

MetaLung: Meticulous affine-transformation-based lung cancer augmentation method

Diana Nam¹, Alexandra Panina², Alexandr Pak¹, Fuad Hajiyev³

¹School of Information Technology and Engineering, Kazakh-British Technical University (KBTU), Almaty, Kazakhstan

²Department of Radiology and Nuclear Medicine, Kazakh Research Institute of Oncology and Radiology, Almaty, Kazakhstan

³School of Information Technologies and Engineering, ADA University, Baku, Azerbaijan

Article Info

Article history:

Received Feb 6, 2024

Revised May 17, 2024

Accepted Jun 5, 2024

Keywords:

Affine transformation

Data augmentation

Image segmentation

Lung cancer

Medical image processing

ABSTRACT

The limitation of medical image data in open source is a big challenge for medical image processing. Medical data is closed because of confidential and ethical issues, also manual labeling of medical data is an expensive process. We propose a new augmentation method named MetaLung (Meticulous affine-transformation-based lung cancer augmentation method) for lung CT image augmentation. The key feature of the proposed method is the ability to expand the training dataset while preserving clinical and instrumental features. MetaLung shows a stable increase in image segmentation quality for three CNN-based models with different computational complexity (U-Net, DeepLabV3, and MaskRCNN). Also, the method allows in reduce the number of False Positive predictions.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Alexandra Panina

Department of Radiology and Nuclear Medicine, Kazakh Research Institute of Oncology and Radiology
Almaty, Kazakhstan

Email: doctorpanina@gmail.com

1. INTRODUCTION

Lung cancer is one of the most lethal forms of cancer, contributing significantly to global cancer-related mortality. Lung cancer's severity often results from late-stage diagnoses, as the disease may not manifest symptoms until it has progressed. Despite advances in treatment modalities, the overall prognosis for lung cancer remains challenging, underscoring the importance of prevention, early detection, and ongoing research efforts to improve outcomes for those affected. There is a significant problem with the timely classification of lung cancer and subsequent patient management. The application of artificial intelligence in tasks related to lung cancer diagnosis can significantly improve indicators and enhance survival through timely diagnosis. The confidentiality of medical data poses a major barrier to its public domain use. Additionally, it is crucial to note that data labeling, an essential step in training artificial intelligence models, is a costly process that demands the expertise of qualified clinicians.

Computer vision engineers apply data augmentation techniques to increase the dataset's size and variability, avoiding overfitting of neural networks. There are two main approaches to generating new images to improve the quality of image processing: the application of affine transformation to make changes in existing images and drawing new images. Generative adversarial networks (GAN) [1] are a powerful tool for new instance generation. GANs excel at generating realistic data by training a generator network to create data instances that are indistinguishable from real ones by a discriminator network. They were applied for the generation of computed tomography images of liver [2] and lung cancer [3]. Toda *et al.* [3] have also applied GANs to increase the quality of lung cancer detection. The authors of LCDAE: data augmented ensemble

framework for lung cancer classification [4] proposed a hybrid method that allows to increase in the dataset of histopathological lung images [5] and uses them for multiclass classification between three types of lung cancer: adenocarcinoma, benign, squamous cell carcinoma. The LCDAE includes free main parts. The first one is generative model LDCGAN (LDCGAN). LSCGAN is standard GAN architecture, the generator network transforms random vectors into images, while the discriminator assesses the authenticity of generated and real data through adversarial training. The application of GAN allows to increase the dataset from 15,000 to 45,000 images. 80% of all images are saved for training and validation steps, while the remaining 20% is fed to data augmented ensemble model DA-END (DA-ENM). DA-ENM is a combination of affine transformation, such as random horizontal flip, erosion, contrast, brightness, saturation, and rotation. The last step of DA-ENM is data normalization, which is necessary for the correct training of the neural network. The last part of hybrid data augmentation (HDA). HAD is the combination of the first two parts, which allows receiving a significant increase in image classification. The quality of LCDAE for lung cancer classification was proved with 5 convolutional neural networks (DenseNet121 [6], GoogleNet [7], ResNet101 [8], VGG19_BN, VGG16_BN, VGG16 [9]) at least by 4.94% for VGG19_BN and maximum by 19.96% for GoogleNet. The histopathological image received by biopsy is not recommended for all potential lung cancer patients. The preliminary step is computed tomography screening. The authors of the next observed article [3] proposed an application of semi-conditional InfoGAN [10] for lung CT generation. The authors used the center of each lesion as an input. The authors first extracted volume of interest (VOI) as a preprocessing step. The VOI is a cubic region with sides equal to twice the diameter of the lesion, aiming to encompass information from the neighboring structures while constraining the scope of the analysis area. Then the center slides of each VOI were collected in the dataset. The dataset of original images consists of 7644 2-dimensional CT scans with 3 possible types of cancer (Adenocarcinoma, Squamous Cell Carcinoma, Small Cell Lung Cancer, 2,548 scans per class) provided by Fujita Health University Hospital. The dataset is closed because of ethical reasons. The authors of InfoGAN proposed an application of latent code instead of random vectors as an input. It involves incorporating additional variables into the generator's space, allowing the model to learn disentangled and interpretable features. These latent codes influence the generation process, enabling the generator to produce outputs where specific codes correspond to meaningful and understandable characteristics. An application of InfoGAN allows for increasing average accuracy for the classification of lung cancer lesions for 3 lung cancer types (Adenocarcinoma, Squamous Cell Carcinoma, Small Cell Lung Cancer) from 34.2% to 57.7%. Also, the authors provided a comparison of InfoGAN with Wasserstein GAN (WGAN) [11] with an average accuracy of 54.7%.

Although GANs are potent instruments for new instance generation, their application presents several challenges. Firstly, many images in the training set are required for effective GAN training. Secondly, GANs demand significant computational power as they involve two deep networks internally. Also, it is difficult to control and evaluate the resulting images of GAN. Especially for medical data generation because the parameters of generated instances have a direct effect on the final disease. In the medical field is the distribution between different types of instances on the medical image. Hounsfield Units in lung CT scans measure the radiodensity of tissues, aiding in the differentiation between air, lung parenchyma, and pathological conditions based on their attenuation characteristics. On the lung computer tomography image, the cancer has a value near 0 in Hounsfield Units, while the calcinates with the same form and location have a value significantly higher. So, for the control of lung cancer augmentation, it is significantly important to control this distribution, because its breaking could cause the incorrect training of the model when features of the disease could be extracted incorrectly. The main problem with this is that it is difficult to catch before real-life implementation, so it is better to prevent it. Also, as evident from the observation, lung cancer classification quality could be significantly improved by the application of data augmentation techniques. However, image classification provides the type of illness. By the use of image segmentation, we can check not only the class of the possible illness but also the location and radiological features of the affected by cancer area.

We proposed a novel lung cancer augmentation method MetaLung (Meticulous affine-transformation-based lung cancer augmentation method) which could be applied for lung cancer segmentation. The method is constructed based on the combination of six affine transformation techniques: mirroring the image, replacing cancer for free space, rotating the image, rotating cancer, adding noise to image, and adding noise to cancer which were preliminary tested with the same convolutional-based models. An application of affine transformation allows to increase in the size and variability of the train set at least by 5 times. While usual affine transformation allows image changes, without changing the position of cancer among other lung instances, we also proposed replacing cancer for free space of the lung, as a sub-method of MetaLung, which allows the generation of new images with different locations of lung cancer among other lung cancer instances. Table 1 provides the main advantages of MetaLung augmentation method.

Table 1. Advantages of MetaLung

Advantage	Description
1 Saving the distribution between different instances on CT images.	CT image contains instances of different types: lung area, bones, and fats. Some of them, for example, cancer and calcinates, could have similar forms, however the density is different. The mixture between them could cause incorrect diagnostics in the future.
2 Low computational complexity.	Computed aided diagnostics should reduce the burden on the medical sector. The problem of timely medical diagnosis is most common in developing countries, in which the economic component is a crucial factor. Therefore, it is necessary to use less expensive models and methods.
3 Reducing false positive results.	Processing computed tomography images in the case of pulmonary patients is an intermediate step before prescribing a biopsy for the final diagnosis of cancer. However, a biopsy is an invasive procedure that can cause harm to the patient, so its use in cases where there is no need should be minimized. An application of MetaLung shows a stable increase in precision for all observed models.
4 Increasing in size and variability in train set	An application of MetaLung shows the stable increase in DICE and intersection over union (IoU) for all observed models.

To prove the stability of MetaLung for lung cancer segmentation tasks we specifically collect the dataset of Kazakhstani local patients. The dataset contains 972 labeled CT scans with corresponding cancer masks in DICOM format. Then we trained three convolutional neural networks (U-Net [12], DeepLab V3 [13], Mask RCNN [14]) with different computational complexities and the number of trainable parameters for lung cancer segmentation task with and without the application of MetaLung. We used 4 metrics for the evaluation of image segmentation: DICE, IoU, Precision, and Recall. MetaLung shows a stable increase in segmentation quality for all of the observed deep-learning models. We also opened in public the data set which combines local Kazakhstani data with LIDC-IDRI. The dataset could be used for lung cancer classification and detection models [15].

2. METALUNG: METICULOUS AFFINE-TRANSFORMATION-BASED LUNG CANCER AUGMENTATION METHOD

There are various methods to augment the training dataset, such as GANs or affine transformations. We do not use GANs due to several difficulties:

- GANs require a diverse training dataset
- A lot of computational resources
- Medical professionals should additionally verify the generated images

CT images contain a lot of instances, such as lung area, bones, and fats. Also, there are air, water, and instances. The standard format for CT images is DICOM. The density of each pixel calculated in Hounsfield units (HU) Figure 1 provides the distribution between different lung instances on CT images in HU. So the value of air on the DICOM image is approximately near -1000HU, Lung area near -500HU. Lung cancer has a similar form to calcinates, which are not dangerous for human beings. However, in the CT image, both instances have serious differences in HU value.

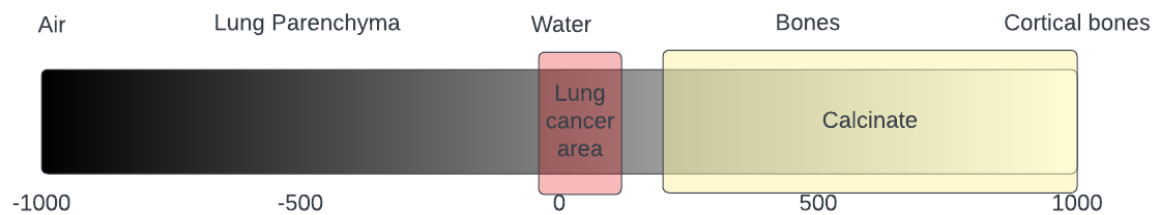


Figure 1. The distribution between lung instances

Saving the distribution between instances on CT images is crucial for new CT image generation because a mixture between them could mix values on different instances. We combined 6 affine transformation methods, including mirroring the image, replacing cancer for free space, rotating the image, rotating cancer, adding noise to the image, and noise to cancer, with random parameters that do not break the distribution in the density of lung instances on CT image. The description of affine transformation methods combined in MetaLung is provided in Table 2.

Table 2. Affine transformation methods used for comparative analysis

N	Method name	Description	Random parameters
1	Mirroring the image	All CT images are provided in [512, 512] format. The image is mirror-flipped according to the vertical central line with the x-axis 128.	n/a
2	Replacing cancer for free space	Lung and vessels on CT images are preliminarily segmented via a thresholding algorithm. It allows us to get information about the free space from cancer and the vessels area of the lung. The cancer is flipped from the central point to 180 degrees and moved to the free area of the opposite side of the lung if it existed. Only in case there is no free space on the opposite part, the cancer is moved to the same side of the lung. An example of the application of the second method is provided in Figure 2. This method is not applicable to the image in case there is not enough free space to move the cancer because of vessels, calcinates, or other lung structures.	n/a
3	Rotating the image	The whole image is rotated to random angle from 0 to 90.	$angle \in [0,90]$
4	Rotating cancer	The cancer area is rotated to a random angle from 0 to 90. Other parts of the image are not changed and saved as it was.	$angle \in [0,90]$
5	Adding noise to the image	The noise with the random value from 0 to 20 is summarized in the whole image. We do not use a big value for noise generation because all instances (lung, bone, cancer, vessels) have fixed value in HU on the CT image. So, we could not apply extreme changes to the image because we want to save this distribution on the image. The same values are used for Method 6.	$noise \in [0,20]$
6	Adding noise to cancer	The noise with the random value from 0 to 20 is summarized to the cancer area only.	$noise \in [0,20]$

MetaLung augmentation method allows to generate at least 5 images from 1. The method used as an input CT image in HU with a corresponding binary mask with lung cancer. The method operates with 2D CT slices with standard side [512, 512]. An output of MetaLung is the set of new CT images with new masks. Adding noise to the image and adding noise to cancer do not require the generation of a new mask, because the location of the cancer is not changed. Replace cancer to free space is the novel method used in MetaLung. It allows the generation of new CT images with different locations of lung cancer among other lung CT instances. However, this method does not apply to all cases, because it replaced the cancer with a space without any vessels, that could not exist. The pseudocode of the algorithm is provided above. Also, the application and visualization of the MetaLung Method for one CT example is available by the link [16].

Algorithm 1. MetaLung (CT image, cancer mask)

```

FUNCTION Mirroring the image (CT image, cancer mask)
    new CT image = mirror flip (CT image)
    new cancer Mask = mirror flip (cancer mask)
    RETURN new CT image, new cancer mask

FUNCTION Replacing cancer for free space (CT image, cancer mask)
    RoI = CT image * cancer mask
    IF (check vessels on the opposite side == True)
        new CT image = replace RoI with free part
        generate new cancer mask
        RETURN new CT image, new cancer mask
    ELSE IF (check vessels on the same side == True)
        new CT image = replace RoI with free part
        generate new cancer mask
        RETURN new CT image, new cancer mask

FUNCTION Rotating the image (CT image, cancer mask)
    angle = random (min = 1, max = 90)
    new CT image = rotate (CT image, angle)
    new cancer mask = rotate (cancer mask, angle)
    RETURN new CT image, new cancer mask

FUNCTION Rotating the cancer (CT image, cancer mask)
    angle = random (min = 1, max = 90)
    new cancer mask = rotate(cancer mask, angle)
    old RoI = CT image * cancer mask
    new RoI = flip (CT image, angle) * new mask
    new Ct image = replace(old Roi, new Roi)
    RETURN new CT image, new cancer mask

FUNCTION Adding noise to the image (CT image, cancer mask)
    noise = random array (min = 0, max = 20, size = [512,512])
    new CT image = noise + CT image
    RETURN new CT image, cancer mask

FUNCTION Adding noise to cancer (CT image, cancer mask)
    noise = random array (min = 0, max = 20, size = [512,512])
    new CT image = noise * cancer mask + CT image
    RETURN new CT image, cancer mask

```

3. METHOD

We proposed a novel method for lung cancer augmentation MetaLung Meticulous affine transformation-based lung cancer augmentation method. The method is applicable for increasing the size and variability of a training set of computed tomography images with lung cancer. MetaLung could be used to increase the quality of training models for segmentation. To check the stability of the proposed method we used three convolutional-based models for lung cancer segmentation which were trained with a similar distribution between train and test sets. The pipeline of the lung segmentation framework is shown in Figure 2. The results of training models without data augmentation were used as baseline results. Then we applied MetaLung to increase the size of the train set and retrain the models. We used the same hardware and test set for baseline models and models with MetaLung augmentation. All of the provided steps have been provided [16].

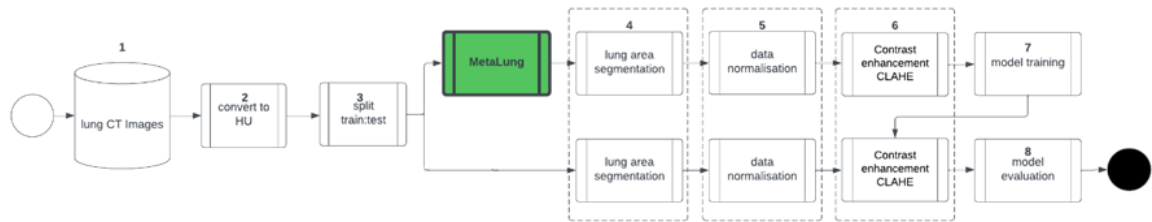


Figure 2. Pipeline of lung cancer segmentation framework

3.1. Data collection

We collected data from Kazakhstan lung cancer patients at various stages. The data was provided by the Kazakh Research Institute of Oncology and Radiology [17]. All data was anonymized before the usage. Due to a lack of data, we supplemented it with information from the publicly available LIDC – IDRI dataset [18]. All the data was annotated in a specific manner. Since we decided to tackle the segmentation task, a medical professional manually outlined the polygons containing cancer in the images. Additionally, each polygon was assigned a value according to lung-RADS.

Lung-RADS is a system used for categorizing and reporting lung cancer screening results. There are four classes which depend on the size of the affected area. The general information about the lung-RAGS System and the distribution between the data in the dataset are provided in Table 3. As it is shown in Table 3, LR4B has the biggest size of the nodule and LR2 is the smallest one correspondingly. Because LR4A and LR4B have a large size is it easier to diagnose them. So, the classification among Lung-RADS would be used in model evaluation.

Table 3. Description and number of samples according to Lung-RADS system

Class according to Lung RADS	Size	Recommendations	Number of images in the dataset
LR2	4-6 mm	1 year control	142
LR3	6-8 mm	6 month control	138
LR4A	Less than 1.5 mm	3 month control	177
LR4B	more than 1.5 mm	biopsy is recommended	515

The dataset consists of 972 images from 71 patients. The dataset contains of CT slice, serial number, binary mask with cancer area, corresponding lung RADS class, and patient ID. All data was anonymized because of ethical considerations. CT images were provided in DICOM format. CT images were taken on four possible equipment: SIEMENS, GE MEDICAL SYSTEMS, TOSHIBA, UIH.

3.2. Data splitting

The original dataset with all CT images from each patient contains information about unique patient ID. We split the data based on the number of patients by the way when all labeled CT images with lung cancer could belong to a train or a test set only. The data also was labeled according to the Lung-RADS System as well with 4 possible classes (LR2, LR3, LR4A, LR4B). We used 4 patients for each class for the test set and the remaining 55 for to train set. We saved only CT slices with lung cancer and continue to work in 2D space. We did not take into account the patient ID for the next experiments and worked with each slide as with a separate image. The train set consists of 708 images and corresponding masks, and the test set contains 264. The dataset could be available [15]. Exploratory data analysis of test set could be available [16].

3.3. Converting the image to hounsfield unit

All CT images have been provided in DICOM format. DICOM format allows to contain information with a maximum pixel value much bigger than [0, 255], so it contains information about the different types of instances on computed tomography images. HU in DICOM represent the quantitative measure of radiodensity in a medical image, allowing for the standardized characterization of tissue attenuation based on X-ray attenuation coefficients. CT images were taken on four possible equipment: SIEMENS, GE MEDICAL SYSTEMS, TOSHIBA, UIH. The data is presented in DICOM format of size [512,512]. Since the CT images were collected using various medical equipment, they are not standardized by default. So all CT scans should be converted to Hounsfield Units for further calculations as in (1).

$$HU = \frac{\mu_X - \mu_{water}}{\mu_{water} - \mu_{air}} * 1000 \tag{1}$$

Where μ_{water} and μ_{air} are linear attenuation coefficients for water and air under standard conditions.

Figure 3 shows the changes in density distribution from the original DICOM image after converting HU. Figure 3(a) shows the histogram of the density distribution of one CT image before converting the image to HU. Figure 3(b) shows the density distribution after standardization the DICOM image to HU. As it is shown in Figure 3(a), the main part of the image has the value -1000HU which belongs to air.

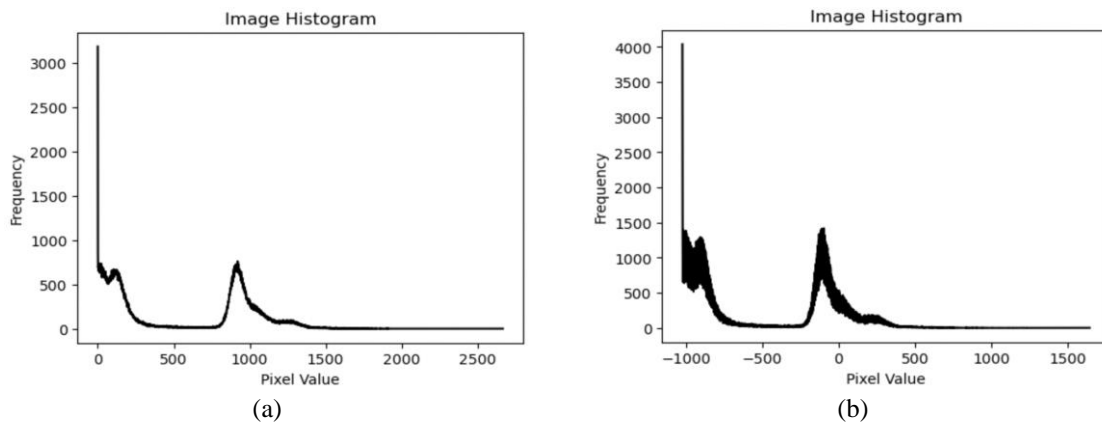


Figure 3. Histogram of DICOM image (a) before converting to HU and (b) after converting to HU

3.4. Lung segmentation

To increase the result of model training thresholding-based segmentation was applied to train and test set both to delete all non-related to the lung parts. It should be mentioned that this step is optional and could be skipped. Also, we did not evaluate the quality of lung segmentation.

Although external factors could influence the CT image and lead to noise appearance, the lung has a very big distance from the bone area in HU. This factor allows segmentation of the lung area via the thresholding method without additional application of neural network to reduce computation complexity. The block scheme of lung segmentation is provided in Figure 4. As the received binary mask has the same size as the input image it could be multiplied with CT to get the lung area only. Figure 5 shows the lung segmentation by the thresholding method. The result of the lung segmentation algorithm from the original CT image shown in Figure 5(a) is provided in Figure 5(b).

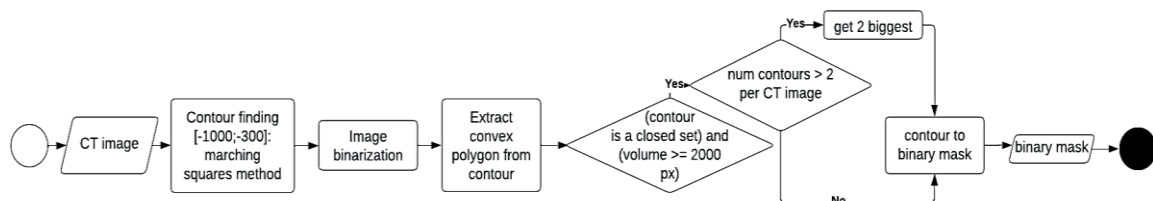


Figure 4. Block scheme of the lung segmentation method

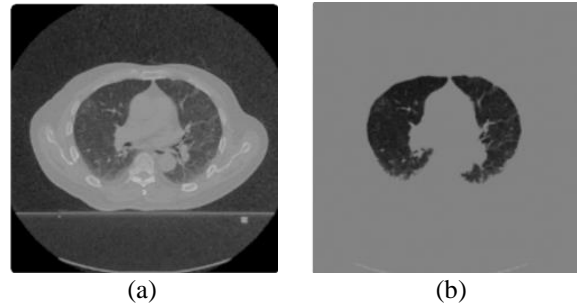


Figure 5. CT images (a) original and (b) after lung segmentation

3.5. Data normalization

Data normalization is a crucial preprocessing step for deep neural networks involving the scaling of input data to a standard range, often between 0 and 1. Normalization helps CNNs converge faster during training, mitigates issues related to varying input scales, and promotes better generalization by ensuring consistent input representations across the dataset. For all of the next experiments (2) was used for image normalization.

$$\text{Norm}(img) = \frac{img - \min(img)}{\max(img) - \min(img)} \quad (2)$$

Where img is the original image.

3.6. Contrast limited adaptive histogram equalization CLAHE

Contrast limited adaptive histogram equalization (CLAHE) [19] is a widely used image enhancement technique in medical image processing. Application of CLAHE shows an increase of lung illness classification based on CT and X-ray images [20]. It addresses the challenge of uneven illumination and varying contrast levels present in medical images. By adaptively dividing the image into smaller regions and applying histogram equalization independently to each, CLAHE avoids over-amplification of noise while enhancing local contrast. This makes it particularly valuable for highlighting subtle details in medical images, aiding clinicians in diagnosis. The non-linear nature of CLAHE makes it well-suited for scenarios where different regions of an image may require distinct contrast adjustments, enhancing the overall interpretability of medical imagery. CLAHE could be described by (3).

$$CLAHE(img) = \text{Interp}(\text{ClipHistEq}(img)) \quad (3)$$

Where img is an input image.

ClipHistEq involves performing histogram equalization on the image, but with a constraint on the contrast amplification to avoid over-enhancement.

Interp represents an interpolation function.

3.7. Lung cancer segmentation models

3.7.1. MASK RCNN for lung cancer segmentation

Mask region-based convolutional neural network (R-CNN) is a deep learning architecture primarily designed for object detection. It extends the Faster R-CNN [21] framework by incorporating an additional branch for predicting segmentation masks alongside bounding box coordinates and class labels. The architecture consists of three main components: a backbone network (feature extractor like ResNet [8]), a region proposal network (RPN) [21] for generating candidate object proposals, and the segmentation mask branch responsible for fine-grained pixel-level predictions. Mask R-CNN utilizes a two-stage process, first generating region proposals and then refining them for precise object localization and segmentation. The architecture's flexibility allows it to handle tasks such as object detection, instance segmentation, and keypoint estimation, making it a versatile solution for a variety of computer vision applications. MASK RCNN shows state-of-the-art results for different medical image processing tasks [22].

3.7.2. U-Net

U-Net is a convolutional neural network architecture designed for semantic segmentation tasks in medical image processing. Its distinctive U-shaped architecture consists of a contracting path to capture context and a symmetric expansive path to enable precise localization. The model's skip connections facilitate

the fusion of low-level features with high-level contextual information, aiding in capturing fine details. The network's ability to handle limited data through data augmentation and transfer learning further contributes to its popularity in various image segmentation applications. U-Net is widely employed in medical image processing when the ability of the neural network to train on a small dataset is significantly important [23].

3.7.3. DeepLab V3

DeepLabv3 is a state-of-the-art semantic segmentation architecture designed for image segmentation. The network used atrous (dilated) convolutions to effectively capture multi-scale contextual information without increasing computational complexity. DeepLabv3 further integrates a dilated spatial pyramid pooling module, enabling the model to capture context at multiple scales. Finally, it employs a decoder module to refine segmentation predictions, producing highly detailed and accurate segmentations. DeepLabv3 has demonstrated exceptional performance in various computer vision tasks [24].

We used CNN-based models with different computational complexity and architecture features to check the stability of MetaLung in the different circumstances. The number of trainable parameters is responsible for computational complexity of the neural network. The number of trainable parameters for observed models is shown in Table 4.

Table 4. Number of trainable parameters for observed models

Model	Number of parameters
U-Net	4,39E+07
DeepLab V3	1,82E+07
Mask RCNN	3,96E+07

3.8. Model evaluation

We used 4 metrics for evaluation of the quality of image segmentation. DICE, as in (4) and, IoU, as in (5) have been used for evaluation the similarity between ground true and predicted masks. Image segmentation could be described as pixel-wise segmentation problem. It allows to calculate precision, as in (6) and recall, as in (7) to evaluate the distribution between false positive and false negative predictions among one predicted mask. Precision focuses on minimizing false positives, measuring the accuracy of positive predictions among those predicted as positive. Recall, on the other hand, aims to minimize false negatives, assessing the model's ability to capture all actual positive instances by minimizing the instances incorrectly predicted as negative. In the context of CT image diagnostics for lung cancer detection, precision is crucial as it helps minimize false positive predictions. Reducing false positives is particularly essential because an erroneous prediction may lead to unnecessary and invasive procedures such as biopsy, which can be traumatic and highly invasive.

$$\text{DICE} = \frac{2 * |X \cap Y|}{|X| + |Y|} \quad (4)$$

$$\text{IoU} = \frac{|X \cap Y|}{|X \cup Y|} \quad (5)$$

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (6)$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (7)$$

3.9. Hardware

We used the same equipment for experiments with all baseline and models with an application of MetaLung. All experiments have been done on: NVIDIA A100 80GB GPU with CUDA version 11.7, paired with an AMD EPYC 7663 56-Core Processor for CPU tasks, RAM 1,5 Ti. The hardware equipment has been temporary provided by Institute of Information and Computational Technologies (IICT), Almaty, Kazakhstan.

4. RESULTS AND DISCUSSION

4.1. Visual explanation of MetaLung augmentation method for single lung cancer CT image

An application of MetaLung allows to generate 5 or 6 new images from one. An input of MetaLung is an original CT image with a corresponding binary mask with lung cancer. Figure 6 provides an input CT image of MetaLung. Figure 7 shows the ground true mask with labeled cancer area. Figure 8 shows the labeled

part in the CT image for better visualization. Figures 9 to 20 show the transformation of the original image after the application of MetaLung. Figures 9 to 11 show changes in the original CT image and ground true mask after Mirroring the image. Figures 12 to 14 show resulted data after the application Replacing cancer to free space. Figures 15 to 17 show updated data after image rotation. Figures 18 to 20 show changes after cancer rotation.

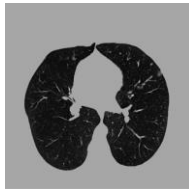


Figure 6. Original CT image



Figure 7. Original ground true mask



Figure 8. Labeled cancer area



Figure 9. CT image after mirroring the image

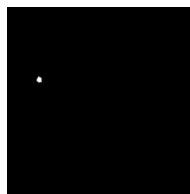


Figure 10. Mask after mirroring the image



Figure 11. Labeled area after mirroring the image

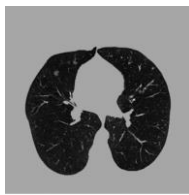


Figure 12. CT image after replacing cancer for free space



Figure 13. Mask after replacing cancer for free space



Figure 14. Labeled area after replacing cancer for free space

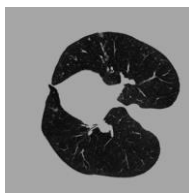


Figure 15. CT image after rotating image

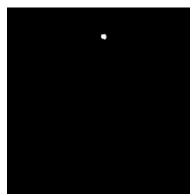


Figure 16. Mask after rotating the image

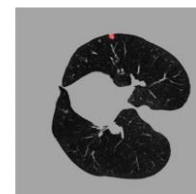


Figure 17. Labeled area after rotating the image

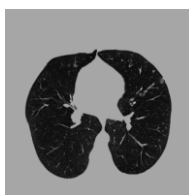


Figure 18. CT image after rotating cancer

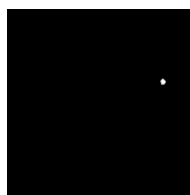


Figure 19. Mask after replacing rotate cancer



Figure 20. Labeled area after rotating cancer

We visualize the CT image, mask, and labeled area to mirroring the image, replacing cancer for free space, rotating the image, and rotating cancer because it allows to catch the changes in lung cancer location in 2-dimensional space. Also new mask has been generated for all of the provided methods. However, we used histogram visualization for adding noise to the image and adding noise to cancer, because it is impossible for the human eye to check the small changes in density change with the random value from 0 to 20. An application of adding noise to the image is shown in Figure 21. Figure 21(a) shows the histogram of the original CT image. Figure 21(b) shows the histogram of a new CT image after the application of Noise to the image. An application of adding noise to the cancer area is shown in Figure 22. Figure 22(a) provides the histogram of the lung cancer area only. Figure 22(b) provides the histogram of the lung cancer area after the application of noise to cancer.

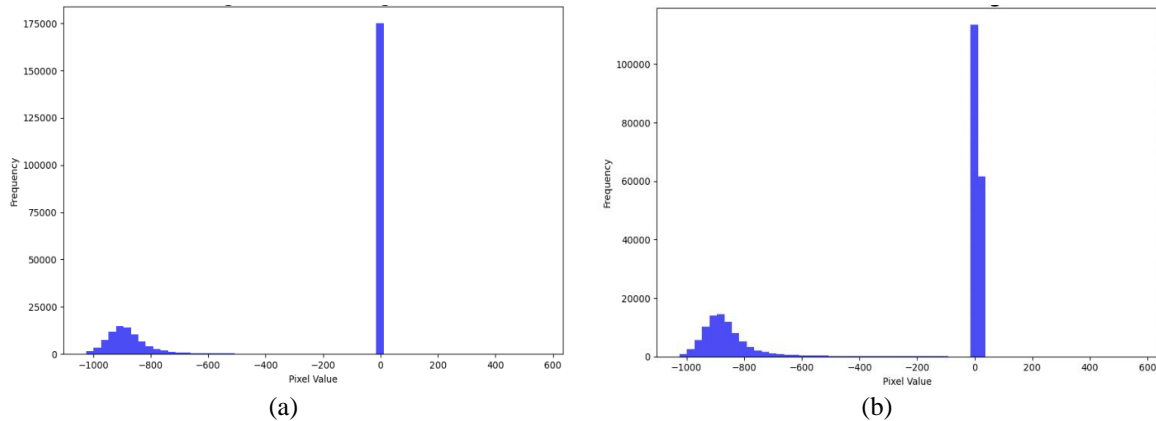


Figure 21. Histogram of CT image (a) original and (b) CT image after adding noise to the image

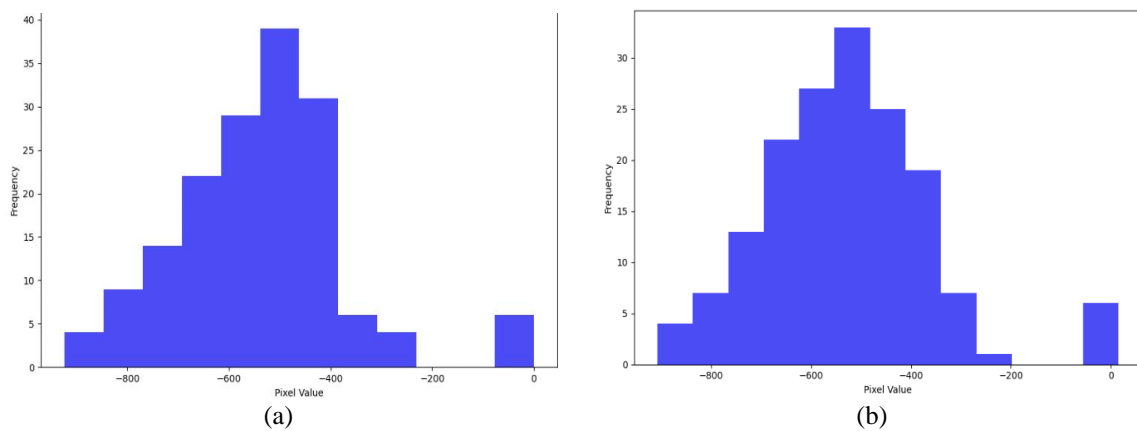


Figure 22. Histogram of cancer area only (a) original CT image and (b) after adding noise to cancer

4.2. Application of MetaLung to increase the size and variability of dataset with lung cancer CT images

An application of the MetaLung allows us to increase the train set from 708 to 4,865, which is 6.87 times more than in the original set. This number also could be increased by the combination and permutation of affine transformations with each other. Figure 23 shows the distribution among cancer locations and the number of changes because of increasing the number of images and corresponding masks. We summarized all masks of the original image and constructed a heatmap based on the received results and received heatmap shown in Figure 23(a). Then we applied the same to the dataset updated by MetaLung. The distribution is shown in Figure 23(b). Also, as it is shown in Figure 23(b), an application of MetaLung allows for to equalization of the distribution between cancers in the left and right lungs in the training dataset, thus making the training set more balanced and increasing the diversity. The increase in the size and diversity of training data leads to an increase in the quality of medical image processing.

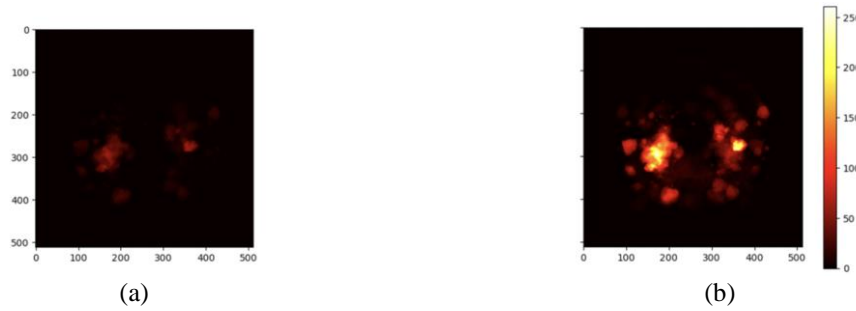


Figure 23. Heatmap of cancer distribution (a) original train set and (b) extended train set by MetaLung application

4.3. Impact of application of MetaLung for lung cancer segmentation

First, we trained three convolutional-based models (mask RCNN, U-Net, and Deeplab V3) without data augmentation for comparison reasons. We used these results as a baseline for the next experiments. Then we applied seven augmentation techniques including the MetaLung to generate new images to increase the train set. We used 4 metrics for the evaluation of segmentation results: DICE, IoU, precision, and recall. We used DICE and IoU to evaluate the similarity between ground true and predicted masks. And precision and recall for evaluation of the distribution between false positive and false negative predictions. All weights for models could be found [25]. The results for baseline models and models with the application of MetaLung are shown in Table 5. Also, we calculated DICE separately for all of the lung-RADS classes to evaluate which of the classes is the most difficult for diagnosis. The DICE score calculated for each lung-RADS class is shown in Table 6.

Table 5. Average results if metrics for lung cancer segmentation

Model	DICE		IoU		Precision		Recall	
	Baseline	MetaLung	Baseline	MetaLung	Baseline	MetaLung	Baseline	MetaLung
U-Net	0.3708	0.4004	0.312	0.3418	0.8175	0.8676	0.359	0.3684
DeepLab V3	0.3308	0.4136	0.269	0.3498	0.7906	0.8634	0.3177	0.3817
Mask RCNN	0.5117	0.6004	0.4609	0.5361	0.5372	0.6864	0.5093	0.5896

Table 6. DICE metrics for lung cancer segmentation among Lung RADS classes

Model	LR2		LR3		LR4A		LR4B	
	Baseline	MetaLung	Baseline	MetaLung	Baseline	MetaLung	Baseline	MetaLung
U-Net	0.1997	0.1346	0.365	0.36	0.1979	0.3655	0.6127	0.5497
DeepLab V3	0.0645	0.0458	0.3417	0.3889	0.1773	0.3971	0.5813	0.5755
Mask RCNN	0.0	0.2665	0.2259	0.4032	0.7226	0.7501	0.5961	0.6475

4.4. Discussion

Although an application of CAD systems can significantly increase the speed and the quality of lung cancer detection, the problem of data deficit presents a significant barrier to realizing the full potential of these systems. Traditional data augmentation approaches, such as GAN or affine transformations allow a significant increase in the size and the variability of the dataset for the next deep learning model training. However, they do not take into account specific radiological features of lung cancer, such as the density in HU, which does not allow to fully control the process of decision making of neural network.

We proposed a novel method MetaLung for lung cancer data augmentation, which takes into account the density distribution of different instances, such as lung area, fat, cancer, and calcinates, on lung CT images. An application of MetaLung allows to increase in the size and the variability of the train set and increases the quality of lung cancer segmentation, which were proven with three convolutional neural networks (U-Net, Mask RCNN, and DeepLab V3) with different architecture and the number of trainable parameters based on DICE, IoU, precision, and recall metrics in comparison with baseline models without data augmentation techniques. However, it should be noted that despite MetaLung showing high results on the entire dataset for all three considered models, the segmentation quality can be improved for individual classes according to Lung-RADS. Additionally, the quantity of generated images may be increased using the proposed method by combining the employed transformations. Also, the method is applicable for CT images

in DICOM format only, because it contains the information about the human organs in Hounsfield Units, while the standard format, such as PNG or JPG with RGB format, breaks the density distribution by itself.

Our study demonstrates the proposed MetaLung data augmentation method for lung cancer segmentation. We worked with positive cases with existing lung cancer only. We assume that the proposed method also is applicable for lung cancer classification on the CT image.

The use of augmentation methods for processing medical data is inherently complicated by the fact that it is necessary to strictly control the generated results, since they can have a further negative impact on the diagnosis, even if quality metrics show an increase. In solving problems of disease detection, data is usually collected for one type of disease or several, but the number of such diseases in real life is much higher. We proposed a novel data augmentation MetaLung which allows not only an increase in image segmentation quality, but saves the radiological differences between lung cancer and other instances presented on lung CT image.

5. CONCLUSION

We introduce MetaLung, a novel data augmentation technique designed to augment the size and enhance the variability of CT image datasets focusing on lung cancer. MetaLung combines previously unused affine transformations with a groundbreaking method that replaces cancerous regions with free space. Notably, our approach employs strict boundaries for all random transformations applied within MetaLung. The efficacy of MetaLung was evaluated using three CNN-based neural networks: U-Net, DeepLabV3, and Mask RCNN. Our method demonstrates significant improvements in DICE, IoU, precision, and recall metrics for lung cancer segmentation. Additionally, we have made our meticulously labeled dataset publicly available, specifically curated for lung cancer segmentation and classified in accordance with the lung reporting and data system (Lung-RADS). We posit that our newly proposed method holds promise for enhancing lung cancer classification tasks as well.

ACKNOWLEDGMENT

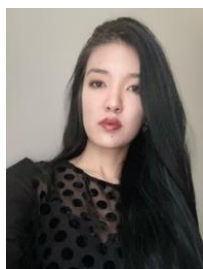
This work was supported by the Ministry of Education and Sciences of the Republic of Kazakhstan under the following grant #AP14871214. The funders HAD no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.




REFERENCES

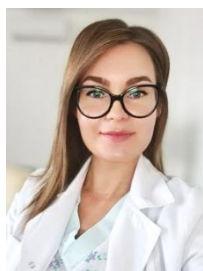
- [1] I. Goodfellow *et al.*, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.
- [2] Q. Xiao and L. Zhao, "End-to-end 3D liver CT image synthesis from vasculature using a multi-task conditional generative adversarial network," *Applied Sciences*, vol. 13, no. 11, p. 6784, Jun. 2023, doi: 10.3390/app13116784.
- [3] R. Toda *et al.*, "Synthetic CT image generation of shape-controlled lung cancer using semi-conditional InfoGAN and its applicability for type classification," *International Journal of Computer Assisted Radiology and Surgery*, vol. 16, no. 2, pp. 241–251, Feb. 2021, doi: 10.1007/s11548-021-02308-1.
- [4] Z. Ren, Y. Zhang, and S. Wang, "LCDAE: data augmented ensemble framework for lung cancer classification," *Technology in Cancer Research & Treatment*, vol. 21, p. 153303382211243, Jan. 2022, doi: 10.1177/15330338221124372.
- [5] A. A. Borkowski, M. M. Bui, L. B. Thomas, C. P. Wilson, L. A. DeLand, and S. M. Mastorides, "Lung and colon cancer histopathological image dataset (lc25000)," *arXiv preprint arXiv:1912.12142*, 2019.
- [6] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4700–4708, 2017.
- [7] C. Szegedy *et al.*, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.
- [8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
- [9] S. Karen, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv: 1409*, 2014.
- [10] X. Chen, Y. Duan, R. Houthoofd, J. Schulman, I. Sutskever, and P. Abbeel, "Infogan: Interpretable representation learning by information maximizing generative adversarial nets," in *Advances in Neural Information Processing Systems*, vol. 29, 2016.
- [11] Y. Onishi *et al.*, "Automated pulmonary nodule classification in computed tomography images using a deep convolutional neural network trained by generative adversarial networks," *BioMed Research International*, vol. 2019, pp. 1–9, Jan. 2019, doi: 10.1155/2019/6051939.
- [12] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," 2015, pp. 234–241.
- [13] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv preprint arXiv:1706.05587* 5, 2017.
- [14] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct. 2017, pp. 2980–2988, doi: 10.1109/ICCV.2017.322.
- [15] D. Nam, A. Panina, and A. Pak, "Lung cancer segmentation dataset with Lung-RADS class," *Mendeley Data*, 2024.
- [16] "MetaLung: Meticulous affine-transformation-based lung cancer augmentation method," [Online]. Available: <https://github.com/namdiana/MetaLung--data-augmentation-method-for-lung-cancer-segmentation/>.
- [17] "Kazakh research institute of oncology and radiology," [Online]. Available: <http://www.onco.kz/>.




- [18] S. G. Armato, "Data from LIDC-IDRI [Data set]," *The Cancer Imaging Archive*.
- [19] S. M. Pizer *et al.*, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355–368, Sep. 1987, doi: 10.1016/S0734-189X(87)80186-X.
- [20] A. N. Setty, R. T. Mathad, K. Shenthar, and L. Likhith, "Evaluation of filtering and contrast in X-ray and computerized tomography scan lung classification," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 33, no. 3, p. 1715, Mar. 2024, doi: 10.11591/ijeecs.v33.i3.pp1715-1725.
- [21] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.
- [22] A.-A. Nayan *et al.*, "A deep learning approach for brain tumor detection using magnetic resonance imaging," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 1, p. 1039, Feb. 2023, doi: 10.11591/ijece.v13i1.pp1039-1047.
- [23] A. I. Sapitri, S. Nurmaini, S. Sukemi, M. N. Rachmatullah, and A. Darmawahyuni, "Segmentation atrioventricular septal defect by using convolutional neural networks based on U-NET architecture," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 3, p. 553, Sep. 2021, doi: 10.11591/ijai.v10.i3.pp553-562.
- [24] W. Yujie, S. Simon, Y. Jahow, and D. L. Yu, "Skin lesion segmentation using atrous convolution via DeepLab v3," *arXiv preprint arXiv:1807.08891*, 2018.
- [25] D. Nam, A. Panina, and A. Pak, "Pre-trained weights for application of MetaLung (Meticulous affine-transformation-based lung cancer augmentation method) and baseline models," *zenodo*, 2024.

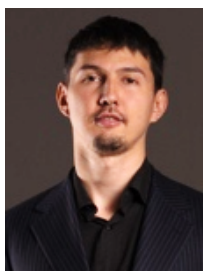
BIOGRAPHIES OF AUTHORS






Diana Nam    is a Ph.D. student and holds a Master's degree in Technical Sciences. She is a lecturer at the Kazakhstan-British Technical University, specializing in computer vision and deep learning. With expertise in machine learning and deep learning, she focuses on medical image processing. She can be contacted at email: d.nam@kbtu.kz.






Alexandra Panina    was born in Almaty, Kazakhstan, in 1989. She received her bachelor's degree in medicine from the Kazakh National Medical University in 2012 and completed her residency in radiology at the Kazakh Research Institute. Currently a candidate for the PhD degree in Medicine. Since 2018, she has been working as a radiologist in the Department of Radiology and Nuclear Medicine at the Kazakh Research Institute of Oncology and Radiology in Almaty, Kazakhstan. She is the author of 3 publications indexed in Scopus with an H-index of 3. Her research interests include early diagnosis of cancer. She can be contacted at email: doctorpanina@gmail.com.



Alexandr Pak    was born in Almaty city, Kazakhstan in 1983. He received his BS degree in Physics from Kazakh National University in 2005 and a Ph.D. degree specializing in 'Applied Mathematics from the Institute of Mathematics in 2010. He has been an associate professor at Kazakh-British Technical University (KBTU) since 2019. His Ph.D. thesis is 'The Computer Methods and algorithms of Data Analysis with a long-range dependency.' Since 2018, he has worked as a Head of the Laboratory at the Research Institute of Information and Computational Technologies, Committee of Science of the Ministry of Science and Education of the Republic of Kazakhstan, in Almaty, Kazakhstan. He is the author of 15+ publications indexed on Scopus with an h-index of 5. His research interests include natural language processing, computer vision, and deep learning. He can be contacted at email: a.pak@kbtu.kz.



Fuad Hajiyev    School of Information Technologies and Engineering, ADA University, Baku, Azerbaijan. He can be contacted at email: fhajiyev@ada.edu.az.