9-24-2021

# Computer-aided diagnosis of low grade endometrial stromal sarcoma (LGESS)

Xinxin Yang
*San Jose State University*

Mark Stamp
*San Jose State University*, mark.stamp@sjsu.edu

## Recommended Citation

# Computer-aided diagnosis of low grade endometrial stromal sarcoma (LGESS)

Xinxin Yang, Mark Stamp [*]

*Department of Computer Science, San Jose State University, USA*

## ARTICLE INFO

## ABSTRACT

Low grade endometrial stromal sarcoma (LGESS) accounts for about 0.2% of all uterine cancer cases. Approximately 75% of LGESS patients are initially misdiagnosed with leiomyoma, which is a type of benign tumor, also known as fibroids. In this research, uterine tissue biopsy images of potential LGESS patients are preprocessed using segmentation and stain normalization algorithms. We then apply a variety of classic machine learning and advanced deep learning models to classify tissue images as either benign or cancerous. For the classic techniques considered, the highest classification accuracy we attain is about 0.85, while our best deep learning model achieves an accuracy of approximately 0.87. These results clearly indicate that properly trained learning algorithms can aid in the diagnosis of LGESS.

## 1. Introduction

Cancer is the second leading cause of death in the United States, accounting for 21.6% of all deaths in a 2017 survey conducted by the Center for Disease Control (CDC) [19]. The tremendous medical costs of cancer treatments and the harm the disease brings to patients and their families makes cancer a vital area of medical research.

Low grade endometrial stromal sarcoma (LGESS) is a tumor comprised of endometrial stromal cells. It is rare, accounting for approximately 0.2% of uterine cancers [7,20]. Most patients with LGESS have a good prognosis, with a 5-year survival rate of about 80% after surgical removal of the tumor. However, LGESS has a relatively high recurrence rates, at about 60%, and the disease-related death rate is estimated to be between 15% and 25% [15,32].

When diagnosing LGESS, it is difficult to differentiate LGESS from benign leiomyoma, which is also known as fibroids. Consequently, only about 10% of LGESS patients are correctly diagnosed, whereas 75% are misdiagnosed with preoperative leiomyoma [22]. Many cases even remain misdiagnosed postoperative [11]. More accurate and automatic image analysis methods would be useful to diagnose LGESS, assess treatment efficacy, and lower cancer-related costs to the healthcare system. Due to the inherent difficulty of the problem, learning algorithms for LGESS are sure to be less than 100% accurate. This makes such techniques more suited as an aid in cancer risk assessment, rather than as definitive diagnostic tools [25]. Learning models can reduce the workload of healthcare professionals by automating tedious tasks, such as tumor segmentation. Moreover, learning algorithms may be more capable than humans at analyzing smaller, more subtle structures in patient images [1]. Also, computers can analyze larger feature sets in a shorter amount of time, potentially allowing for a more nuanced analysis of image structure that is not easily perceived by a human viewer. These capabilities have been showcased in a wide variety of use cases, including tumor segmentation, determination of tumor malignancy, and prediction of survivability in afflicted patients [27].

In this research, we apply machine learning and deep learning methods to classify soft tissue images of potential LGESS patients. Our goal is to provide tools for improved diagnostic accuracy of LGESS tumors. In fact, our results confirm that machine learning techniques can be used to greatly improve the diagnostic accuracy of LGESS tumors, as compared to current practice.

The remainder of this paper is organized as follows: Section 2 contains relevant background material, including a review of the research literature related to cancer image analysis, and a brief introduction to the machine learning and deep learning algorithms that we employ in our experiments. Specifically, we consider the classic machine learning techniques of multilayer perceptron (MLP), random forest, XGBoost, support vector machines (SVM), and principal component analysis (PCA) in combination with SVM. In the realm of deep learning algorithms, we consider convolutional neural networks (CNN), residual networks (ResNet), AlexNet, and DenseNet.

---

* Corresponding author.
  *E-mail address:* mark.stamp@sjsu.edu (M. Stamp).

Section 3 gives an overview of the cancer image dataset used in this research, and we discuss the preprocessing strategies that we employ. This data preprocessing workflow, includes region of interest (ROI) segmentation, image patch extraction, and stain normalization. In Section 4, we provide details on our machine learning and deep learning experiments, and we analyze the results. Section 5 concludes the paper, and we provide suggestions for possible future work related to automatic analysis of LGESS images.

## 2. Background

In this section, we discuss relevant background topics. First, we discuss related work, then we provide a brief overview of each of the classic machine learning and deep learning algorithms that we use for the analysis of LGESS images.

### 2.1. Related work

Machine learning has found widespread use in cancer classification and diagnosis [17,27]. Although machine learning and deep learning research has been conducted on many different types of cancer, we are aware of no existing studies pertaining to LGESS. This may be due to the rarity of LGESS, as compared to other cancers that are typically studied in this type of research.

Mesrabadi and Faez [29] apply multilayer perceptrons (MLP), AlexNet, and support vector machines (SVM) to classify potential prostate cancer images. In their experiments, AlexNet yielded a classification accuracy of 86.3%, compared to 81.1% for SVM and 79.3% for MLPs. Kharya et al. [23] note that generic MLPs are the most widely used prediction technique in medical forecasting. Kharya et al. also consider the relative advantages and disadvantages of decision trees, naïve Bayes, MLPs and SVMs for breast cancer detection.

Ashhar et al. [4] consider the efficacy of deep learning algorithms for early detection of lung cancer. Lung cancer is almost always diagnosed at advanced stages—early and accurate screening would be beneficial to patient outcomes. Ashhar et al., apply five state-of-the-art CNN-based architectures to a dataset of computer tomography (CT) lung cancer images. They find that GoogleNet gives the best results, with an accuracy, specificity, sensitivity, and AUC of 94.53%, 99.06%, 65.67%, and 86.84%, respectively.

Vijayarajeswari et al. [36] apply Hough transforms to mammogram images to detect features that are potentially indicative of breast cancer, and the images are then classified using an SVM. They attain 94% accuracy with this strategy, far surpassing the classification accuracy of a SVM on unmodified images. The Wisconsin diagnostic breast cancer (WDBC) dataset was used as the source of their images [33].

In [16], Ghoneim et al. conduct research on cervical cancer detection and classification—cervical cancer is one of the leading causes of cancer death among women. They extract relevant features from cervical cancer images using CNNs, then classify the images using extreme learning machines (ELM), multi-layer perceptrons (MLP), and autoencoder (AE) based classifiers. The best performance in these experiments came from a CNN-ELM-based system, which attained an accuracy of 99.5% on a binary classification problem and 91.2% accuracy on a multiclass problem.

Chaturvedi et al. [12] propose a classification method for skin cancer that achieves better evaluation results than previous studies and outperforms human dermatologists. Their implementation of the MobileNet model achieved an overall accuracy of 83.1% for multiclass experiments involving seven classes.

Bharat et al. [6] apply traditional machine learning classifiers, including $k$-nearest neighbor ($k$-NN), naïve Bayes, classification and regression trees (CART), and SVMs to the problem of predicting and diagnosing breast cancer. Interestingly, they conclude that these machine learning algorithms behave quite differently depending on the application. Specifically, they find that $k$-NN has the best diagnostic

results, while naïve Bayes and logistic regression perform well when applied to specific breast cancer diagnosis problems. These results generally agree with the research of Rana et al. [31].

According to Maglogiannis et al. [26], SVM is the best technique for predicting recurrence or non-recurrence of breast cancer. In their research, SVM attains an accuracy of 96.91%, specificity of 97.67%, and sensitivity of 97.84%, outperforming the Bayesian and MLP classifiers it was compared against.

The results of these studies paint a promising picture for the predictive abilities of machine learning and deep learning in the cancer domain. In this research, we apply deep learning and machine learning techniques to the diagnosis of LGESS, based on images.

### 2.2. Learning algorithms

Based on our literature review, we will consider both classic machine learning models and various cutting-edge deep learning techniques. In this section, we provide a brief description of each learning model that we employ in our classification experiments.

#### 2.2.1. Multilayer perceptron
Multilayer perceptrons (MLP) are a basic type of artificial neural network (ANN). An MLP includes an input layer and an output layer, with one or more hidden layers in between. MLPs can form the basis of deep neural networks (DNN), where multiple hidden layers are used [37].

#### 2.2.2. Random forest
A random forest combines multiple decision trees, where each component decision tree is trained on a subset of the data and features [8]. Every decision tree is itself a classifier, and hence if *N* decision trees comprise a random forest, we will have *N* classification results. A random forest typically uses a simple voting scheme among the component decision trees to determine the classification.

#### 2.2.3. XGBoost
Extreme gradient boosting (XGBoost) [13] has attracted considerable attention because of its efficiency and high prediction accuracy, at least in some cases. XGBoost, which is based on gradient boosted decision trees (GBDT), has proven more effective than simpler boosting techniques, such as AdaBoost.

#### 2.2.4. SVM
Support vector machines (SVM) date to the early 1960s, but only became practical in the mid-1990s [5]. Prior to the recent re-emergence of deep learning, SVM was regarded as one of the most successful machine learning algorithms, and it is still competitive in many problem domains.

In a binary classification problem, an SVM attempts to construct a separating hyperplane between the classes, while maximizing the "margin," i.e., the minimum distance between the separating hyperplane and the training samples. The strength of SVMs derive largely from the so-called kernel trick, which enables operations to take place in a higher dimensional space—where there is a better chance of the data being linearly separable—without paying a significant performance penalty.

#### 2.2.5. PCA
Principal component analysis (PCA) is one of the most widely used dimensionality reduction algorithms [34]. PCA is based on eigenvalue analysis, with the underlying assumption being that linear combinations of features with larger variances are more relevant for classification than those with smaller variance.

#### 2.2.6. CNN
As the name suggests, convolutional neural networks (CNN) include

convolutional layers, which consist of filters that enable the efficient analysis of local structure [14]. Fig. 1 illustrates a case where five convolutions are applied to three-dimensional data.

By applying multiple convolutional layers, progressively higher levels of abstraction are obtained. CNNs were designed for image analysis, but the technique generally performs well in any task where local structure dominates.

### 2.2.7. AlexNet

AlexNet is a specific CNN architecture. The AlexNet architecture is illustrated in Fig. 2.

From Fig. 2, we see that AlexNet has eight layers, not including pooling layers, with the first five being convolutional layers, while the last three are fully connected. There are 650,000 neurons and some 60 million parameters in AlexNet. See Ref. [24] for additional details on the AlexNet architecture.

### 2.2.8. DenseNet

The dense convolutional network (DenseNet) architecture was proposed in 2017 [21]. DenseNet connects each layer in a feed-forward manner. In contrast to a typical convolutional neural network, in DenseNet, the input of each layer includes the output of all previous layers. Fig. 3 illustrates the DenseNet architecture.

### 2.2.9. ResNet

Residual networks (ResNet) were first proposed in 2015, and in that same year, a ResNet architecture won first place in the classification task for the ImageNet competition. ResNet can be viewed as a collection of sub-networks that can be stacked to form a deep network. The use of these sub-networks, or "shortcut connections," enables efficient training of very deep networks. Fig. 4 illustrates a sub-network structure of ResNet. The identity mapping in Fig. 4 represents a shortcut connection.

## 3. Dataset and preprocessing

The dataset used in our experiments was obtained from the Cancer Imaging Archive, an organization offering data for cancer research [10]. Our specific dataset includes 888 uterine tissue biopsy images taken from 250 potential LGESS patients, formatted as SVS files. A separate annotations file contains the clinical diagnosis of each tissue image.

SVS is a file format for whole slide images (WSI). In a WSI, a microscope slide is scanned to create a single high-resolution digital file [28]. Most WSIs have a resolution of 100, 000 × 100, 000 pixels. WSIs are usually stored in a pyramid structure, where each level of the pyramid holds different downsampled versions of the original image [9]. The more "downsampled" an image, the lower the resolution and the less magnified it appears to be. Fig. 5 illustrates the pyramid structure of a WSI image. In this research, we use the OpenSlide library [30], which provides methods to read and access WSI images stored in a variety of file formats, including SVS.

WSIs offer clear visualization of tumor characteristics, including tissue infiltration, lymph node metastasis, and degree of differentiation. Such images are helpful for the diagnosis, prognosis, grading, and staging of tumors [35].

The WSIs contained in the Cancer Imaging Archive dataset must undergo a color standardization stage before our classification algorithms can be applied, since different production processes and scanning machines result in color variations. These color differences can cause problems for algorithms that are not robust to such variations, even if the differences are imperceptible to the human eye.

The way we deal with color difference is by color normalization, which is also referred to as stain normalization. This involves normalizing all pictures to the color distribution of a template picture [3]. Stain normalization is a standard practice in the analysis of biopsy images using machine learning techniques. Without such normalization, insignificant color variations are likely to result in decreased classification accuracy.

Before applying color normalization to our WSIs, the regions of interest (ROI) of each image need to be identified. This is done via standard Gaussian filtering and contour extraction techniques. These processing steps are illustrated in Fig. 6.

The details of steps 1 through 3 of the process outlined in Fig. 6 are as follows.

1. Segmentation of the target region from the image.
   (a) Use the OpenSlide library to read the level 2 image of the WSI. This second level of the WSI pyramid still has good resolution, but contains less data and is thus easier and more efficient to
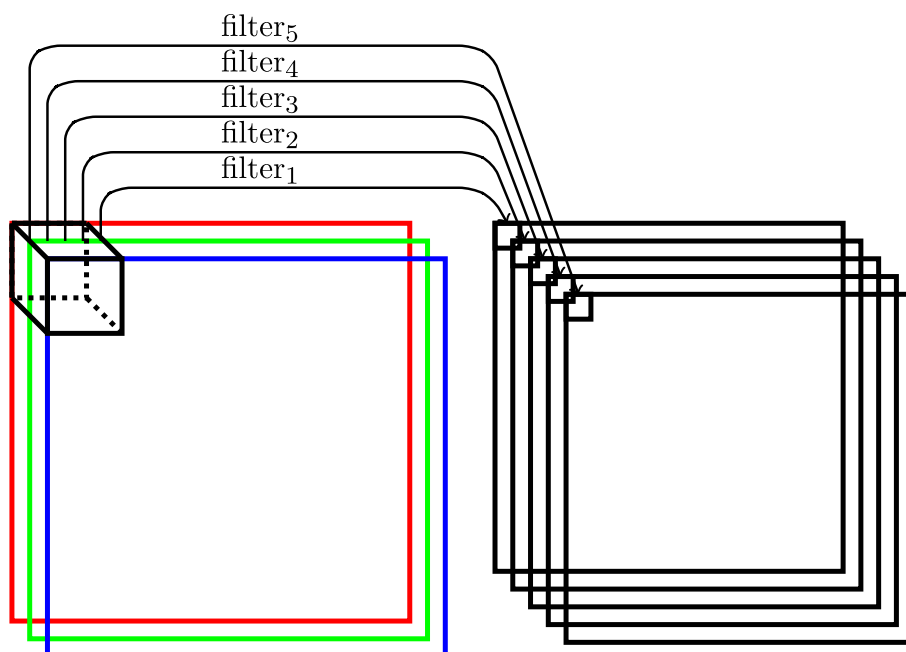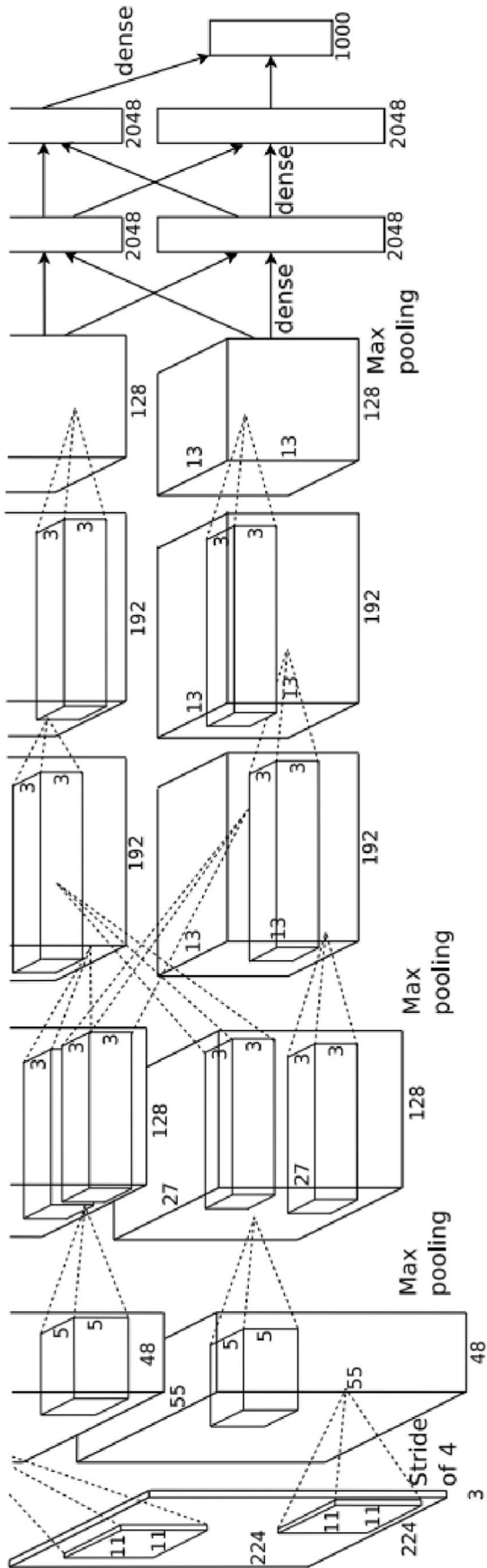


**Fig. 1.** Illustration of convolutions.
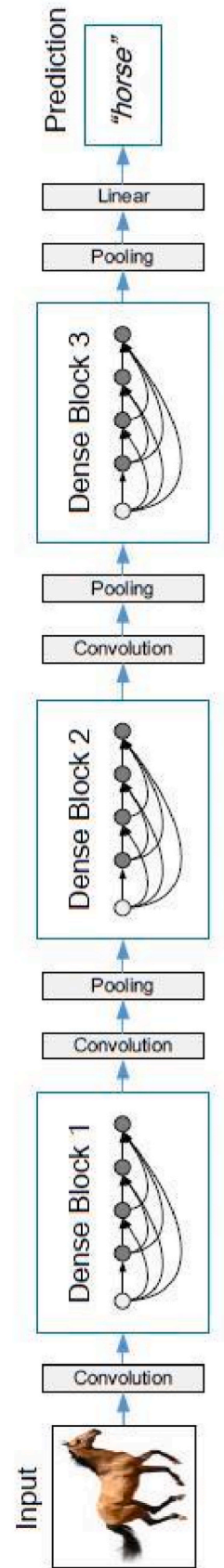
**Fig. 2.** AlexNet architecture [24].



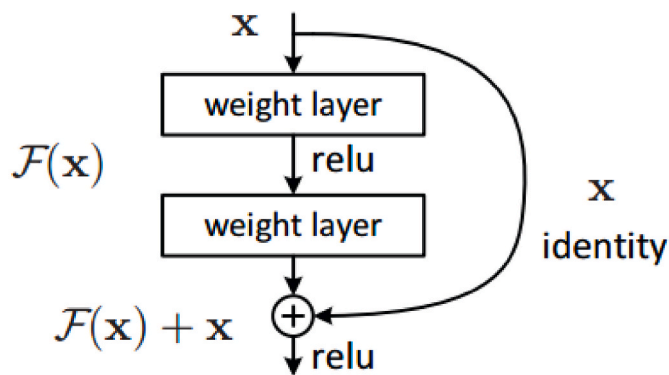**Fig. 3.** DenseNet architecture [21].
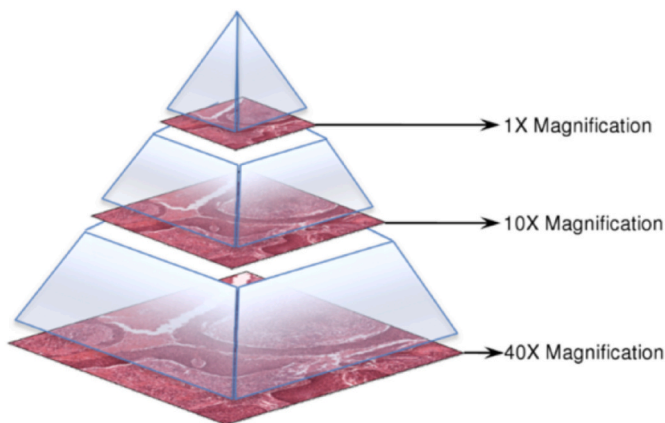
**Fig. 4.** ResNet sub-network [18].



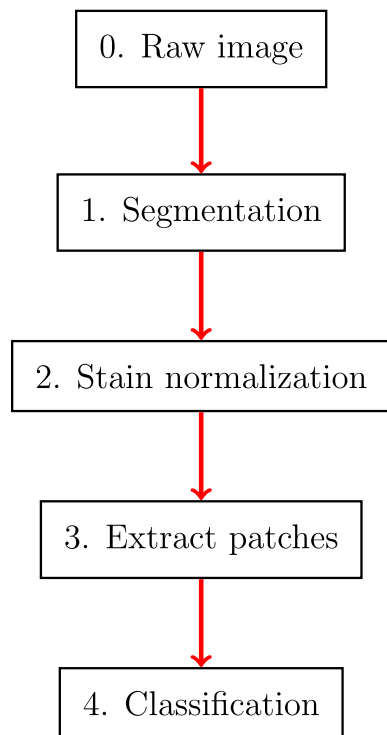**Fig. 5.** Whole slide image structure.



**Fig. 6.** Processing and classification of images.

process. We save the level 2 image in portable network graphics (PNG) format.

(b) Transform the image from the RGB color space to grayscale.

(c) Apply a Gaussian filter to normalize the image. We must make sure the filter threshold preserves the image contours.

(d) Calculate the area of each contour and remove all contours that have an area below a specified threshold.

(e) Obtain the final image mask based on the contours acquired from the previous steps. This image mask comprises the tissue regions.

Fig. 7 below shows examples of extracted contours and the resulting image masks. In each row, the first image is the original image in the dataset, the second image is the mask we obtain from the original image, and the last image is the original image with the contours marked.

2. Extract patches from an image's target region.

We apply the previously-generated mask to the image, then walk through the image and cut patches of a predefined length and width. All patches extracted must have a certain area occupied by the masked image. Patches below this area threshold consist mostly of empty space and are therefore discarded. Examples of extracted patches are given in Fig. 8.

3. Staining normalization.

We apply Vahadane's staining normalization to achieve color normalization [3]. The steps involved in this process are as follows:

(a) Optical density calculation.

(b) Unsupervised staining density estimation.

(c) Color normalization.

(d) Normalized pixel intensity calculation.

Examples of images before and after color normalization are given in Fig. 9.

Fig. 9 shows the color of an original image, labeled "Source," while the image labeled "Target" indicates the color we would like our image to become. The image labeled "Result" shows the original source image after color normalization. Note that the color of the "Result" image is similar to that of the "Target," as desired.

Fig. 10 shows 12 image examples from our dataset before and after color normalization. Note that in each case, the "before" image is directly above the corresponding "after" image.

Before normalization, we observe that the colors in the image can vary greatly. After the stain normalize process, the colors are clearly much more uniform.

## 4. Experiments and results

After preprocessing the images and cutting them into patches, our dataset includes 4205 tumor images and 1459 benign images. We use 5-fold cross validation, which serves to avoid any potential issues of bias in the data, and to maximize the amount of training data. For each fold of an experiment, this gives us 4529 samples in the training set (3363 tumor images and 1166 benign images) and 1133 samples in the test set (841 tumor images and 292 benign images).

The metrics we use to evaluate the performance of our machine learning classifiers are precision, recall, F1 score, and the accuracy. Precision is the proportion of true positives among those predicted as positive, while recall measures how accurately the positive samples are predicted. In cancer detection, it is imperative that models have a high recall. The F1 score is the harmonic mean of the precision and recall, while accuracy is self-explanatory.

For the basic techniques considered in the next section, we also consider receiver operating characteristic ROC analysis. An ROC curve is
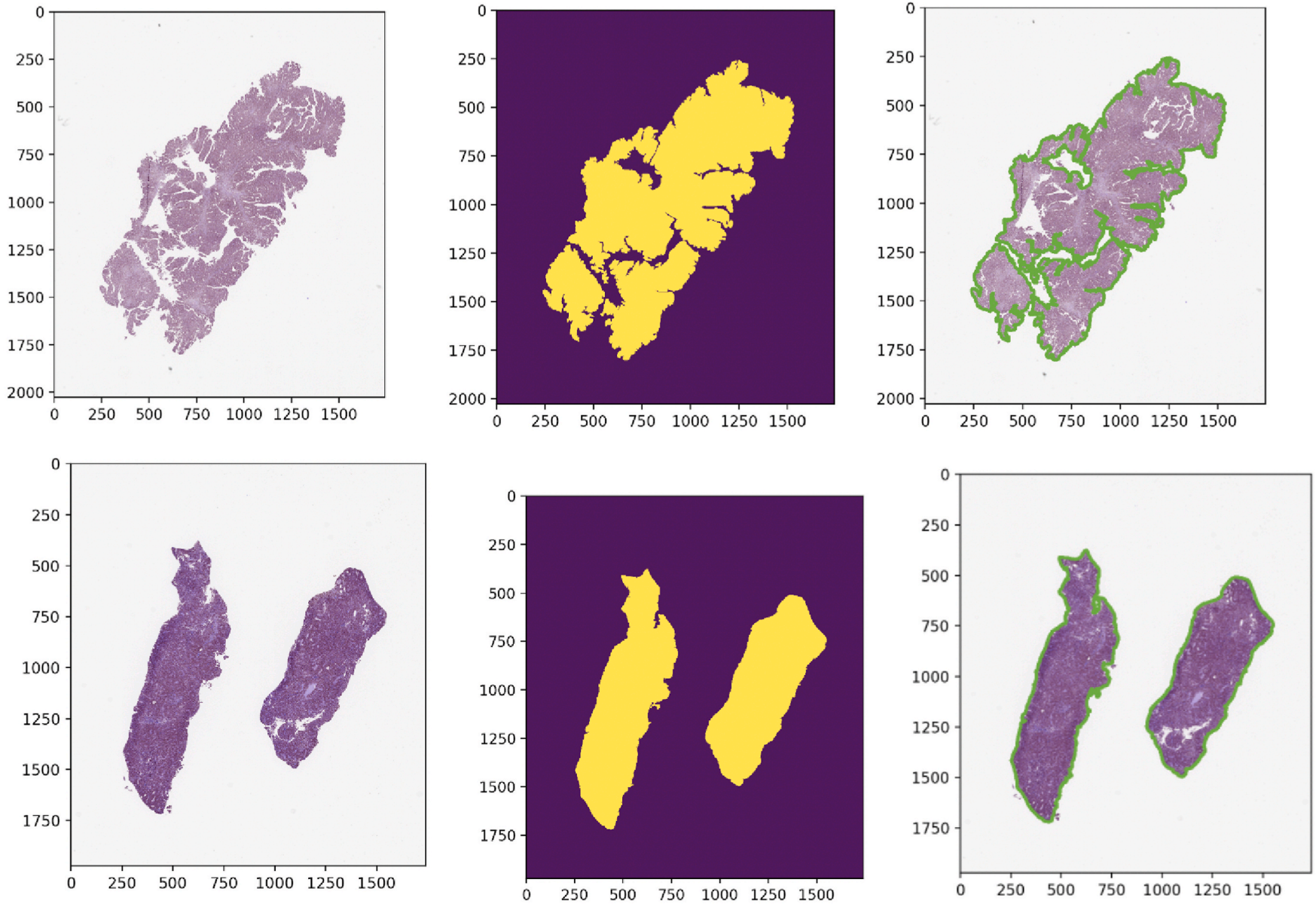
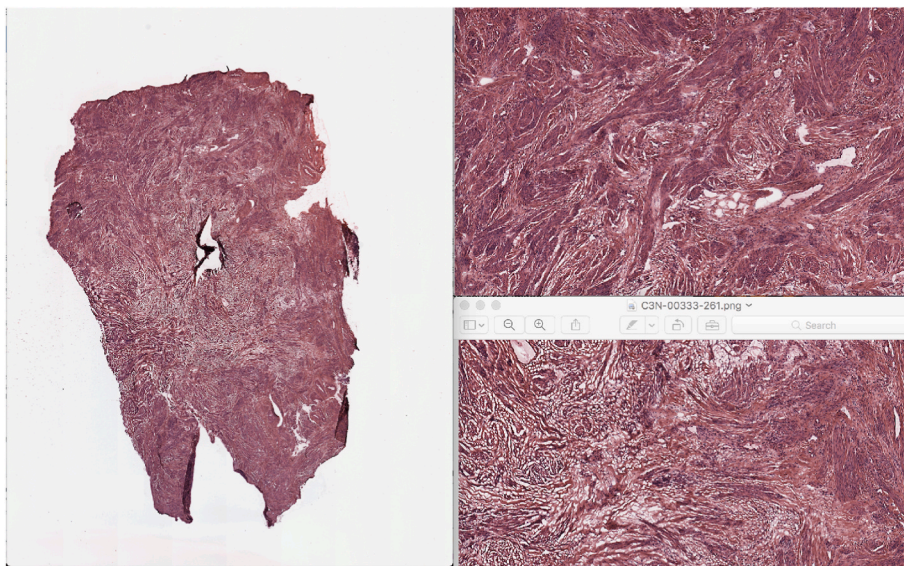**Fig. 7.** Examples of contour and mask extracted from raw image.

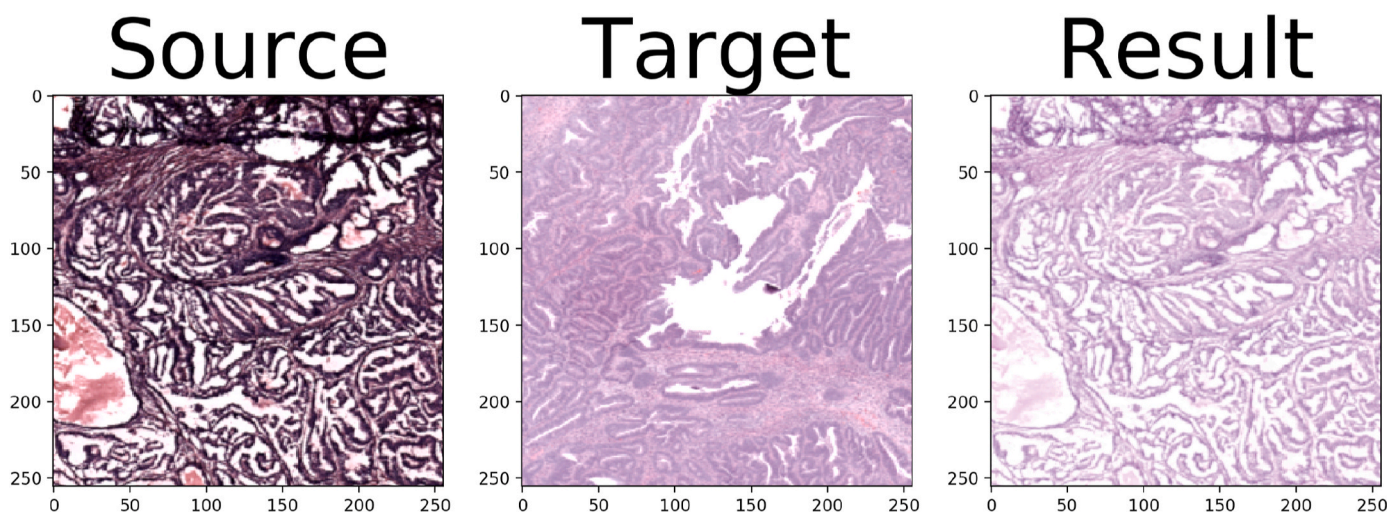**Fig. 8.** Patches cut from one image.



**Fig. 9.** Color normalization of an image.

a plot of the true positive rate versus the false positive rate as the threshold varies through all possible values. The area under the ROC curve (AUC) can be interpreted as the probability that a randomly selected positive instance scores higher than a randomly selected negative instance. The AUC provides a threshold-independent means of comparing classifiers. An AUC of 1.0 implies that there is a threshold for which ideal separation is attained, while an AUC of 0.5 means that the binary classifier is no better than a coin flip. If an AUC value of $0 \leq x < 0.5$ is obtained, by simply switching the sense of the classifier, we obtain an AUC of $1 - x > 0.5$, and hence the AUC can be no worse than 0.5.

### 4.1. Basic techniques

In this section, we will discuss our experiment result using "basic" or standard machine learning techniques. We include as basic MLP, random forest, XGBoost, and SVM models, as well as a model that combines PCA for dimensionality reduction with SVM for classification. Here, we discuss the specific architecture for each of these models and report the test accuracy. At the end of this section, we provide a more detailed comparison of these basic models.

#### 4.1.1. Multilayer perceptron

In this experiment, we imported MLPClassifier from the Scikit-Learn library. We use a fully connected neural network with three hidden layers–the first, second, and third hidden layers have 600, 800, and 300 neuron, respectively. We use rectified linear units (RELU) as the activation function for each neuron, and we train the model for 600 epochs. In this case, we obtain a test accuracy of 0.5702.

#### 4.1.2. Random forest

We build our random forest model using RandomForestClassifier from the module sklearn.ensemble. We have experimented with the number of trees and find that 100 gives the best accuracy. The maximum depth of the tree was set to none so that nodes are expanded until all leaves are pure or until all leaves contain less than elements. All other parameters of the model were set to default values. In this case, we obtain a test accuracy of 0.7873.

#### 4.1.3. XGBoost

Our XGBoost model is trained using the xgboost package in Python. We obtain the best results with max_depth set to 6, and the objective function binary:logistic. We achieve a test accuracy of 0.8147.

Fig. 10. Color normalization examples.

### 4.1.4. Support vector machine

We use the SVM model from the Scikit-Learn. We found no significant difference between any kernel functions tested, and hence for the sake of efficiency, we select the linear kernel. We set all other parameters to default except that we enable probability estimates. In our SVM experiment, we achieve a test accuracy of 0.8455.

### 4.1.5. PCA with SVM

In this experiment, we import the PCA model from the sklearn. decomposition module. We combine PCA with SVM using the make_pipeline method from the sklearn.pipeline module. We find that the best results are obtained with 300 components from PCA and using the RBF kernel for the SVM. In this model, we obtain a test accuracy of 0.8535.

### 4.1.6. Comparison of basic techniques

Table 1 provides a more detailed comparison of the basic techniques discussed above. The training and test accuracies are also given in the form of a bar graph in Fig. 11. We see that SVM and PCA with SVM achieve the best test accuracy, with both at approximately 0.85.

Fig. 12 gives ROC curves for each of the basic classifiers considered above. SVM performs best, marginally outperforming the "PCA with SVM" model.

### 4.2. Advanced techniques

In this section, we discuss the performance of more advanced learning techniques. Specifically, we consider CNN, AlexNet, DenseNet, ResNet, and a ResNet model that includes realtime data augmentation.

### 4.2.1. Convolutional neural network

Our CNN model is implemented in Tensorflow and includes two convolution layers, a learning rate of 0.005, a max_pool size of two, and a final fully connected layer. We experimented with various batch sizes, number of generation, and optimizers. The loss and accuracy plots for our best result (which used the Adam optimizer) are shown in Fig. 13. In this case, we achieve a training accuracy of 0.8143 test accuracy of 0.7860.

### 4.2.2. AlexNet

We use the original AlexNet architecture with a learning rate of 0.001, softmax as the loss function, and the Adam optimizer. We train the model for 50 epochs. The best training and test accuracy we attain with AlexNet on our LGESS image dataset are 0.9153 and 0.8314, respectively. Fig. 14 shows the loss and accuracy plots for this case.

### 4.2.3. DenseNet

We built our DenseNet based on the Keras library. Each convolution layer is followed by a max pooling layer. Three dense blocks are used, and there is a classification layer at the end. We found that a learning rate to 0.001 and 16 filters gave the best results. We trained the model for 30 epochs. This DenseNet, when applied to our LGESS dataset, yields a training accuracy of 0.8683 and a testing accuracy of 0.8457. Fig. 15
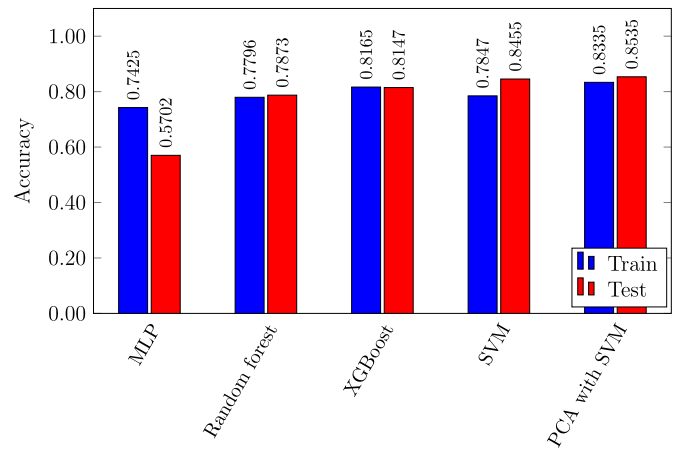


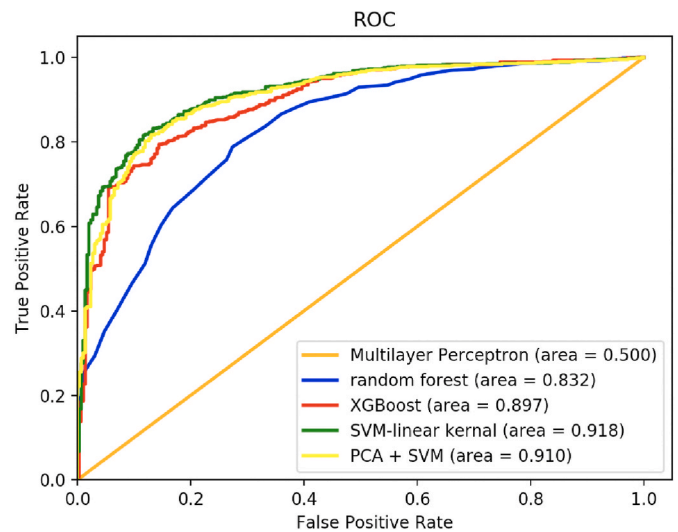**Fig. 11.** Accuracies for basic techniques.



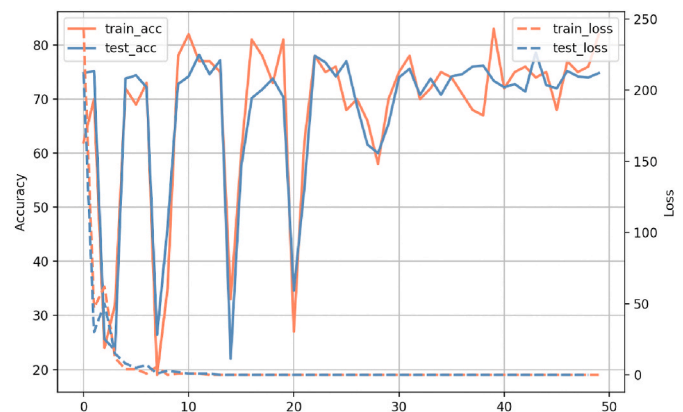**Fig. 12.** ROC analysis of basic classifiers.



**Fig. 13.** CNN accuracy and loss graphs.

details the performance of our DenseNet model.

### 4.2.4. ResNet

We used the Keras library to build our ResNet. This particular ResNet is built with a convolution layer, max pooling layer, basic block layers, and an average pooling layer. The basic block layer includes a

**Table 1**
Accuracy, precision, recall, and F1 score for basic techniques.

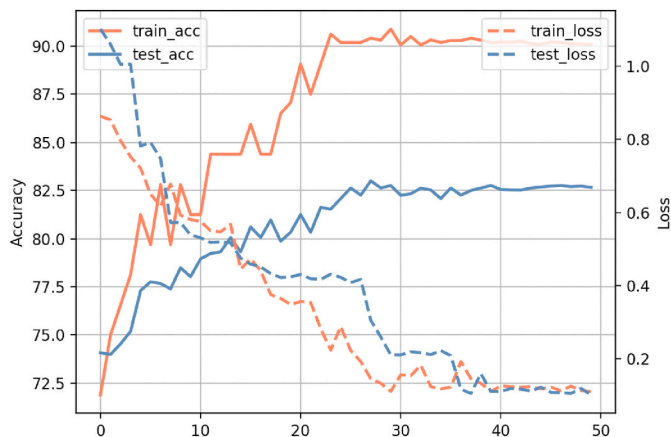| Technique | Mode | Precision | Recall | F1 | Accuracy |
|---|---|---|---|---|---|
| MLP | Train | 0.7425 | 1.0000 | 0.8523 | 0.7425 |
| | Test | 0.7493 | 0.6326 | 0.6860 | 0.5702 |
| Random forest | Train | 0.7800 | 0.9795 | 0.8684 | 0.7796 |
| | Test | 0.7820 | 0.9893 | 0.8735 | 0.7873 |
| XGBoost | Train | 0.8254 | 0.9548 | 0.8854 | 0.8165 |
| | Test | 0.8171 | 0.9667 | 0.8856 | 0.8147 |
| SVM | Train | 0.8510 | 0.8598 | 0.8554 | 0.7847 |
| | Test | 0.8698 | 0.9301 | 0.8989 | 0.8455 |
| PCA with SVM | Train | 0.9201 | 0.9675 | 0.9432 | 0.8335 |
| | Test | 0.8567 | 0.9632 | 0.9068 | 0.8535 |

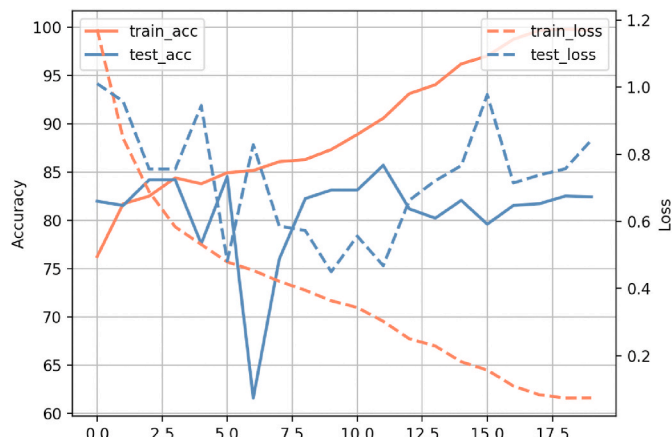**Fig. 14.** AlexNet accuracy and loss graphs.



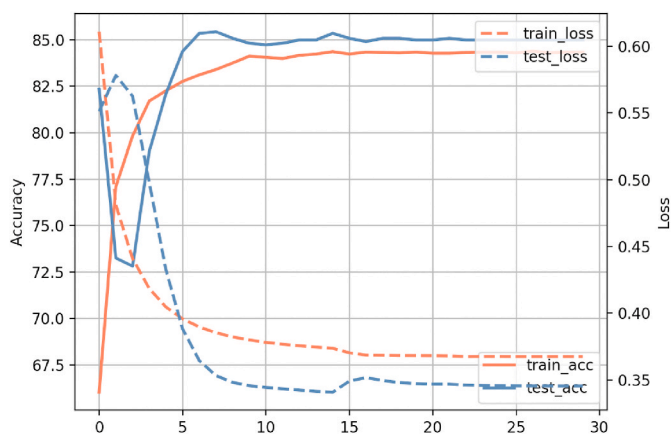**Fig. 16.** ResNet accuracy and loss graphs.



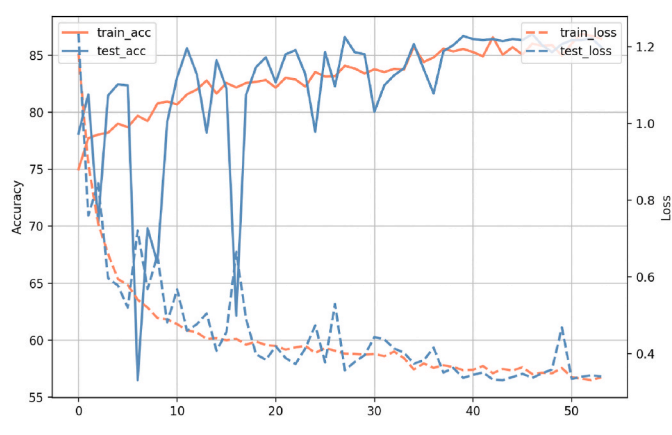**Fig. 15.** DenseNet accuracy and loss graphs.



**Fig. 17.** ResNet with data augmentation accuracy and loss graphs.

convolution layer, a batch normalization layer, and an activation layer. We trained our model for 20 epochs and obtained a best training accuracy of 0.9978 and testing accuracy of 0.8570. Fig. 16 shows the performance of this ResNet model on our dataset.

*4.2.5. ResNet with realtime data augmentation*

Since ResNet yielded the best performance, we further test it with realtime data augmentation. In general, data augmentation can reduce overfitting, as well improving classification accuracy [38]. However, there is no reliable way to know in advance whether data augmentation will yield improved results, so experimentation is required.

In our realtime data augmentation, images in the existing dataset are copied, randomly rotated, shifted, or flipped to form an augmented dataset. ResNet with data augmentation yielded a slightly improved accuracy of 0.8788 in training and 0.8685 in testing. The accuracy and loss plots for this model are shown in Fig. 17.

Fig. 18 compares the training and test accuracy for each advanced technique considered in this section. We note that ResNet with real time data augmentation has the highest classification accuracy at 0.8685.

*4.3. Discussion*

Fig. 19 summarizes the test performance of all of the machine learning and deep learning algorithms we tested on our LGESS image dataset. PCA with SVM is the best of the basic techniques with a test accuracy of 0.8535, while the best of the advanced techniques, namely, ResNet with data augmentation, improves on this result by 1.5%.
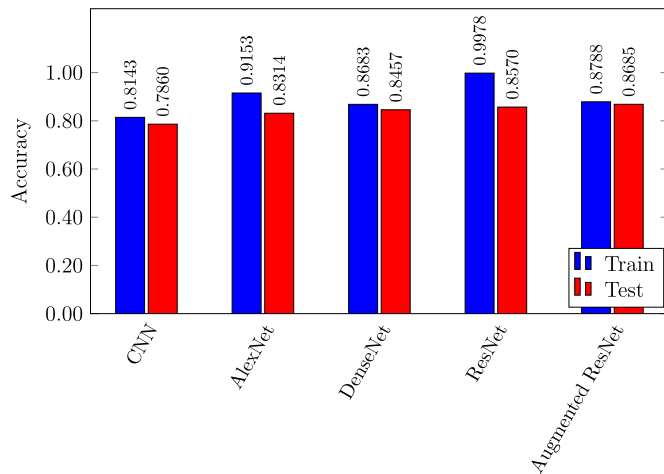


**Fig. 18.** Accuracies for advanced techniques.

## 5. Conclusion

Cancer results in massive health care expenses and a staggering death toll [2]. Early detection of cancer can significantly reduce mortality. Therefore, fast, easy-to-use, and high precision tools for automatic cancer screening could save lives and lower healthcare costs.

We are not aware of previous research into the efficacy of machine learning or deep learning models in screening for LGESS. Our results
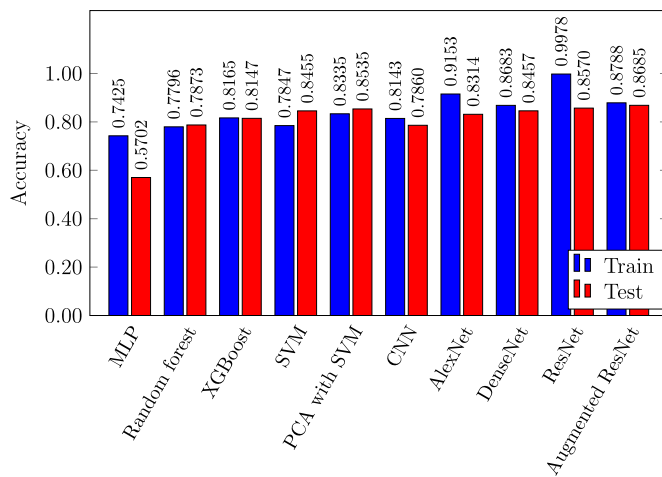
**Fig. 19.** Comparison of all techniques.

indicate that several models have the potential to provide surprisingly accurate results in a realistic setting. In particular, we found that specific SVM and ResNet architectures can achieve test accuracies in excess of 0.85.

Today, 75% of LGESS patients are incorrectly diagnosed with benign leiomyoma, leading to a delay in needed treatment and lower chances of longterm survival. In this work, we have demonstrated that machine learning can, at a minimum, serve as a second opinion, and thereby prevent many incorrect diagnoses.

In future studies, additional characteristics can be extracted from images, which would enable learning algorithms to work with a more refined and detailed feature space. Moreover, there exists a plethora of additional machine learning and deep learning algorithms that can be considered. Such further research could benefit individuals who suffer from the rare, but potentially deadly, disease of low grade endometrial stromal sarcoma.

## Declaration of Competing interest

None.

## Acknowledgements

## References

[1] Abhimanyu S. Ahuja, The Impact of Artificial Intelligence in Medicine on the Future Role of the Physician, PeerJ, 2019, e7702.

[2] American Cancer Society, The economic impact of cancer. https://www.cancer.org/cancer/cancer-basics/economic-impact-of-cancer.html, 2018.

[3] D. Anand, G. Ramakrishnan, A. Sethi, Fast GPU-enabled color normalization for digital pathology, in: 2019 International Conference on Systems, Signals and Image Processing, IWSSIP, 2019, pp. 219–224.

[4] Sarah Mohd Ashhar, Siti Salasiah Mokri, Ashrani Aizzuddin Abd Rahni, Aqilah Baseri Huddin, Noraishikin Zulkarnain, Nor Aniza Azmi, Thanuja Mahaletchumy, Comparison of deep learning convolutional neural network (CNN) architectures for CT lung cancer classification, International Journal of Advanced Technology and Engineering Exploration 8 (2021) 126–134.

[5] Kristin P. Bennett, Colin Campbell, Support vector machines: hype or hallelujah? SIGKDD Explorations 2 (2) (2000) 1–13.

[6] Anusha Bharat, N. Pooja, R. Anishka Reddy, Using machine learning algorithms for breast cancer risk prediction and diagnosis, in: 3rd International Conference on Circuits, Control, Communication and Computing, I4C, 2018, pp. 1–4.

[7] Dorry Boll, H. Rob, A. Verhoeven, A. van der Aa Maaike, Patrick Pauwels, E. Henrike, Karim-Kos, Jan Willem W. Coebergh, Helena C. van Doorn, Incidence and survival trends of uncommon corpus uteri malignancies in The Netherlands, Int. J. Gynecol. Canc. 22 (4) (2012) 599–606.

[8] Leo Breiman, Adele Cutler, Random forests™. https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm.

[9] CAMELYON16, 2021. https://camelyon16.grand-challenge.org/Data/.

[10] The cancer imaging archive: national cancer institute clinical proteomic tumor analysis consortium (cptac). https://www.cancerimagingarchive.net/, 2018.

[11] K.L. Chang, G.S. Crabtree, S.K. Lim-Tan, R.L. Kempson, M.R. Hendrickson, Primary uterine endometrial stromal neoplasms. a clinicopathologic study of 117 cases, The American Journal of Surgical Pathology 14 (5) (1990) 415–438.

[12] Saket S. Chaturvedi, Kajol Gupta, Prakash S. Prasad, Skin lesion analyser: an efficient seven-way multi-class skin cancer classification using mobilenet. https://arxiv.org/abs/1907.03220, 2020.

[13] Tianqi Chen, Carlos Guestrin, XGBoost: a scalable tree boosting system. https://arxiv.org/abs/1603.02754, 2016.

[14] Daphne Cornelisse, An intuitive guide to convolutional neural networks. https://medium.freecodecamp.org/an-intuitive-guide-to-convolutional-neural-networks-260c2de0a050, 2018.

[15] A. Gadducci, E. Sartori, F. Landoni, P. Zola, T. Maggino, A. Urgesi, Endometrial stromal sarcoma: analysis of treatment failures and survival, Gynecol. Oncol. 63 (2) (1996) 247–253.

[16] Ghoneim Ahmed, Ghulam Muhammad, M. Shamim Hossain, Cervical cancer classification using convolutional neural networks and extreme learning machines, Future Generat. Comput. Syst. 102 (2020) 643–649.

[17] L. Maryellen, Giger. Machine learning in medical imaging, J. Am. Coll. Radiol. 15 (3, Part B) (2018) 512–520.

[18] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2016, pp. 770–778.

[19] Healthline, What are the 12 leading causes of death in the United States?. https://www.healthline.com/health/leading-causes-of-death, 2018.

[20] M.R. Hendrickson, F.A. Tavassoli, R.L. Kempson, W.G. McCluggage, U. Haller, R. A. Kubik-Huch, Mesenchymal Tumours and Related Lesions, WHO classification of Tumours Pathology and Genetics Tumours of the Breast and Female Genital Organs, 2003, pp. 233–236, 2007.

[21] G. Huang, Z. Liu, L. Maaten, K. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2017, pp. 4700–4708.

[22] Reena Jain, Swaraj Batra, Ayesha Ahmad, Arifa Anwar Elahi, Monika Gupta, Poonam Saith, Low grade endometrial stromal sarcoma: a case report, Iran. J. Med. Sci. 40 (1) (2015) 81–84.

[23] S. Kharya, D. Dubey, S. Soni, Predictive machine learning techniques for breast cancer detection, Int. J. Comput. Sci. Inf. Technol. 4 (6) (2013) 1023–1028.

[24] Alex Krizhevsky, Ilya Sutskever, Hinton Geoffrey, ImageNet classification with deep convolutional neural networks, Neural Information Processing Systems 25 (2012).

[25] Loria Keith, Putting the AI in radiology, Radiology Today 19 (1) (2018) 10. https://www.radiologytoday.net/archive/rt0118p10.shtml.

[26] Ilias Maglogiannis, Elias Zafiropoulos, Ioannis Anagnostopoulos, An intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers, Appl. Intell. 30 (2009) 24–36.

[27] George D. Magoulas, Andriana Prentza, Machine Learning in Medical Applications, Springer, 2001, pp. 300–307.

[28] M.B.F. Bioscience, What is whole slide imaging?. https://www.mbfbioscience.com/whole-slide-imaging, 2021.

[29] Hengame Abbasi Mesrabadi, Karim Faez, Improving early prostate cancer diagnosis by using artificial neural networks and deep learning, in: 4th Iranian Conference on Signal Processing and Intelligent Systems, ICSPIS, 2018, pp. 39–42.

[30] OpenSlide, 2021. https://openslide.org.

[31] Mandeep Rana, Pooja Chandorkar, Alishiba Dsouza, Nikahat Kazi, Breast cancer diagnosis and recurrence prediction using machine learning techniques, Int. J. Renew. Energy Technol. 4 (4) (2015) 372–376.

[32] Pietro Renzulli, Rosemarie Weimann, Jean-Pierre Barras, Thierry-Pierre Carrel, Daniel Candinas, Low-grade endometrial stromal sarcoma with inferior vena cava tumor thrombus and intracardiac extension: radical resection may improve recurrence free survival, Surgical Oncology 18 (2009) 57–64.

[33] Mihir Sewak, Priyanka Vaidya, Chien-Chung Chan, Zhong-Hui Duan, SVM approach to breast cancer classification, in: Second International Multi-Symposiums on Computer and Computational Sciences, IMSCCS, 2007, pp. 32–37.

[34] Jonathon Shlens, A tutorial on principal component analysis. http://www.cs.cmu.edu/~elaw/papers/pca.pdf.

[35] Abhishek Vahadane, Tingying Peng, Sethi Amit, Shadi Albarqouni, Lichao Wang, Maximilian Baust, Katja Steiger, Anna Melissa Schlitter, Irene Esposito, Nassir Navab, Structure-preserving color normalization and sparse stain separation for histological images, IEEE Trans. Med. Imag. 35 (8) (2016) 1962–1971.

[36] R. Vijayarajeswari, P. Parthasarathy, S. Vivekanandan, A.A. Basha, Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform, Measurement 146 (2019) 800–805.

[37] Charles Wallis, History of the perceptron. https://web.csulb.edu/~cwallis/artificialn/History.htm, 2017.

[38] Sebastien C. Wong, Gatt Adam, Victor Stamatescu, D. Mark, McDonnell, Understanding data augmentation for classification: when to warp?. https://arxiv.org/pdf/1609.08764.pdf?, 2016.