A Multi-Resolution Deep Learning Framework for Lung Adenocarcinoma Growth Pattern Classification

Najah Alsubaie[1,2], Muhammad Shaban[1], David Snead[3], Ali Khurram[4], and Nasir Rajpoot[1,3]

[1] Department of Computer Science, Warwick University, UK
[2] Department of Computer Science, Princess Nourah University, KSA
[3] Department of Pathology, University Hospitals Coventry and Warwickshire, UK
[4] School of Clinical Dentistry, University of Sheffield, Sheffield

Abstract. Identifying lung adenocarcinoma growth patterns is critical for diagnosis and treatment of lung cancer patients. Growth patterns have variable texture, shape, size and location. They could appear individually or fused together in a way that makes it difficult to avoid inter/intra variability in pathologists reports. Thus, employing a machine learning method to learn these patterns and automatically locate them within the tumour is indeed necessary. This will reduce the effort, assessment variability and provide a second opinion to support pathologies decision. To the best of our knowledge, no work has been done to classify growth patterns in lung adenocarcinoma. In this paper, we propose applying deep learning framework to perform lung adenocarcinoma pattern classification. We investigate what contextual information is adequate for training using patches extracted at several resolutions. We find that both cellular and architectural morphology features are required to achieve the best performance. Therefore, we propose using multi-resolution deep CNN for growth pattern classification in lung adenocarcinoma. Our preliminarily results show an increase in the overall classification accuracy.

Keywords: Lung adenocarcinoma, histology subtype, growth patterns, deep learning, digital pathology, whole slide image analysis

1 Introduction

Lung adenocarcinoma is one of the most common types of cancer in the world. The main characteristic that makes it distinguishable form the other types of non-small cell lung cancers is the presence of specific tumour morphology patterns [4, 5]. These patterns are called growth patterns or adenocarcinoma histology subtypes. In 2011, a new lung adenocarcinoma classification system was proposed by a joint group of IASLC/ATS/ERS [5]. The classification system proposes using the predominant growth pattern for prognosis purposes.

According to the latest 2015 WHO of lung tumour [9], lung adenocarcinoma has five growth patterns: acinar, papillary, micro-papillary, solid and lepidic.

Figure 2 shows examples of theses patterns. Acinar pattern is identified by a glandular formation. The columnar shape tumour cells form acini and tubules. Solid tumour is identified as a sheet and nests of tumour cells. Lepidic pattern is composed of neoplastic cells growing along the alveolar walls. It often has no architectural complexity and does not contain stromal, vascular or lymphatic invasion. Papillary is identified by the presence of a papillary structure with fibrovascular cores that replace the alveolar. At least one blood vessel should present within papillary structure, psmmoma bodies might present as well. The micropapillary pattern has smaller papillae. It has no fibrovascular cores or blood vessels. Several studies show that these patterns are correlated with patient survival [13, 11]. They show that solid [6, 3] and micro-papillary have the worst prognosis [4]. These studies agree that micro-papillary is an independent factor for overall survival, while lepidic predominant cases have good survival [8].

Adenocarcinoma tumour could contain one or more of these patterns in the same biopsy. The clinical routine for identifying the presence of these patterns is normally by visual examination of the tumour under microscope. Growth pattern is identified by presence of the percentage of each subtype in 5% increment [10]. The predominant pattern is then assigned to the case. The examination process is tedious and often leads to inter/intra variability [7, 12]. This is due to the difficulty of identifying one type from another and the localisation of these patterns when they are mixed in one tumour region. Therefore, an automatic subtype classification is important to provide an objective assessment that could support pathologists decision and provide a second opinion.

In this paper, we propose a deep learning based framework that mimics the pathologists clinical routine in growth pattern assessment. Our method identifies all possible patterns by examining the tissue at several resolutions. We evaluate the performance of each network individually. We find that cellular and morphological architecture are both useful for network to learn these patterns. Thus, we propose a multi-resolution deep CNN where we train our network using images from different resolutions to improve the overall accuracy. We measure the pattern classification accuracy based on the pathologist annotations and the preliminary results are promising. The remaining of the paper is organised as follows: Section 2 describes the proposed method. Section 3 gives the experimental settings and discusses preliminary results. Finally, Section 4 summarises the findings and suggests future directions.

2 Methodology

Pathologists look into the tissue under the microscope at several magnifications by zooming in and out until they could recognise what pattern they are looking at. They zoom into high magnification to examine cellular features, then gradually move to lower magnifications to identify regional morphology.

In deep learning, a similar approach is performed. Each layer calculates image features at one resolution. Next layer, more abstract features are calculated. However, in deep learning the context is limited to the provided image. There-

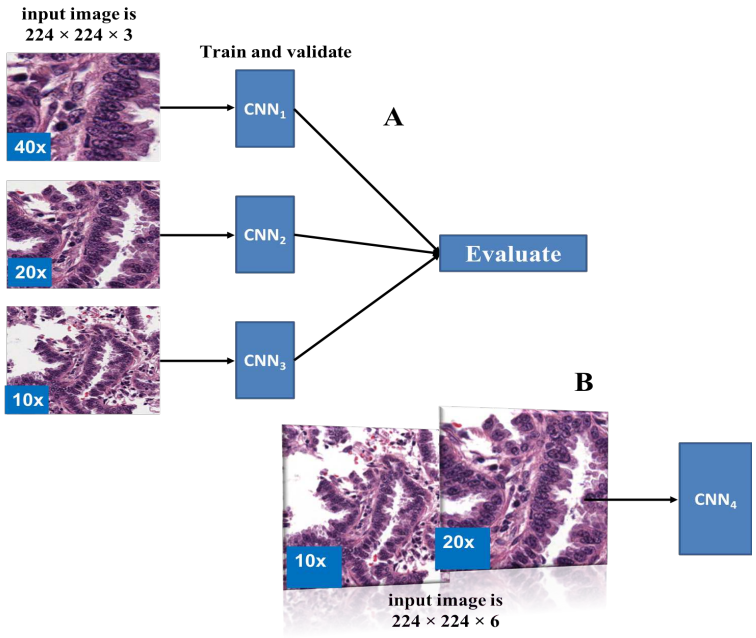

Fig. 1. Algorithm flowchart. A) A neural network is trained independently on dataset extracted at four different resolution. We then evaluate the accuracy of each network which indicates which context is adequate for training. B) Network is trained using image resolutions that provide highest performance in A.

fore, a question that we would like to answer in this paper is as follows: Can deep learning framework exploit patterns from more than one resolutions? Given training images from different resolutions, how this could affect the performance of deep learning network. In the following sections, we present a framework to answer above questions.

2.1 The Datasets

In these experiments we used lung adenocarcinoma images from the MICCAI 2017 CPM Challenge[1]. The dataset contains a total number of 10 images of H&E lung adenocarcinoma sample images. The key point in this experiment is patch generation. Figure 2 shows one sample image extracted at four different resolutions for each growth pattern. Tissue specimens contain other type of tissue

[1] http://miccai.cloudapp.net/competitions/

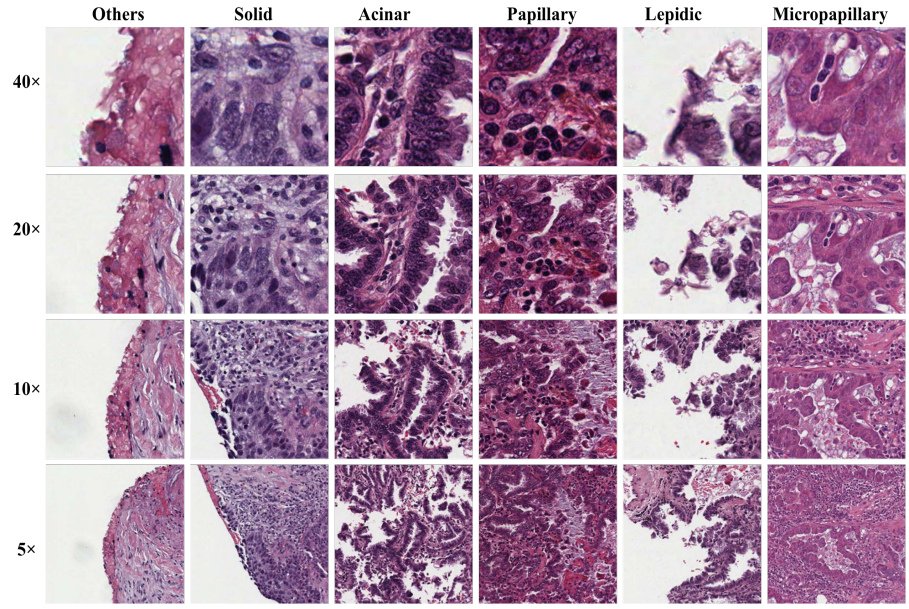

Fig. 2. Data set extraction. Each row represents one dataset. Each column represents images from one class extracted in a decreasing magnification. Notice, the context is gradually increasing as we move one magnification level while maintaining same image size i.e. (224×224). In 40× dataset, cellular information is noticeable, however it is very difficult to identify patterns at this level. As we decrease the magnification, the cells start to show some architectural arrangements.

that does not comprise one of the growth patterns. These tissue could be one of the following:

- background: this includes white background.
- non tumour tissue: this includes all other type of tissues, i.g: stromal tissue, smooth muscles, necrosis, etc.
- group of tumour cells that does not form pattern.

  Therefore, we would like our network to be able to classify a given image into one of the following six classes: acinar, papillary, micropapillary, solid and others, where others are all images that either background, non tumour tissue, or tumour cells that did not form any pattern.

  Next, two pathologists DS and AK annotated 10 whole slide images of H&E lung adenocarcinoma. We then generated patches using the annotated regions. Figure 3 shows an example of patch extraction. In Figure 3, we expand the context within each image by gradually moving from the highest resolution (top image) towards lower resolution bottom image. All images are of the same size so we could employ the same network architecture for training. Therefore, we could evaluate the performance of the CNNs when trained on each resolution, independently.

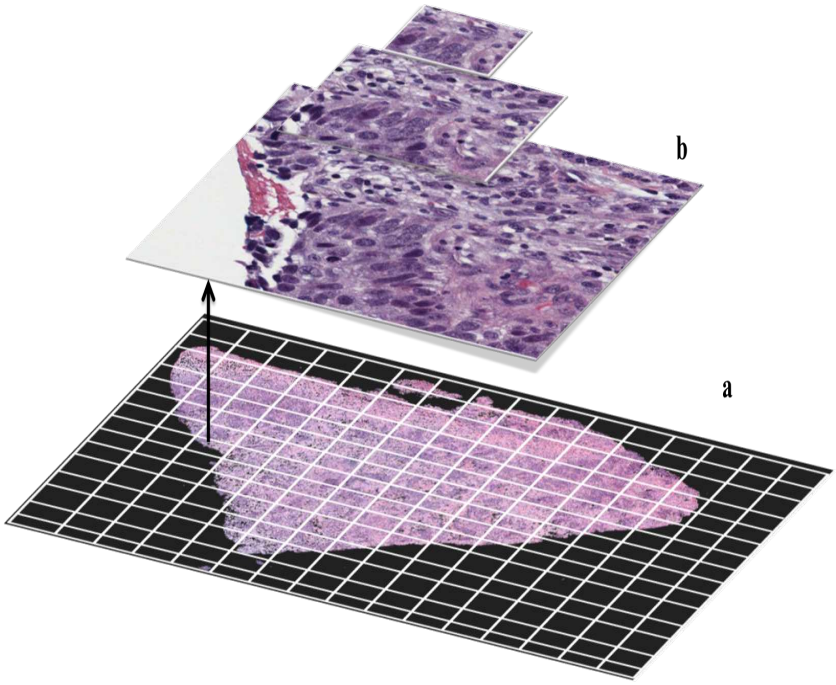

Fig. 3. Patch extraction for lung adenocarcinoma growth pattern classification. a) is the WSI. The WSI is divided into small squares of equal sizes. Then, center of each square is used as reference point to extract patches at each of the required magnification levels. (b) patches are extracted from the same center point denoted by the arrow in (a).

from the same annotated slides, We prepared three different training datasets (40×, 20×, 10×) using the three whole slide image resolutions. In each dataset, we have six classes: solid, acinar, papilary, lepidic, micropapillary, and others (non pattern). Figure 2 shows examples of images in each dataset.

3 Experimental Results and Discussion

Images are augmented by applying Gaussian filter, rotation and flipping at different angles. Data are divided into 80% for training and 20% for testing. In total we had 7K images per class for training and 2K images per class for validation. Then, we train ResNet [1] architecture on each dataset independently. Training and validation parameters were fixed in all the datasets. Image size is fixed to 224×224, all RGB channels is used as features for training, Adam method [2] was used for stochastic optimization, learning rate is 0.001. Max number of epochs is 50 with maximum number of training steps 1000, training batch size is 64, evaluation batch size is 32.

Table 1. Validation accuracy for Resent trained on four different magnifications 40×, 20×, 10×, and 20× with 10×. In each network, best and worst accuracy per pattern is highlighted with bold and italic; respectively

Magnification	Solid	Acinar	Papillary	Lpidic	Micropapillary	Others	Mean
40×	89%	84%	**93%**	83%	*81%*	87%	86%
20×	94%	**96%**	*73%*	**96%**	95%	83%	90%
10×	94%	94%	*77%*	94%	**98%**	84%	90%
20× 10×	95%	**99%**	*79%*	88%	96%	87%	**91%**

We evaluate each network independently in terms of accuracy. Table 1 shows the accuracy for each network. Network trained on 40× images have the lowest accuracy. The mean performance is increased when training network on images of 20× and then also when trained on 10× magnification. The best performance for classification of acinar and lepidic is achieved by 20× network while 10× network has the best performance for classifying micropapillary pattern. From this table we can conclude that the contexts within 10× and 20× images are good enough for the network to learn each pattern. Therefore, training a network using images from both magnifications could improve the overall accuracy. This is shown in the last row in Table 1. The network is able to classify classes (solid, acinar and others) better than previous networks.
Figure 5 shows the confusion matrix for each network. From the confusion matrices we can notice the following: first, networks are mostly confused between others and micropapillary classes. However, when combining 20× and 10× together, we are able to provide both cellular features and pattern morphology to the learning process. Thus, in 5 (d) the others vs. micropapillary confusion is reduced significantly. Second, acinar and papillary are also confusing classes for the networks. This is due to the significant similarity between these two classes and the possibility that one class might contain instances of the other. Thus, increasing the training dataset could improve the accuracy. We evaluated the the proposed framework on the whole slide images. Results are shown in Figure 4.

4 Conclusions

In this work, we propose a deep learning approach for classification of growth patterns in lung adenocarcinoma images. We evaluate the performance of deep learning framework when trained on an increasing contextual information. We find that network requires both cellular information and their morphological architecture to differentiate between growth patterns. We train a deep CNN on images extracted from two magnification levels and the overall accuracy increased. In the future work, we plan to increase training patches in our dataset and provide a quantitative evaluation of the network on whole slide images.

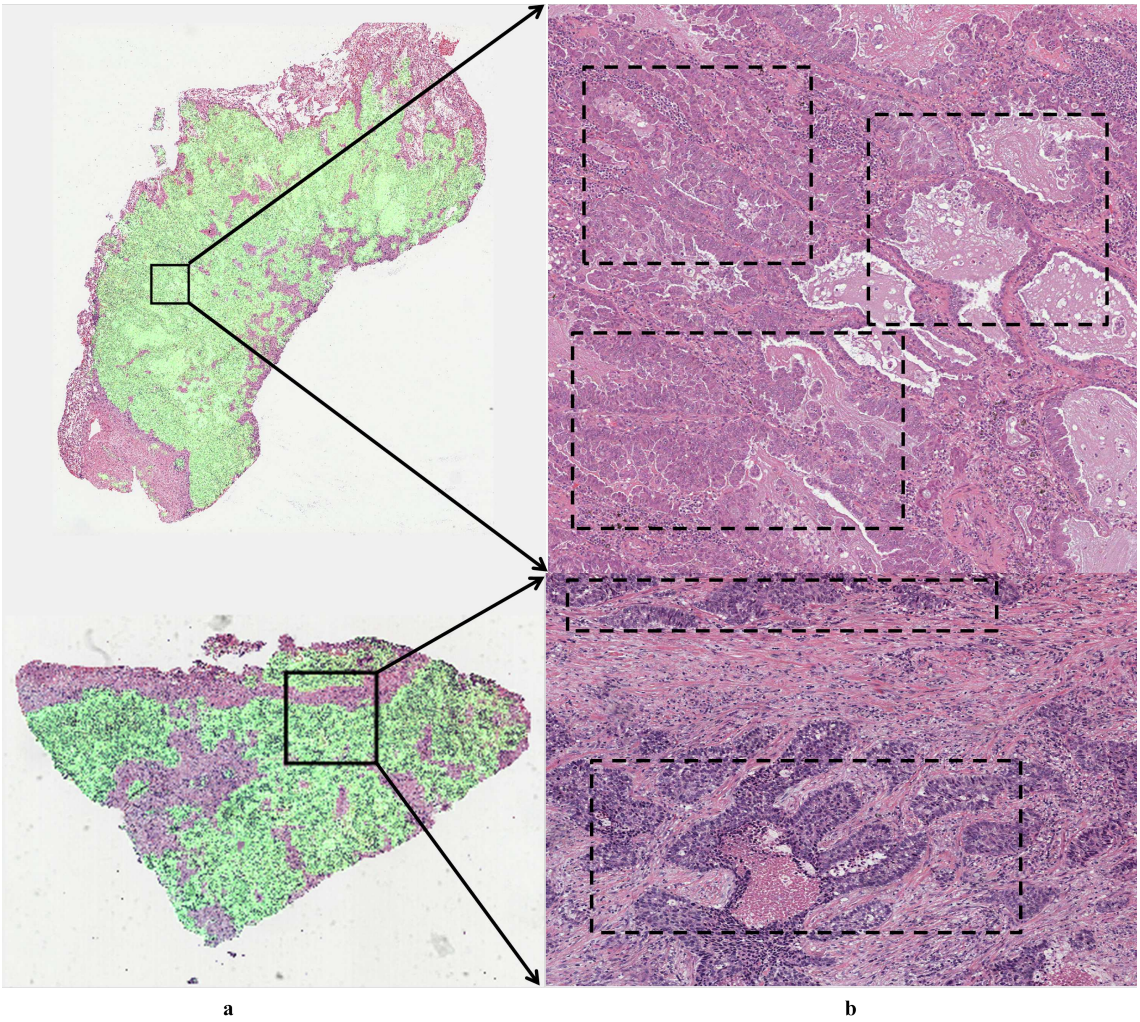

Fig. 4. (a, b) shows the results for two whole slide images using ResNet architecture trained on images from 20× and 10× magnifications. Top row shows an image with micropapillary predominant pattern. The bottom row shows an image with solid predominant pattern. The highlighted regions are shown in column b.

References

1. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778. IEEE (jun 2016)
2. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. CoRR abs/1412.6980 (2014), http://arxiv.org/abs/1412.6980

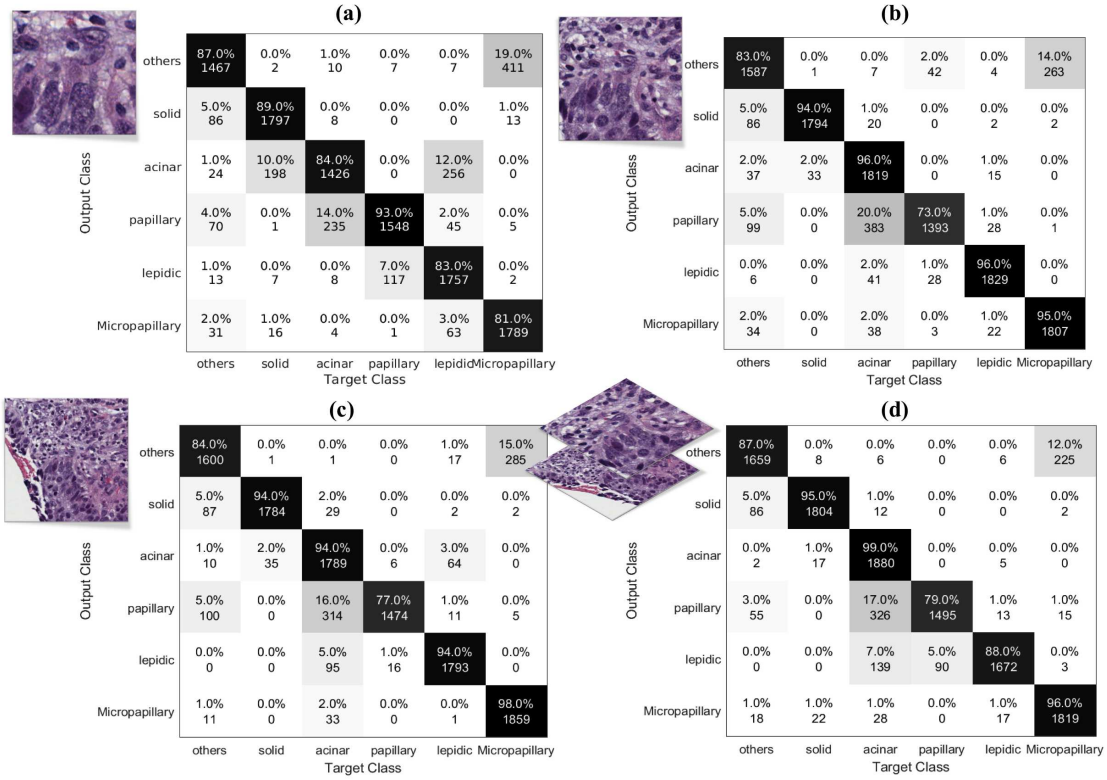

Fig. 5. (a, b, c, and d) shows the confusion matrix for ResNet architecture trained on four different magnifications 40×, 20×, 10×, and images from 20× and 10× magnifications; respectively. Highest misclassification is highlighted with red circle. A sample training image is also shown on the top left corner of each confusion matrix.

3. Riquet, M., Foucault, C., Berna, P., Assouad, J., Dujon, A., Danel, C.: Prognostic value of histology in resected lung cancer with emphasis on the relevance of the adenocarcinoma subtyping. Ann. Thorac. Surg. 81(6), 1988–1995 (jun 2006)
4. Russell, P.A., Wainer, Z., Wright, G.M., et al.: Does lung adenocarcinoma subtype predict patient survival?: A clinicopathologic study based on the new International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary lung adeno. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer 6(9), 1496–504 (2011)
5. Russell, P.A., Wright, G.M.: Predominant histologic subtype in lung adenocarcinoma predicts benefit from adjuvant chemotherapy in completely resected patients: discovery of a holy grail? Annals of translational medicine 4(1), 16 (2016)
6. Solis, L.M., Behrens, C., et al.: Histologic patterns and molecular characteristics of lung adenocarcinoma associated with clinical outcome. Cancer 118(11), 2889–99 (2012)

Multi-Resolution CNN for Lung Growth Pattern Classification 9

7. Thunnissen, E., Beasley, M.B., Borczuk, A.C., et al.: Reproducibility of histopathological subtypes and invasion in pulmonary adenocarcinoma. An international interobserver study. Modern Pathology 25(12), 1574–1583 (2012)
8. von der Thüsen, J.H., Tham, Y.S., et al.: Prognostic significance of predominant histologic pattern and nuclear grade in resected adenocarcinoma of the lung: potential parameters for a grading system. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer 8(1), 37–44 (2013)
9. Travis, W.D., Brambilla, E., Nicholson, A.G., et al.: The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer 10(9), 1243–60 (sep 2015)
10. Travis, W.D., Brambilla, E., Noguchi, M., et al.: International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society International Multidisciplinary Classification of Lung Adenocarcinoma. Journal of Thoracic Oncology 6(2), 244–285 (2011)
11. Van Schil, P.E., Sihoe, A.D., et al.: Pathologic classification of adenocarcinoma of lung. Journal of Surgical Oncology 108(5), 320–326 (2013)
12. Warth, A., Stenzinger, A., von Brünneck, A.C., Goeppert, B., Cortis, J., Petersen, I., Hoffmann, H., Schnabel, P.A., Weichert, W.: Interobserver variability in the application of the novel IASLC/ATS/ERS classification for pulmonary adenocarcinomas. European Respiratory Journal 40(5), 1221–1227 (2012)
13. Zugazagoitia, J., Enguita, A.B., Nuñez, J.A., Iglesias, L., Ponce, S.: The new IASLC/ATS/ERS lung adenocarcinoma classification from a clinical perspective: current concepts and future prospects (2014)