu, J.; Li, Q.; Zhang, H.; Meng, Y.; Liu, Y.F.; Jiang, H.; Ma, C.; Liu, F.; Fang, X.; Li, J.; et al. Contrast-enhanced computed tomography radiomics and multilayer perceptron network classifier: An approach for predicting CD20+ B cells in patients with pancreatic ductal adenocarcinoma. *Abdom. Radiol.* **2022**, *47*, 242–253. [CrossRef]
23. Erickson, B.J.; Korfiatis, P.; Akkus, Z.; Kline, T.L. Machine learning for medical imaging. *Radiographics* **2017**, *37*, 505. [CrossRef]
24. Khajuria, R.; Quyoom, A.; Sarwar, A. A comparison of deep reinforcement learning and deep learning for complex image analysis. *J. Multimed. Inf. Syst.* **2020**, *7*, 1–10. [CrossRef]
25. Tyagi, A.K.; Chahal, P. Artificial intelligence and machine learning algorithms. In *Research Anthology on Machine Learning Techniques, Methods, and Applications*; IGI Global: Hershey, PA, USA, 2022; pp. 421–446.

26. Sarker, I.H. Machine learning: Algorithms, real-world applications and research directions. *SN Comput. Sci.* **2021**, *2*, 160. [[CrossRef](#)] [[PubMed](#)]
27. Montavon, G.; Samek, W.; Müller, K.R. Methods for interpreting and understanding deep neural networks. *Digit. Signal Process.* **2018**, *73*, 1–15. [[CrossRef](#)]
28. Rossi, F.; Conan-Guez, B. Functional Multi-Layer Perceptron: A Nonlinear Tool for Functional Data Analysis. *Neural Netw. Off. J. Int. Neural Netw. Soc.* **2005**, *18*, 45–60. [[CrossRef](#)]
29. Wahab, N.; Khan, A. Multifaceted fused-CNN based scoring of breast cancer whole-slide histopathology images. *Appl. Soft Comput.* **2020**, *97*, 106808. [[CrossRef](#)]
30. Zainudin, Z.; Shamsuddin, S.M.; Hasan, S. Deep layer CNN architecture for breast cancer histopathology image detection. In *International Conference on Advanced Machine Learning Technologies and Applications*; Springer: Cham, Switzerland, 2019; pp. 43–51.
31. Dabeer, S.; Khan, M.M.; Islam, S. Cancer diagnosis in histopathological image: CNN based approach. *Inform. Med. Unlocked* **2019**, *16*, 100231. [[CrossRef](#)]
32. Graziani, M.; Lompech, T.; Müller, H.; Andrearczyk, V. Evaluation and comparison of CNN visual explanations for histopathology. In Proceedings of the 35th AAAI Conference on Artificial Intelligence, Virtual Event, 2–9 February 2021; pp. 195–201.
33. Zhang, L.; Wu, Y.; Zheng, B.; Su, L.; Chen, Y.; Ma, S.; Hu, Q.; Zou, X.; Yao, L.; Yang, Y.; et al. Rapid histology of laryngeal squamous cell carcinoma with deep-learning based stimulated Raman scattering microscopy. *Theranostics* **2019**, *9*, 2541. [[CrossRef](#)]
34. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
35. Lauande, M.G.M.; Teles, A.M.; da Silva, L.L.; Matos, C.E.F.; Junior, G.B.; de Paiva, A.C.; de Almeida, J.S.; Oliveira, R.; Brito, H.; Nascimento, A.; et al. Classification of Histopathological Images of Penile Cancer using DenseNet and Transfer Learning. In Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications—Volume 4: VISAPP, Online, 6–8 February 2022; pp. 976–983.
36. Reenadevi, R.; Sathiya, T.; Sathiyabhama, B. Breast cancer histopathological image classification using augmentation based on optimized deep ResNet-152 structure. *Ann. Rom. Soc. Cell Biol.* **2021**, *25*, 5866–5874.
37. Wang, C.-W.; Lee, Y.-C.; Chang, C.-C.; Lin, Y.-J.; Liou, Y.-A.; Hsu, P.-C.; Chang, C.-C.; Sai, A.-K.; Wang, C.-H.; Chao, T.-K. A weakly supervised deep learning method for guiding ovarian cancer treatment and identifying an effective biomarker. *Cancers* **2022**, *14*, 1651. [[CrossRef](#)]
38. Samee, N.A.; Atteia, G.; Meshoul, S.; Al-antari, M.A.; Kadah, Y.M. Deep Learning Cascaded Feature Selection Framework for Breast Cancer Classification: Hybrid CNN with Univariate-Based Approach. *Mathematics* **2022**, *10*, 3631. [[CrossRef](#)]
39. Allegra, A.; Tonacci, A.; Sciacotta, R.; Genovese, S.; Musolino, C.; Pioggia, G.; Gangemi, S. Machine learning and deep learning applications in multiple myeloma diagnosis, prognosis, and treatment selection. *Cancers* **2022**, *14*, 606. [[CrossRef](#)] [[PubMed](#)]
40. Xie, X.; Fu, C.C.; Lv, L.; Ye, Q.; Yu, Y.; Fang, Q.; Zhang, L.; Hou, L.; Wu, C. Deep convolutional neural network-based classification of cancer cells on cytological pleural effusion images. *Mod. Pathol.* **2022**, *35*, 609–614. [[CrossRef](#)] [[PubMed](#)]
41. Gadermayr, M.; Koller, L.; Tschuchnig, M.; Stangassinger, L.M.; Kreutzer, C.; Couillard-Despres, S.; Oostingh, J.G.; Hittmair, A. MixUp-MIL: Novel Data Augmentation for Multiple Instance Learning and a Study on Thyroid Cancer Diagnosis. *arXiv* **2022**, arXiv:2211.05862.
42. Jimenez-del-Toro, O.; Otálora, S.; Andersson, M.; Eurén, K.; Hedlund, M.; Rousson, M.; Müller, H.; Atzori, M. Analysis of histopathology images: From traditional machine learning. In *Biomedical Texture Analysis*; Academic Press: Cambridge, MA, USA, 2017; pp. 281–314.
43. Komura, D.; Ishikawa, S. Machine learning methods for histopathological image analysis. *Comput. Struct. Biotechnol. J.* **2018**, *16*, 34–42. [[CrossRef](#)] [[PubMed](#)]
44. Lee, J.; Warner, E.; Shaikhouni, S.; Bitzer, M.; Kretzler, M.; Gipson, D.; Pennathur, S.; Bellovich, K.; Bhat, Z.; Gadegbeku, C.; et al. Unsupervised machine learning for identifying important visual features through bag-of-words using histopathology data from chronic kidney disease. *Sci. Rep.* **2022**, *12*, 4832. [[CrossRef](#)] [[PubMed](#)]
45. Ahamed, S.B.; Shamsudheen, S.; Balobaid, A.S. Deep Learning Approaches, Algorithms, and Applications in Bioinformatics. In *Applications of Machine Learning and Deep Learning on Biological Data*; Taylor & Francis: Abingdon, UK, 2023; p. 1.
46. Bilal, M.; Nimir, M.; Snead, D.; Taylor, G.S.; Rajpoot, N. Role of AI and digital pathology for colorectal immuno-oncology. *Br. J. Cancer* **2022**, *128*, 3–11. [[CrossRef](#)] [[PubMed](#)]
47. Banerji, S.; Mitra, S. Deep learning in histopathology: A review. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2022**, *12*, e1439. [[CrossRef](#)]
48. Lu, Z.; Xu, S.; Shao, W.; Wu, Y.; Zhang, J.; Han, Z.; Feng, Q.; Huang, K. Deep-learning-based characterization of tumor-infiltrating lymphocytes in breast cancers from histopathology images and multiomics data. *JCO Clin. Cancer Inform.* **2020**, *4*, 480–490. [[CrossRef](#)]
49. Subramanian, H.; Subramanian, S. Improving diagnosis through digital pathology: Proof-of-concept implementation using smart contracts and decentralized file storage. *J. Med. Internet Res.* **2022**, *24*, e34207. [[CrossRef](#)]
50. Han, Z.; Wei, B.; Zheng, Y.; Yin, Y.; Li, K.; Li, S. Breast cancer multi-classification from histopathological images with structured deep learning model. *Sci Rep.* **2017**, *7*, 4172. [[CrossRef](#)]
51. Ahmedt-Aristizabal, D.; Armin, M.A.; Denman, S.; Fookes, C.; Petersson, L. A survey on graph-based deep learning for computational histopathology. *Comput. Med. Imaging Graph.* **2022**, *95*, 102027. [[CrossRef](#)] [[PubMed](#)]

52. Bouteldja, N.; Klinkhammer, B.M.; Bülow, R.D.; Droste, P.; Otten, S.W.; von Stillfried, S.F.; Moellmann, J.; Sheehan, S.M.; Korstanje, R.; Menzel, S.; et al. Deep learning-based segmentation and quantification in experimental kidney histopathology. *J. Am. Soc. Nephrol.* **2021**, *32*, 52–68. [[CrossRef](#)] [[PubMed](#)]
53. Chen, M.; Zhang, B.; Topatana, W.; Cao, J.; Zhu, H.; Juengpanich, S.; Mao, Q.; Yu, H.; Cai, X. Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. *NPJ Precis. Oncol.* **2020**, *4*, 14. [[PubMed](#)]
54. Schaer, R.; Otálora, S.; Jimenez-del-Toro, O.; Atzori, M.; Müller, H. Deep learning-based retrieval system for gigapixel histopathology cases and the open access literature. *J. Pathol. Inform.* **2019**, *10*, 19. [[CrossRef](#)]
55. Chaunzwa, T.L.; Hosny, A.; Xu, Y.; Shafer, A.; Diao, N.; Lanuti, M.; Christiani, D.C.; Mak, R.H.; Aerts, H.J. Deep learning classification of lung cancer histology using CT images. *Sci. Rep.* **2021**, *11*, 5471. [[CrossRef](#)]
56. Mahesh, B. Machine learning algorithms-a review. *Int. J. Sci. Res. IJSR* **2020**, *9*, 381–386.
57. Sari, C.T.; Gunduz-Demir, C. Unsupervised feature extraction via deep learning for histopathological classification of colon tissue images. *IEEE Trans. Med. Imaging* **2018**, *38*, 1139–1149. [[CrossRef](#)]
58. Huss, R.; Coupland, S.E. Software-assisted decision support in digital histopathology. *J. Pathol.* **2020**, *250*, 685–692. [[CrossRef](#)]
59. Ciga, O.; Xu, T.; Martel, A.L. Self supervised contrastive learning for digital histopathology. *Mach. Learn. Appl.* **2022**, *7*, 100198. [[CrossRef](#)]
60. Byra, M.; Jarosik, P.; Szubert, A.; Galperin, M.; Ojeda-Fournier, H.; Olson, L.; O'Boyle, M.; Comstock, C.; Andre, M. Breast mass segmentation in ultrasound with selective kernel U-Net convolutional neural network. *Biomed. Signal Process. Control.* **2020**, *61*, 102027. [[CrossRef](#)]
61. Jiang, Y.; Sui, X.; Ding, Y.; Xiao, W.; Zheng, Y.; Zhang, Y. A semi-supervised learning approach with consistency regularization for tumor histopathological images analysis. *Front. Oncol.* **2022**, *12*, 7200. [[CrossRef](#)]
62. Nugroho, A.; Suhartanto, H. Hyper-parameter tuning based on random search for densenet optimization. In Proceedings of the 2020 7th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), Semarang, Indonesia, 24–25 September 2020; pp. 96–99.
63. Khanam, J.J.; Foo, S.Y. A comparison of machine learning algorithms for diabetes prediction. *ICT Express* **2021**, *7*, 432–439. [[CrossRef](#)]
64. Li, Y.; Wei, C.; Ma, T. Towards explaining the regularization effect of initial large learning rate in training neural networks. In Proceedings of the NIPS'19: Proceedings of the 33rd International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019.
65. Chikontwe, P.; Sung, H.J.; Jeong, J.; Kim, M.; Go, H.; Nam, S.J.; Park, S.H. Weakly supervised segmentation on neural compressed histopathology with self-equivariant regularization. *Med. Image Anal.* **2022**, *80*, 102482. [[CrossRef](#)]
66. Buddhavarapu, V.G. An experimental study on classification of thyroid histopathology images using transfer learning. *Pattern Recognit. Lett.* **2020**, *140*, 1–9. [[CrossRef](#)]
67. Kumar, S.K. On weight initialization in deep neural networks. *arXiv* **2017**, arXiv:1704.08863.
68. Zhang, H.; Feng, L.; Zhang, X.; Yang, Y.; Li, J. Necessary conditions for convergence of CNNs and initialization of convolution kernels. *Digit. Signal Process.* **2022**, *123*, 103397. [[CrossRef](#)]
69. Singh, P.; Muchahari, M.K. Solving multi-objective optimization problem of convolutional neural network using fast forward quantum optimization algorithm: Application in digital image classification. *Adv. Eng. Softw.* **2023**, *176*, 103370. [[CrossRef](#)]
70. Srikantamurthy, M.M.; Rallabandi, V.P.; Dudekula, D.B.; Natarajan, S.; Park, J. Classification of benign and malignant subtypes of breast cancer histopathology imaging using hybrid CNN-LSTM based transfer learning. *BMC Med. Imaging* **2023**, *23*, 19. [[CrossRef](#)]
71. Gour, M.; Jain, S.; Shankar, U. Application of Deep Learning Techniques for Prostate Cancer Grading Using Histopathological Images. In *Computer Vision and Image Processing: 6th International Conference, CVIP 2021, Rupnagar, India, December 3–5, 2021, Revised Selected Papers, Part I*; Springer International Publishing: Cham, Switzerland, 2022; pp. 83–94.
72. Krishna, S.; Krishnamoorthy, S.; Bhavsar, A. Stain normalized breast histopathology image recognition using convolutional neural networks for cancer detection. *arXiv* **2022**, arXiv:2201.00957.
73. Uthatham, A.; Yodrabum, N.; Sinmaroeng, C.; Titijaroonroj, T. Automatic Lymph Node Classification with Convolutional Neural Network. In Proceedings of the 2022 14th International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, Indonesia, 18–19 October 2022; pp. 1–6.
74. Li, X.; Cen, M.; Xu, J.; Zhang, H.; Xu, X.S. Improving feature extraction from histopathological images through a fine-tuning ImageNet model. *J. Pathol. Inform.* **2022**, *13*, 100115. [[CrossRef](#)]
75. Wu, C.T.; Lin, P.H.; Huang, S.Y.; Tseng, Y.J.; Chang, H.T.; Li, S.Y.; Yen, H.W. Revisiting alloy design of low-modulus biomedical β -Ti alloys using an artificial neural network. *Materialia* **2022**, *21*, 101313. [[CrossRef](#)]
76. Wright, L.; Demeure, N. Ranger21: A synergistic deep learning optimizer. *arXiv* **2021**, arXiv:2106.13731.
77. Wang, P.; Li, P.; Li, Y.; Wang, J.; Xu, J. Histopathological image classification based on cross-domain deep transferred feature fusion. *Biomed. Signal Process. Control.* **2021**, *68*, 102705. [[CrossRef](#)]
78. Akossi, A.; Wang, F.; Teodoro, G.; Kong, J. Image registration with optimal regularization parameter selection by learned auto encoder features. In Proceedings of the 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Nice, France, 13–16 April 2021; pp. 702–705.

79. Moshkov, N.; Mathe, B.; Kertesz-Farkas, A.; Hollandi, R.; Horvath, P. Test-time augmentation for deep learning-based cell segmentation on microscopy images. *Sci. Rep.* **2020**, *10*, 5068. [[CrossRef](#)] [[PubMed](#)]
80. Nguyen, C.P.; Vo, A.H.; Nguyen, B.T. Breast cancer histology image classification using deep learning. In Proceedings of the 2019 19th International Symposium on Communications and Information Technologies (ISCIT), Ho Chi Minh City, Vietnam, 25–27 September 2019; pp. 366–370.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.