This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3040155, IEEE Access

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2020.Doi Number

Metabolic Imaging based Sub-Classification of Lung Cancer

Mustafa Bicakci[1], Oguzhan Ayyildiz[2], Zafer Aydin[3], Alper Basturk[4], Seyhan Karacavus[5] and Bulent Yilmaz[2]

[1] Computer Engineering Department, Hasan Kalyoncu University, Gaziantep, Turkey
[2] Electrical and Electronics Engineering Department, Abdullah Gul University, Kayseri, Turkey
[3] Computer Engineering Department, Abdullah Gul University, Kayseri, Turkey
[4] Computer Engineering Department, Erciyes University, Kayseri, Turkey
[5] Nuclear Medicine Department, University of Health Sciences, İstanbul, Turkey

Corresponding author: Mustafa Bicakci (e-mail: mustafa.bicakci@hku.edu.tr).

This study was supported by TUBITAK (The Scientific and Technological Research Council of Turkey) under Project No: 113E188.

ABSTRACT Lung cancer is one of the deadliest cancer types whose 84% is non-small cell lung cancer (NSCLC). In this study, deep learning-based classification methods were investigated comprehensively to differentiate two subtypes of NSCLC, namely adenocarcinoma (ADC) and squamous cell carcinoma (SqCC). The study used 1457 ^{18}F-FDG PET images/slices with tumor from 94 patients (88 men), 38 of which were ADC and the rest were SqCC. Three experiments were carried out to examine the contribution of peritumoral areas in PET images on subtype classification of tumors. We assessed multilayer perceptron (MLP) and three convolutional neural network (CNN) models such as SqueezeNet, VGG16 and VGG19 using three kinds of images in these experiments: 1) Whole slices without cropping or segmentation, 2) cropped image portions (square subimages) that include the tumor and 3) segmented image portions corresponding to tumors using random walk method. Several optimizers and regularization methods were used to optimize each model for the diagnostic classification. The classification models were trained and evaluated by performing stratified 10-fold cross validation, and F-score and area-under-curve (AUC) metrics were used to quantify the performance. According to our results, it is possible to say that inclusion of peritumoral regions/tissues both contributes to the success of models and makes segmentation effort unnecessary. To the best of our knowledge, deep learning-based models have not been applied to the subtype classification of NSCLC in PET imaging, therefore, this study is a significant cornerstone providing thorough comparisons and evaluations of several deep learning models on metabolic imaging for lung cancer. Even simpler deep learning models are found promising in this domain, indicating that any improvement in deep learning models in machine learning community can be reflected well in this domain as well.

INDEX TERMS Convolutional Neural Networks, Deep Learning, PET imaging, Subtype Classification, Non-Small Cell Lung Cancer

I. INTRODUCTION

84% of lung cancers are non-small cell lung cancer (NSCLC) [1]. Adenocarcinoma (ADC) and squamous cell carcinoma (SqCC) are the two major subtypes of NSCLC. While 40% of lung cancers are ADC, 25-30% are of SqCC [2]. Understanding the effects of cytotoxic and biological agents suggests that subtype-specific treatment methods may be developed in the future. From this point of view, it shows the importance of the differentiation of ADC and SqCC [3]. Small lung biopsies (bronchoscopic, needle or core biopsies) and cytology specimens are used for lung cancer diagnosis. In general, standard morphological criteria by routine microscopy are essential to differentiate these two subtypes. However, bad morphology, especially in small specimens, can sometimes lead to difficult differentiation of tumors. On the other hand, sampling error that may occur during the biopsy procedure would be problematic in representing intratumoral heterogeneity.

Positron emission tomography (PET) is a functional imaging approach that has been widely used in medicine, and

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3040155, IEEE Access

provides significant diagnostic benefits. It is also a very effective and efficient method for staging and therapy of tumors. According to radiomics, medical images carry more information than obtained by visual examination [4]. With high-resolution PET images, it has become possible to obtain inferences with image processing methods. In this regard, it has become possible to recognize tumor characteristics using PET images.

Machine learning examines the algorithms that recognize complex patterns, and make predictions from the available data (such as the medical images in our case) to come up with intelligent decisions. In oncology, machine learning methods are used in different applications such as cancer prognosis and prediction [5], survival analysis [6], drug response [7], gene expression [8] and subtype differentiation [9]. As a machine learning methodology, artificial neural networks (ANN) based approaches involve learning and prediction algorithms that mimic the human brain in terms of recognition and decision-making. ANN methods are used in oncology in different applications such as tumor detection [10] and diagnosis [11]. Deep learning (DL) is an advanced neural network type with more layers to allow higher levels of abstraction. Deep convolutional neural networks (CNN) have brought breakthroughs in image-based studies [12]. They are highly successful in solving difficult problems such as recognizing objects in real world images [13], [14]. In recent years, there have been increasing number of cancer related studies using deep learning such as cancer detection and gene identification [15], skin cancer classification [16], histopathological diagnosis [17]. In addition, there are several studies [18], [19] showing that deep learning approaches are more effective and successful when compared to other machine learning methods in cancer-related classification problems. Diagnosis and classification studies using CT images [20], [21] are also available for lung cancer. However, CT images do not reflect the metabolic and heterogeneity information about the tumor which is highly critical in tumor subtype determination and is accessible when PET imaging is used. We should note that there are limited number of DL-based studies using PET images aiming lung cancer diagnosis [22]. Especially, subtype classification in NSCLC has not been explored using PET images and DL. Moreover, no studies have examined the contribution of peritumoral regions in this problem so far.

In this study, three different experiments were carried out to examine the contribution of peritumoral tissue in PET images/slices in subtype classification of tumors. Firstly, the success of classification was investigated by using PET images without any segmentation or cropping. Secondly, region-of-interests (ROIs) were cropped as square subimages including the tumors and peritumoral tissue. And finally, images containing only tumor tissue (no peritumoral tissue) obtained as a result of the segmentation study performed using "random walk" method were employed [23].

PET images or subimages including different amounts of peritumoral tissue obtained using three abovementioned approaches were employed as the training and test datasets in the classification of two subtypes of NSCLC. The multilayer perceptron (MLP) and convolutional neural network (CNN) based deep learning models were evaluated. As the CNN-based models we studied SqueezeNet, VGG16 and VGG19. SqueezeNet is used to reduce the model size, and the number of parameters while maintaining competitive accuracy. To achieve this, there are three main strategies used in SqueezeNet architecture [24].

Strategy-1: Smaller network by replacing 3x3 filters with 1x1 filters.

Strategy-2: Limited number of input channels (3x3 filters).

Strategy-3: Delayed downsampling for higher classification accuracy of large activation maps.

VGG model got the first place in the localization and classification part of the ImageNet Challenge in 2014 [25]. This model with 16- and 19-layer versions is described as very deep convolutional networks. At the end of the models, fully connected layers are followed by softmax. In addition to SqueezeNet and VGG models MLP models with different number of hidden layers with 64 neurons were examined. Studies published in 1986 [26] and 1987 [27] proposed the MLP architecture that has more than one hidden layer. While choosing these models, we tried to provide a variety of complexity. In this way, we aimed to investigate hyperparameter optimizations on different complexities. VGG models were used as deep architectures with varying depths. On the other hand, we have included SqueezeNet because it is an ambitious model for achieving success with fewer parameters by offering a structure different from the standard CNN architecture. In addition to the CNN models, the number of hidden layers in the MLP model was gradually increased to provide a controlled depth of complexity.

In summary, this study consists of subtype classification studies of NSCLC with deep learning using PET images. While doing this, models used for deep learning have been improved to achieve the most successful results by performing 10-fold stratified cross validation and hyperparameter optimization studies. It can be very useful to apply hyperparameter optimization and cross-validation techniques to increase the success of the models and prevent overfitting. For instance, as a recent study, in identification of clathrin proteins [28], 10-fold cross validation and different optimization methods were performed to improve a CNN-based model. Here, various optimization and regularization techniques were used and success of the models for all three experiments were compared. As a result, these evaluations were very important to see how deep learning methods are promising in subtype classification studies and to examine the contribution of peritumoral regions in PET images to the classification performance.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3040155, IEEE Access

II. MATERIALS AND METHODS

A. PATIENT POPULATION AND PET/CT IMAGING

This study was performed using ^{18}F-FDG PET / CT images of 94 patients with NSCLC. The images were obtained in the Nuclear Medicine Department of Acıbadem Kayseri Hospital in Kayseri between March 2010 and April 2014. The PET / CT imaging device was Siemens Biograph 6, HiRez. 10 to 15 mCi of ^{18}F-FDG was injected to the patients and PET / CT acquisitions were performed approximately 60 minutes after the injection. At 8 or 9 bed positions PET scanning was performed for 2 to 3 minutes at each position. Three-dimensional iterative reconstruction algorithm was used for the reconstruction of the images. Two nuclear medicine experts evaluated the images and 3D whole body projection using the e-Soft software platform (Siemens, USA). This study was approved by the Research Ethics Committee of the Kayseri Research and Training Hospital (KRTH) with protocol number 20.02.2013/55. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study. Five of the patients were female, and 88 were male (mean age was around 63, age interval was 39-84). Until now, no studies were published regarding the tumor variability among male or female NSCLC patients. Thus, we have designed this study without considering the effect of gender on cancer subtype characteristics. Thirty-eight patients were diagnosed with ADC, and the rest were diagnosed with SqCC. For the diagnosis, the specimens were obtained with fine needle or excisional biopsy, and the subtype evaluations were performed in the pathology department of KRTH.

B. IMAGE PREPROCESSING

For this study, three different datasets were prepared to be used in three experiments. In the first experiment, FDG-PET images containing 168x168 pixels obtained from the scans were determined as the first dataset without any processing. For the second experiment, each FDG-PET slice was manually cropped to include tumor and peritumoral tissue, and these ROIs formed our second dataset. Here, instead of specifying a standard bounding box size, the boxes were manually cropped to the extent that the tumor could fit inside. Since tumors did not have a standard shape, the amount of peritumoral areas varied in different slices. For the last experiment, the tumors were segmented on each slice using a standard "random walk" algorithm comprising our last dataset. It was used to distinguish tumor from background semi-automatically. In an unpublished preliminary study, we have tested the performances of segmentation approaches suggested for PET images such as Otsu's, active contour and random walk methods. We found random walk approach to perform the best among these approaches when compared to manual drawing of a nuclear medicine expert. Finally, we worked with a total of 1457 images that consisted of 516 ADC and 941 SqCC subtypes for all experiments. Figure 1 shows sample images from different patients with ADC and SqCC.

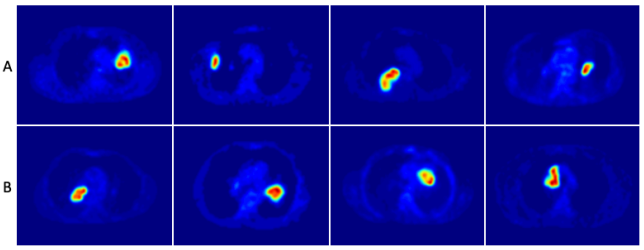


FIGURE 1. (A) Images from different patients with ADC and (B) SqCC.

In order to feed the images with a uniform size to deep learning models as the input, we had to specify a common and appropriate image size. In the first experiment, all images had already a fixed size, as there was no need for further processing. The datasets prepared for the second and third experiments, we considered the largest subimage size as 64x64 pixels that all the tumors in all slices could fit inside. In the bounding box, zero padding was performed by setting all pixel values to zero except ROIs for the second experiment and except segmented regions for the last experiment. At the end of all these processes, the datasets were fed to deep learning models and the training and test procedures were performed. Figure 2 depicts the schematic representation of abovementioned experiments.

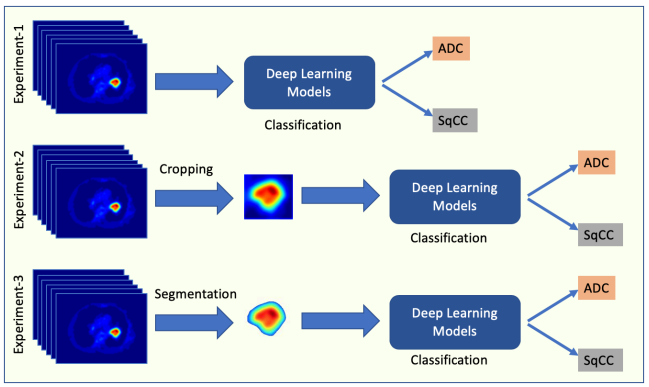


FIGURE 2. Schematic representation of three experiments studied in this work.

C. DATA SPLITTING AND TRAINING

In the literature, 10-fold cross validation method has been commonly used. However, there is no rule on this issue yet. In this study, the dataset was randomly split for stratified 10-fold cross validation which was used in many deep and conventional machine learning based bioimage and biosignal studies [29]. CNN and MLP models were trained from scratch starting with random weights using NVIDIA DGX-1 with NVIDIA Tesla V100 GPUs at Abdullah Gül University, High Performance Computing (HPC) Laboratory. The number of epochs for each training process was 100, and early stopping approach was used to prevent overfitting.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3040155, IEEE Access

D. HYPERPARAMETER OPTIMIZATION

For each model mentioned above several regularization and optimization techniques were used during the training process. The parameters were optimized to increase the classification success, and are listed in Table 1.

TABLE 1. Hyperparameters

Hyperparameters	Values
Optimizers	Momentum / RMSprop / Adadelta / Adam
Regularizations	None / L1 / L2 / Elastic
Dropout	None / 0.5
Batch Size	16 / 32
Learning Rate	1 / 0.1 / 0.01 / 0.001 / 0.0001
Hidden Layer (only MLP)	3 / 7 / 11 / 15

In the training process, stochastic gradient descent (SGD) approach helped us to find the optimum direction for minimizing the cost [30]. The aim of momentum, which is based on an advanced SGD logic, was to accelerate the progression in cases where the gradient did not change direction, and to slow the progression for situations where it changed direction [26]. In addition, some popular optimization methods, such as RMSprop [31], Adam [32], and Adadelta [33], which are similar to momentum, were used to estimate the optimum direction and speed for cost to move towards global minima.

The regularization methods we used were lasso [34], ridge regression [35] and elastic net [36] approaches. Lasso and ridge regression are also known as L1 and L2, respectively. As the penalty term, L1 dealt with the sum of the absolute values of the model parameters, and L2 dealt with the sum of their squares. Elastic net was a convex combination of lasso and ridge regression.

The main idea of the dropout regularization, which is known to prevent overfitting problems [37] especially for deep and complex networks with large number of parameters (such as VGG models), is to randomly drop nodes on the network during training, based on a certain ratio. In this study, the dropout with the ratio of 0.5 was used for the fully connected layers at the end of the models.

Table 2 shows the tuned CNN architectures where 64x64 images were used as the input. As can be seen in this table, the total number of parameters was directly proportional to the depth of the network.

In this study, MLP models with different number of hidden layers with 64 neurons were examined. The classification performance of MLP model was optimized by making its structure much deeper. For this purpose, MLP versions with 3, 7, 11 and 15 hidden layers have been generated and optimization studies have been performed. In the literature there is no rule of thumb for the number of neurons used in MLP architectures. While optimizing the number of hidden layers by using different values, we tried to optimize the number of neurons by the dropout method which randomly drops nodes from the network during the training process, based on the 0.5 ratio.

TABLE 2. Tuned CNN Architectures.

SqueezeNet			VGG16			VGG19		
Layer	Output	Param.	Layer	Output	Param.	Layer	Output	Param.
Input	64x64x1		Input	64x64x1		Input	64x64x1	
Conv.	31x31x64	640	2XConv.	64x64x64	37568	2XConv.	64x64x64	37568
Max Pooling	15x15x64		Max Pooling	32x32x64		Max Pooling	32x32x64	
Fire	15x15x128	11408	2XConv.	32x32x128	221440	2XConv.	32x32x128	221440
Fire	15x15x128	12432	Max Pooling	16x16x128		Max Pooling	16x16x128	
Max Pooling	7x7x128		3XConv.	16x16x256	1475328	4XConv.	16x16x256	2065408
Fire	7x7x256	45344	Max Pooling	8x8x256		Max Pooling	8x8x256	
Fire	7x7x256	49440	3XConv.	8x8x512	5899776	4XConv.	8x8x512	8259584
Max Pooling	3x3x256		Max Pooling	4x4x512		Max Pooling	4x4x512	
Fire	3x3x384	104880	3XConv.	4x4x512	7079424	4XConv.	4x4x512	9439232
Fire	3x3x384	111024	Max Pooling	2x2x512		Max Pooling	2x2x512	
Fire	3x3x512	188992	Flatten	2048		Flatten	2048	
Dropout	3x3x512		FC	4096	8392704	FC	4096	8392704
Fire	3x3x512	197184	Dropout	4096		Dropout	4096	
Dropout	3x3x512		FC	4096	16781312	FC	4096	16781312
Conv.	3x3x2	1026	Dropout	4096		Dropout	4096	
Global Average Pooling	2		FC	2	8194	FC	2	8194
Total Parameters		722370	Total Parameters		39895746	Total Parameters		45205442

To evaluate the performances of models, the F-score and area-under-curve (AUC) metrics were used in terms of correct detection of NSCLC subtype.

III. RESULTS AND DISCUSSION

The results of three experiments carried out within the scope of this study are shown in Table 3. The first experiment was performed on PET images of 168x168 size without any processing. The aim was to observe the success of deep learning models when the images were subject to classification without any extra effort. The purpose of the last two experiments was to examine the contribution of peritumoral regions and segmentation effort to subtype classification. Table 3 shows the performance and run time values of the SqueezeNet, VGG16, VGG19 and MLP models. According to the results of the first experiment shown in Table 3, it requires significantly high run time due to the size of the images. Here, we may conclude that the model performances were relatively low due to extra tissues unrelated to the tumor.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3040155, IEEE Access

TABLE 3. Results from three experiments showing classification performance of the models.

	Experiment-1			Experiment-2			Experiment-3		
Model	Run Time (s)	F-score (%)	AUC (%)	Run Time (s)	F-score (%)	AUC (%)	Run Time (s)	F-score (%)	AUC (%)
SqueezeNet	572	54	52	508	71	66	459	70	69
VGG16	2906	68	63	1165	73	68	523	73	69
VGG19	3758	68	65	1384	**74**	**69**	779	71	70
MLP	287	65	62	231	73	69	72	70	66

According to the results of all experiments, the most successful classification performances were obtained in the second experiment in which peritumoral areas were included. Here, VGG19 was the most successful model with 74% F-score and 69% AUC. In order to demonstrate the training process of this model, the graphs of average accuracy and average loss in 10-fold cross-validation are shown in Figure 3. Although the number of epochs were adjusted as 100, it can be seen in the figure that maximum 60 epochs were run for VGG19 due to early stopping. VGG16 and MLP have an F-score of 73%, followed by SqueezeNet with an F-score of 71%. Also, it is seen that MLP can achieve the success of complex CNN models with less run time thanks to the positive effect of peritumoral areas. When the second and third experiments were compared, the effect of the number of pixels containing information on run time is revealed. In the second experiment, due to the inclusion of tumor tissue as well as peritumoral areas, run time values were slightly higher than those of the third experiment. When the second and third experiments were evaluated together, it is possible to say that peritumoral regions both contribute to the success of models and make segmentation effort unnecessary. Here, the positive effect of peritumoral regions on the classification performance means an important contribution in this field. In a recent prostate study using MRI, the contribution of peritumoral areas was emphasized [38]. Therefore, we think that the effect of peritumoral areas in medical image analysis will be studied more in the future.

The hyperparameter values that led to the results given in Table 3 are shown in Table 4. According to this table, Adadelta was successful as an optimizer for SqueezeNet, while Momentum was successful on all other models. While SqueezeNet used L2 regularization to achieve the best performance, all other models reached the optimum point without using any of L1, L2 or Elastic regularization approaches. While dropout method worked for all CNN models, for MLP it did not work. The optimal batch size and learning rate values were 32 and 0.1 respectively for SqueezeNet and MLP, 16 and 0.001 for VGG models. In addition, it is not surprising that the same combinations succeeded on both VGG16 and VGG19 models, which do not differ except for the number of layers.

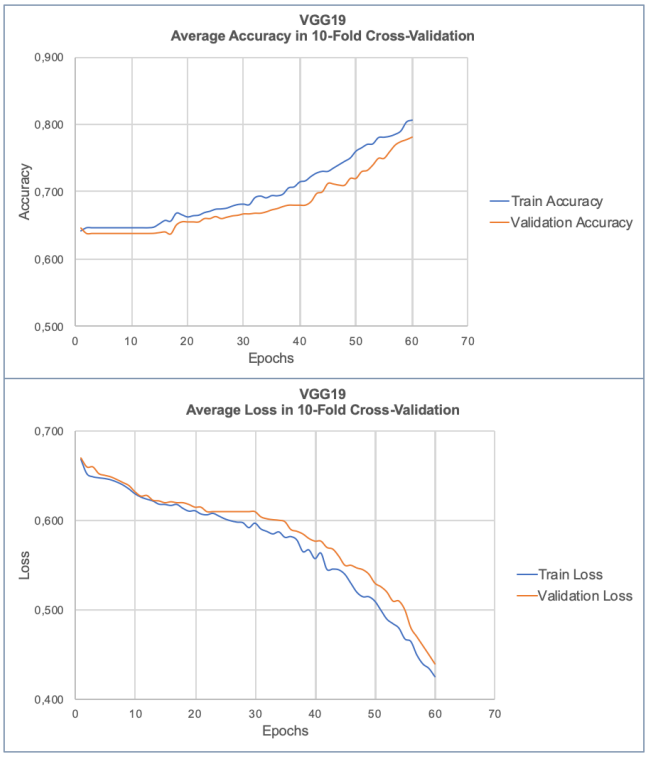


FIGURE 3. The average accuracy and loss of training and validation in the 10-fold cross-validation for VGG19 model in the second experiment.

We recognize that we have only considered the classification of subtypes as ADC and SqCC which constitute approximately 70% of lung cancer, and neglected other possible subtypes such as large cell carcinoma (LCC). However, LCC accounts for only 5–10% [39]. That is why we have limited subtypes to the first two categories.

TABLE 4. Hyperparameter values yielding performances given in Table 3.

Model	Optimizer	Regularization	Dropout	Batch Size	Learning Rate
SqueezeNet	Adadelta	L2	0.5	32	0.1
VGG16	Momentum	None	0.5	16	0.001
VGG19	Momentum	None	0.5	16	0.001
MLP	Momentum	None	None	32	0.1

Since the variety of models used in the study can contribute to the comparative evaluation of the results, it may be useful to examine different models such as ResNet and DenseNet in addition to the deep learning models tested here. Furthermore, it is thought that expanding the datasets using data augmentation techniques may be useful for increasing the subtype classification success.

A study on PET images [40] revealed that noise reduction and partial volume correction (PVC) methods improve the segmentation accuracy. Improvements in segmentation may indirectly lead to an increase in classification performance. Therefore, these methods have the potential to be examined in future classification studies with segmented images.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3040155, IEEE Access

Moreover, the effect of different segmentation approaches may also be investigated in this context. In random walk approach the selection of seed points, one of which should be at a representative background region and the other be on the tumor region, may have an impact on the segmentation accuracy. It is clear that when the semi/automatic approach segments the tumor region incorrectly and misses critical image features related to the subtype then the classification performance would be diminished.

Currently, biopsy is still needed to characterize lung cancer subtypes. In this study, we investigated the feasibility of using PET images directly without the need for pathological studies to classify the subtypes of NSCLC which is a common effort put forward by many researchers working in this field.

IV. CONCLUSION

The importance of this study is to perform subtype tumor classification (NSCLC) from only PET images without the need for a pathological procedure. We investigated the feasibility of several popular deep learning methodologies. To the best of our knowledge, deep learning-based models have not been applied to subtype classification of NSCLC in PET imaging domain. Therefore, this study is a significant contribution providing thorough comparisons and evaluations of several deep learning models on metabolic imaging for lung cancer. Even simpler deep learning models were found to be promising in this field, indicating that any improvement in deep learning models in machine learning community can be reflected well in this domain as well. In addition, no studies investigating the contribution of peritumoral areas to subtype classification of NSCLC using deep learning on PET images were found in the literature. From this perspective, this study includes reviews that shed light on similar studies in the future. This study was carried out on 1457 PET slices obtained from 94 patients. In the future, we aim to increase the success by applying data augmentation methods, and plan to make further contributions to this topic by using various deep learning models such as ResNet and DenseNet, as well as multi-task learning studies. In addition, we aim to examine the performance of deep learning models by making segmentation improvements such as noise reduction and partial volume correction. Finally, we think that better performances can be obtained by combining PET and CT images when compared to using solely PET or CT images. In our research laboratory we plan on such a study in the near future.

REFERENCES

[1] F. C. Detterbeck, D. J. Boffa, and L. T. Tanoue, "The new lung cancer staging system," Chest, vol. 136, no. 1, pp. 260–271, 2009.

[2] V. K. Anagnostou, A. T. Dimou, T. Botsis, E. J. Killiam, M. D. Gustavson, R. J. Homer, D. Boffa, V. Zolota, D. Dougenis, L. Tanoue, S. N. Gettinger, F. C. Detterbeck, K. N. Syrigos, G. Bepler and D. L. Rimm, "Molecular classification of nonsmall cell lung cancer using a 4-protein quantitative assay," Cancer, vol. 118, no. 6, pp. 1607–1618, 2012.

[3] W. D. Travis, E. Brambilla, A. G. Nicholson, Y. Yatabe, J. H. M. Austin, M. B. Beasley, L. R. Chirieac, S. Dacic, E. Duhig, D. B. Flieder, K. Geisinger, F. R. Hirsch, Y. Ishikawa, K. M. Kerr, M. Noguchi, G. Pelosi, C. A. Powell, M. S. Tsao and I. Wistuba, "The 2015 World Health Organization classification of lung tumors: Impact of genetic, clinical and radiologic advances since the 2004 classification.," J. Thorac. Oncol., vol. 10, no. 9, pp. 1243–60, 2015.

[4] V. Kumar, Y. Gu, S. Basu, A. Berglund, S. A. Eschrich, M. B. Schabath, K. Forster, H. J. Aerts, A. Dekker, D. Fenstermacher, D. B. Goldgof, L. O. Hall, P. Lambin, Y. Balagurunathan, R. A. Gatenby and R. J. Gillies, "Radiomics: The process and the challenges," Magn. Reson. Imaging, vol. 30, no. 9, pp. 1234–1248, 2012.

[5] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, and D. I. Fotiadis, "Machine learning applications in cancer prognosis and prediction," Comput. Struct. Biotechnol. J., vol. 13, pp. 8–17, 2015.

[6] C. Parmar, P. Grossmann, J. Bussink, P. Lambin, and H. J. W. L. Aerts, "Machine Learning methods for Quantitative Radiomic Biomarkers," Sci. Rep., vol. 5, pp. 1–11, 2015.

[7] M. P. Menden, F. Iorio, M. Garnett, U. McDermott, C. H. Benes, P. J. Ballester and J. Saez-Rodriguez "Machine Learning Prediction of Cancer Cell Sensitivity to Drugs Based on Genomic and Chemical Properties," PLoS One, vol. 8, no. 4, 2013.

[8] A. C. Haury, P. Gestraud, and J. P. Vert, "The influence of feature selection methods on accuracy, stability and interpretability of molecular signatures," PLoS One, vol. 6, no. 12, pp. 1–12, 2011.

[9] O. Ayyıldız, Z. Aydın, B. Yılmaz, S. Karaçavuş, K. Şenkaya, S. İçer, A. Taşdemir, E. Kaya, "Lung cancer subtype differentiation from positron emission tomography images", Turkish Journal of Electrical Engineering and Computer Science, 28 (1) , 262-274 . DOI: 10.3906/elk-1810-154, 2020.

[10] L. Matulewicz, J. F. A. Jansen, L. Bokacheva, H. A. Vargas, O. Akin, S. W. Fine, A. Shukla-Dave, J. A. Eastham, H. Hricak, J. A. Koutcher and K. L. Zakian, "Anatomic segmentation improves prostate cancer detection with artificial neural networks analysis of 1H magnetic resonance spectroscopic imaging," J. Magn. Reson. Imaging, vol. 40, no. 6, pp. 1414–21, 2014.

[11] H. G. Zadeh, J. Haddadnia, M. Hashemian, and K. Hassanpour, "Diagnosis of breast cancer using a combination of genetic algorithm and artificial neural network in medical infrared thermal imaging," Iran. J. Med. Phys., vol. 9, no. 4, pp. 265–274,

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3040155, IEEE Access

2012.

[12] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[13] Jia Deng, Wei Dong, R. Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conf. Comput. Vis. Pattern Recognit., pp. 248–255, 2009.

[14] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg and L. Fei-Fei "ImageNet Large Scale Visual Recognition Challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, 2015.

[15] P. Danaee, R. Ghaeini, and D. A. Hendrix, "A Deep Learning Approach for Cancer Detection and Relevant Gene Identification," in Biocomputing 2017, 2017, vol. 53, no. 4, pp. 219–229.

[16] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau and S. Thrun "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, 2017.

[17] G. Litjens, C. I. Sanchez, N. Timofeeva, M. Hermsen, I. Nagtegaal, I. Kovacs, C. Hulsbergen-van de Kaa, P. Bult, B. van Ginneken and J van der Laak, "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis.," Sci. Rep., vol. 6, no. April, p. 26286, 2016.

[18] Y. Yuan, Y. Shi, C. Li, J. Kim, W. Cai, Z. Han and D. D. Feng, "DeepGene: an advanced cancer type classifier based on deep learning and somatic point mutations," BMC Bioinformatics, vol. 17, no. S17, p. 476, 2016.

[19] M. Halicek, G. Lu, J. V. Little, X. Wang, M. Patel, C. C. Griffith, M. W. El-Deiry, A. Y. Chen and B. Fei, "Deep convolutional neural networks for classifying head and neck cancer using hyperspectral imaging," J. Biomed. Opt., vol. 22, no. 6, p. 060503, Jun. 2017.

[20] F. Ciompi, K. Chung, S. J. Van Riel, A. A. A. Setio, P. K. Gerke, C. Jacobs, E. T. Scholten, C. Schaefer-Prokop, M. M. W. Wille, A. Marchiano, U. Pastorino, M. Prokop and B. van Ginneken, "Towards automatic pulmonary nodule management in lung cancer screening with deep learning," Sci. Rep., vol. 7, no. March, pp. 1–11, 2017.

[21] Q. Song, L. Zhao, X. Luo, and X. Dou, "Using Deep Learning for Classification of lung nodules on Computed Tomography Images," J. Healthc. Eng., pp. 1–10, 2017.

[22] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginnekan and C. I. Sanchez, "A survey on deep learning in medical image analysis," Med. Image Anal., vol. 42, no. December, pp. 60–88, 2017.

[23] W. Ju, D. Xiang, B. Zhang, L. Wang, I. Kopriva and X. Chen "Random walk and graph cut for co-segmentation of lung tumor on PET-CT Images," IEEE Transactions on Image Processing, vol. 24 no. 12, pp. 5854-5867, 2015.

[24] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," pp. 1–13, 2016.

[25] A. Simonyan, K.; Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv, pp. 1409.1556, 2014.

[26] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," Nature, vol. 323, no. 6088, pp. 533–536, 1986.

[27] R. P. Lippmann, "An Introduction to Computing with Neural Nets," IEEE ASSP Mag., vol. 4, no. 2, pp. 4–22, 1987.

[28] N. Q. K. Le, T. T. Huynh, E. K. Y. Yapp, and H. Y. Yeh, "Identification of clathrin proteins by incorporating hyperparameter optimization in deep learning and PSSM profiles," Comput. Methods Programs Biomed., vol. 177, pp. 81–88, Aug. 2019.

[29] O. S. Lih, V. Jahmunah, T. R. San, E. J. Ciaccio, T. Yamakawa, M Tanabe, M. Kobayashi, O. Fauste and U. R. Acharya, "Comprehensive electrocardiographic diagnosis based on deep learning," Artificial Intelligence in Medicine, vol. 103, 101789, 2020.

[30] H. Robbins, S. Monro, "A Stochastic Approximation Method," Ann. Math. Statist. vol. 22, no. 3, 400-407, 1951.

[31] T. Tijmen and G. E. Hinton, ''Lecture 6.5-RmsProp: Divide the gradient by a running average of its recent magnitude,'' COURSERA: Neural Netw. Mach. Learn., vol. 4, no. 2, pp. 26–31, 2012.

[32] D. P. Kingma and J. Ba, ''Adam: A method for stochastic optimization,'' in Proc. ICLR (Poster), pp. 1–15, 2015.

[33] M. D. Zeiler, ''ADADELTA: An adaptive learning rate method,'' arXiv:1412.6980, 2014.

[34] R. Tibshirani, "Regression shrinkage and selection via the Lasso," Journal of the Royal Statistical Society. Series B (Methodological), pp. 267–288, 1996.

[35] A. E. Hoerl, and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," Technometrics. 12 (1), pp. 55–67, 1970.

[36] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," Journal of the Royal Statistical Society: Series B (Statis- tical Methodology), vol. 67, no. 2, pp. 301–320, 2003.

[37] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, ''Dropout: A simple way to prevent neural networks from overfitting,'' J. Mach. Learn. Res., vol. 15, no. 1, pp. 1929–1958, 2014.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3040155, IEEE Access

[38] A. Algohary, R. Shiradkar, S. Pahwa, A. Purysko, S. Verma, D. Moses, R. Shnier, A.-M. Haynes, W. Delprado, J. Thompson, S. Tirumani, A. Mahran, AR Rastinehad, L. Ponsky, PD Stricker, and A. Madabhushi, "Combination of Peri-Tumoral and Intra-Tumoral Radiomic Features on Bi-Parametric MRI Accurately Stratifies Prostate Cancer Risk: A Multi-Site Study," Cancers, vol. 12, no. 8, p. 2200, Aug. 2020.

[39] C. Zappa and SA Mousa, "Non-small cell lung cancer: current treatment and future advances," Translational Lung Cancer Research, vol. 5, no.3, pp. 288-300, 2016.

[40] Z. Xu, M. Gao, G. Z. Papadakis, B. Luna, S. Jain, D. J. Mollura, U. Bagci, "Joint solution for PET image segmentation, denoising, and partial volume correction," Med. Image Anal., vol. 46, pp. 229–243, 2018.



MUSTAFA BICAKCI was born in Soke, Aydin, Turkey in 1985. He received the B.S. in Computer Systems Teaching from Selcuk University in Konya, Turkey in 2008 and the M.S. degree in Computer and Information Systems from Gazi University, Ankara, Turkey in 2013. He is currently pursuing the Ph.D. degree in Electrical and Computer Engineering at Abdullah Gul University, Kayseri, Turkey. He has been working as an instructor in Department of Computer Engineering at Hasan Kalyoncu University, Gaziantep, Turkey since 2013. His research interests include deep learning, machine learning, data science and biomedical engineering.



OGUZHAN AYYILDIZ was born in Erzurum, Turkey in 1988. He received the B.S. in Electrical and Electronics Engineering from Bilkent University in Ankara, Turkey in 2013 and the M.S. degree in Biomedical Engineering from Erciyes University, Kayseri, Turkey in 2016. He is currently pursuing the Ph.D. degree in Electrical and Computer Engineering at Abdullah Gul University, Kayseri, Turkey. He has been working as an instructor since 2013. His research interests include deep learning, machine learning, biomedical engineering.



DR. ZAFER AYDIN received the B.Sc. and M.Sc. degrees in Electrical and Electronics Engineering from Bilkent University in 1999 and 2001. He received his Ph.D. degree from Electrical and Computer Engineering Department of Georgia Institute of Technology in 2008. Prior to joining Abdullah Gul University as a faculty of Computer Engineering in 2014, he worked as a post-doctoral research fellow at the Genome Sciences Department of Washington University from 2008 to 2011 and as an Assistant Professor in Bahcesehir University from 2011 to 2014. His research concentrates on developing bioinformatics and machine learning methods for protein structure prediction. He has co-authored 11 journals and 15 conference papers.



DR. ALPER BASTURK received his BS degree in Electronics Engineering from Erciyes University, Kayseri, Turkey, in July-1998. He then joined as a research assistant to the Dept. of Electronics Eng. of Erciyes University. He received his MS and PhD degrees in both of Electronics Engineering from Erciyes University in August-2001 and November-2006, respectively. In 2006, he joined the Computer Hardware Division, Department of Computer Engineering, Erciyes University, where he is currently an Associate Professor and head of Computer Hardware Division. Between 2010 and 2011, he was a visiting scholar to the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, New York, USA. He guest-edited several special issues for various journals and has published more than 50 articles in leading journals and conferences. His research areas are deep learning, machine learning, neural networks, fuzzy and neuro-fuzzy systems, intelligent optimization, digital signal and image processing, high performance computing and various applications of these techniques.



DR. SEYHAN KARACAVUS was born in 1973 in Kayseri, Turkey. She graduated from Erciyes University Medical Faculty in 1997. She received her specialty degree at Department of Nuclear Medicine of Erciyes University in 2010. In addition, she received M.Sc. degrees in Biomedical Engineering from Erciyes University in 2015 and she is currently pursuing the Ph.D. degree in Biomedical Engineering at Erciyes University, Kayseri, Turkey. In 2016, she joined Health Sciences University, where she is currently an Associate Professor and head of Department of Nuclear Medicine in Kayseri City Hospital. Her research focuses on Nuclear oncology, radionuclide therapy, theranostic applications, texture analysis and image processing on PET/CT and SPECT images for various diseases evaluation. She published more than 50 refereed journal papers and 70 national and international conference proceedings.



DR. BULENT YILMAZ received his B.Sc. and M.Sc. degrees in Electrical-Electronics Engineering from the Middle East Technical University, Ankara, Turkey in 1997 and 1999, respectively. He acquired his Ph.D. in Bioengineering Department of the University of Utah, Salt Lake City, Utah. He is currently a professor in Electrical and Electronics Engineering department at Abdullah Gul University, Kayseri, Turkey. His current research interests are biomedical signal and image processing applications on brain-computer interfaces, neuromarketing, processing of PET/CT and endoscopic images for cancer and disease assessment.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/