Chest X-ray Image Processing Based on Radiologists' Textual Annotations

Aleksandra Kosareva, Dzmitry Paulenka, Eduard Snezhko
The United Institute of Informatics Problems of the National Academy of Sciences of Belarus
Biomedical Image Analysis Department, https://image.org.by
Minsk, Belarus
Email: kosarevaaleksandra4317@gmail.com, dmitri.pavlenko@gmail.com,
eduard.snezhko@gmail.com

***Abstract*—More than 11,000 chest x-ray images and their corresponding text annotations were analyzed, and the first pilot studies on image processing tailored to text annotations of radiology specialists were conducted. An image processing pipeline for a database and for a neural network has been developed. The prediction of the parameter "Overall percent of abnormal volume" was performed and the mean absolute error (MAE) for the InceptionResNet50V2 neural network model was 11.073.**

***Keywords*—medical image processing, medical image analysis, deep learning, computer-aided diagnosis, chest x-ray, textual annotations of lung lesions**

I. Introduction

In this article the main efforts are made to analyze and prepare Chest X-ray (CXR) images and corresponding *text* data annotations. A total of *11,493* non-empty CXR images were downloaded (in fact there are 13,521 instances, however 2,028 of them were empty and not downloaded from TB Portals [1] website).

On the CASE BROWSER [2] website CXR text annotation can be viewed as in Fig. 1 along the following path:

Patient: № → Case → View Imaging Study → Diagnostic Report.

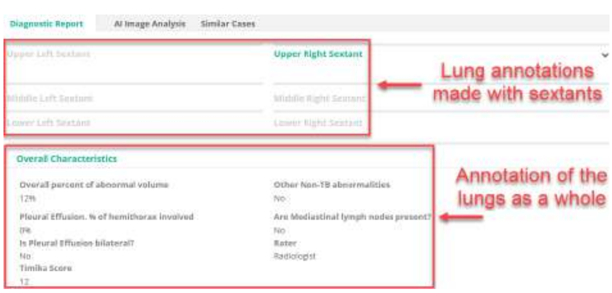

Figure 1. CXR textual annotation on the CASE BROWSER [2] website.

All CXR textual data in corresponding JSON files as well as on the CASE BROWSER [2] and TB DEPOT [3] websites are divided into three blocks of information (Fig. 2):

This work was carried out with the financial support of the ISTC-PR150 "Belarus TB Database and TB Portal" project.

- anonymized patient information (gender, age, country, diagnosis, etc.);
- CXR radiologists' textual annotations, *which is currently being analyzed* in this article;
- treatment history, which includes medications and treatment days.

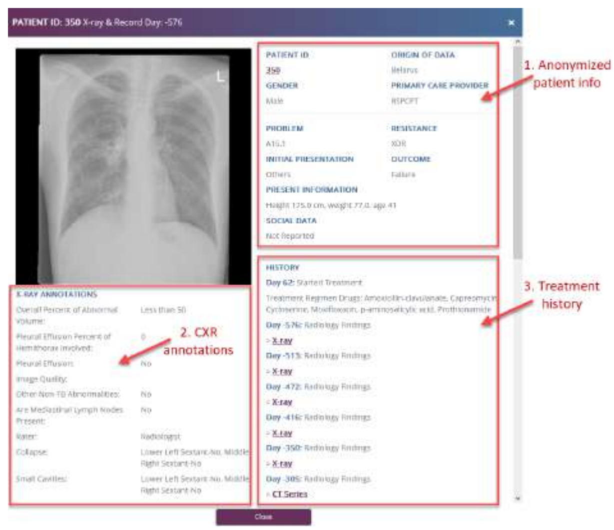

Figure 2. CXR case information on the TB DEPOT [3] website.

Second data block in Fig. 2 or "CXR annotations" can also be roughly divided into three groups:

- CXR annotations, as if it were a computed tomography (CT) scan, in total 106 images;
- CXR annotations of the six lobes of the lung (sextants) with twenty parameters for each sextant, in total 546,364 non-empty text annotations in 9,154 CXR images;
- CXR annotations of the lung as a whole using six parameters, see "Overall Characteristics" tab in Fig. 1, in total 11,387 CXR images.

For clarity, the image categories are shown in the diagram below, Fig. 3.

The results of the analyses of the annotations of these groups are summarized below. Further, after the data

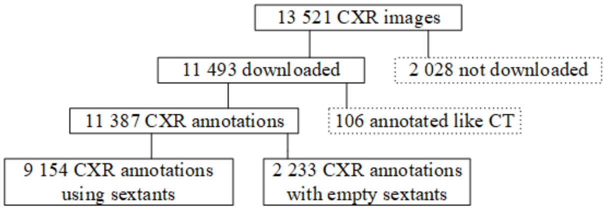

Figure 3. CXR image categories from TB Portals [1] website.

analysis there is a report on the use of the obtained data in solving the prediction of the parameter "Overall percent of abnormal volume" is given.

II. CXR annotations, as if it were a CT

There are 106 images described like a CT. All from Georgia. These images have 19 parameters with disease descriptions, like a CT scan. Annotation parameters for CXR are absent.

For example, in Fig. 4, patient ID 1534 has one annotated image with a study modality CR (Computed Radiography), but the actual description is as if it were a CT (Computed Tomography).

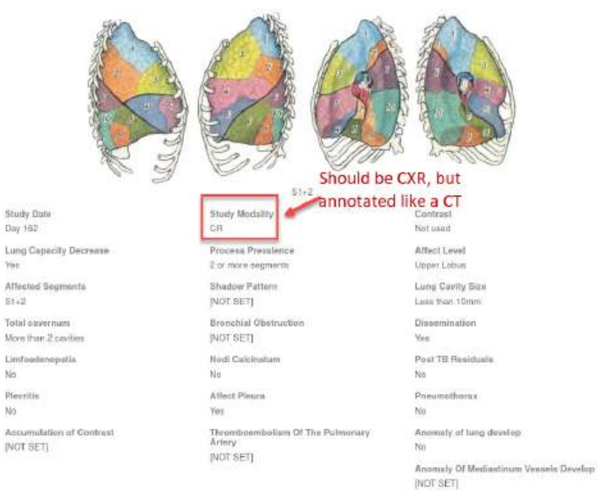

Figure 4. Patient ID 1534 with CXR annotation, as if it were a CT.

It was decided to exclude these images from the further investigations for the following reasons:

- their annotation is different than most other files, meaning that these images cannot participate in neural network training together with everyone else;
- 106 files can participate in separate neural network training, but there is more valuable data to explore;
- it is unclear why these images are described differently, and until this is clarified, it is better to exclude them from further study.

III. CXR annotations of sextants

A. *Lesion names to describe the lungs lobes*

In the CXR image, the lungs are divided into six lobes or sextants:

- Upper Right Sextant;
- Upper Left Sextant;
- Middle Right Sextant;
- Middle Left Sextant;
- Lower Right Sextant;
- Lower Left Sextant.

Each sextant has twenty identical lesion names, which can be found in the corresponding JSON files for each CXR image:

1) Small Cavities
2) Medium Cavities
3) Large Cavities
4) Is any Large cavity belong to a multi-sextant cavity?
5) Can Multiple cavities be seen?
6) Small Nodules
7) Medium Nodules
8) Large Nodules
9) Huge Nodules
10) Is any calcified or partially-calcified Nodule exists?
11) Is any non-calcified Nodule exists?
12) Is any clustered Nodule exists
13) Are multiple Nodules exist?
14) Low/ground glass Density (active fresh nodules)
15) Medium Density (stabilized fibrotic nodules)
16) High Density (calcified nodules, typically sequella)
17) Low/ground glass Density
18) Medium Density
19) High Density
20) Collapse

The data displayed on the CASE BROWSER [2] and TB DEPOT [3] websites and the tag names in JSON files are slightly different, but a mutually unambiguous match can be made even by a non-specialist. The only non-obvious correspondence is that the field "Infiltrate" on the website should correspond to the tag "High Density" in JSON files. This match between the "Infiltrate" field and "High Density" tag was found by the method of exceptions as the last possible match.

Mutually unambiguous correspondence between the data on the websites and in JSON files is given in Tab. I and Tab. III.

The sequential arrangement of the twenty lesions in Tab. I is the same as on the CASE BROWSER [2] website.

The number of non-empty lesions in the lung sextants is listed in Tab. II.

For convenience and to save space, in Tab. II, the word "Sextant" is omitted in table header. The total number of

Table I
Twenty lesions to describe the lungs lobes in CXR images

№	Tags in **JSON** files "info-all.json" [a]	**CASE BROWSER** [2] website	**TB DEPOT** [3] website
		Cavity	
1	Small Cavities	Small Cavities (less than 3 cm)	Small Cavities
2	Medium Cavities	Medium Cavities (3-5 cm)	Medium Cavities
3	Large Cavities	Large Cavities (more than 5 cm)	Large Cavities
4	Is any Large cavity belong to a multi-sextant cavity?	Does any Large cavity belong to a multi-sextant cavity?	Large Cavity Multi Sextant
5	Can Multiple cavities be seen?	Can Multiple cavities be seen?	Multiple Cavities Seen
		Nodules	
6	Small Nodules	Small Nodules (less than 3 mm)	Small Nodules
7	Medium Nodules	Medium Nodules (5-15 mm)	Medium Nodules
8	Large Nodules	Large Nodules (15-30 mm)	Large Nodules
9	Huge Nodules	Huge Nodules (more than 30 mm, tuberculoma)	Huge Nodules
10	Is any calcified or partially-calcified Nodule exists?	Any calcified or partially-calcified Nodules?	Partially Calcified Nodule Exists
11	Is any non-calcified Nodule exists?	Any non-calcified Nodules?	Non Calcified Nodule Exists
12	Is any clustered Nodule exists	Any clustered Nodules (nodules 2-5 mm apart)?	Any Clustered Nodule Exists
13	Are multiple Nodules exist?	Can Multiple Nodules be seen?	Multiple Nodule Exists
14	Low/ground glass Density (active fresh nodules)	Low/ground glass Density (active fresh nodules)	Low Ground Glass Density Active Fresh Nodules
15	Medium Density (stabilized fibrotic nodules)	Medium Density (stabilized fibrotic nodules)	Medium Density Stabalized Fibrotic Nodules
16	High Density (calcified nodules, typically sequella)	High Density (calcified nodules, typically sequella)	High Density Calcified Typically Sequella
		Infiltrate	
17	Low/ground glass Density	Low/ground glass Density	Infiltrate Low/Ground Density
18	Medium Density	Medium Density	Infiltrate Medium Density
19	High Density	High Density	Infiltrate
		Collapse	
20	Collapse	Collapse	Collapse

[a]Lesion names have been left unchanged for accuracy of description.

Table II
The number of non-empty lesions in the lung sextants

Lesion name	Upper Right	Upper Left	Middle Right	Middle Left	Lower Right	Lower Left
Small Cavities	6723	5856	4654	4717	2756	2596
Medium Cavities	6720	5856	4653	4714	2755	2596
Large Cavities	6719	5855	4652	4714	2755	2596
Is any Large cavity belong to a multi-sextant cavity?	6732	5868	4663	4727	2768	2608
Can Multiple cavities be seen?	6732	5867	4661	4727	2768	2606
Small Nodules	6724	5856	4656	4719	2757	2599
Medium Nodules	6724	5855	4653	4715	2755	2596
Large Nodules	6720	5853	4652	4714	2755	2596
Huge Nodules	6720	5854	4652	4714	2755	2596
Is any calcified or partially-calcified Nodule exists?	6733	5868	4666	4727	2768	2609
Is any non-calcified Nodule exists?	6731	5868	4665	4728	2768	2609
Is any clustered Nodule exists	6731	5864	4665	4727	2769	2608
Are multiple Nodules exist?	6730	5862	4661	4723	2768	2608
Low/ground glass Density (active fresh nodules)	6726	5856	4656	4718	2757	2599
Medium Density (stabilized fibrotic nodules)	6724	5857	4654	4716	2755	2596
High Density (calcified nodules, typically sequella)	6720	5853	4652	4714	2754	2596
Low/ground glass Density	6724	5856	4657	4715	2760	2598
Medium Density	6724	5858	4655	4716	2757	2597
High Density	6720	5855	4652	4713	2755	2596
Collapse	6720	5855	4652	4714	2754	2598

text annotations is 546,364. For the right lung: 282,817. For the left lung: 263,547.

The division of the lungs into sextants is made by a radiologist according to the following rules.

For CXR:

- Upper sextants are above the lower edge of the aortic arch;
- Middle sextants are between the lower edge of the aortic arch and the right inferior pulmonary vein);
- Lower sextants are below the right inferior pulmonary vein.

For CT:

- Upper sextants are above keel level;
- Middle sextants are between keel and right inferior pulmonary vein;
- Lower sextants are below right inferior pulmonary vein.

Approximately, *with some reservations*, the lung may be considered to be divided into three equal lobes vertically and into two equal lobes horizontally. It was decided to determine the boundaries of the sextants by calculating the third part of the total area of the lungs. With this choice of boundaries, the sizes of the sextants will be approximately the same in area. In this case the boundaries of the sextant are approximate and not related to lung anatomy. The left lung is where the heart and left arm are located. The right lung is where the right arm is located.

Tab. II shows that the highest number of lung diseases was registered in the upper lung lobes and the lowest number of lung diseases was registered in the lower lung lobes. At the same time, on average, more diseases were registered in the right lung than in the left lung.

According to Radiopaedia "post-primary infections have a strong predilection for the upper zones" [4]. Right versus left lesions asymmetry is not so obvious as top versus bottom and depends on the type (class) of the lesion. Statistical Atlas of Lung Lesions [5], [6], which was made by our team, is still in research. The difficulty in studying the statistical distribution of lesions is that they are of different nature and different biological substrates.

B. *Analysis of sextant lesion names*

According to the CASE BROWSER [2] website 20 lesion names are divided on the four groups (classes): "Cavity", "Nodules", "Infiltrate" and "Collapse" (see Tab. I).

Additional analysis of the sextant lesions list is shown in Fig. 5.

In Fig. 5, equally significant sections of text are highlighted with the same color. The 20 parameters are specially grouped in such a way that the same significant values are next to each other. Here are the definitions for these diseases.

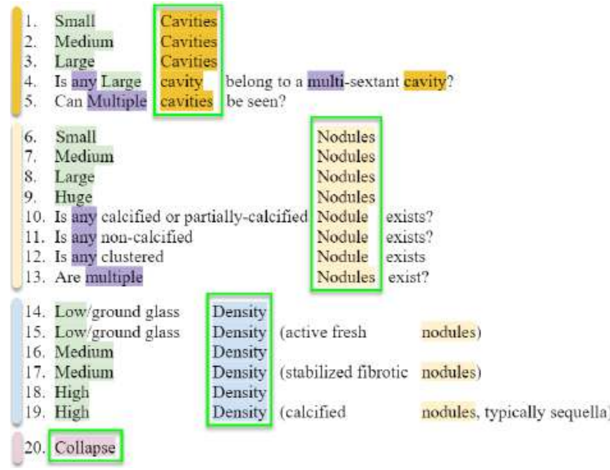

Figure 5. Graphical result of lesion names analysis for lung lobes.

A lung *cavity* or pulmonary cavity is an abnormal, thick-walled, air-filled space within the lung. Cavities in the lung can be caused by infections, cancer, autoimmune conditions, trauma, congenital defects, or pulmonary embolism. The most common cause of a single lung cavity is lung cancer. Bacterial, mycobacterial, and fungal infections are common causes of lung cavities [7].

According to the glossary of terms for chest imaging proposed by the Fleischner Society, a lung *nodule* is defined as an approximately rounded opacity more or less well-defined measuring up to 3 cm in diameter [8]. Rounded lesions measuring more than 3 cm in diameter are termed *lung masses* and should be considered indicative of lung cancer until histologically proven otherwise. Lung mass approach differs from that of nodules [9].

The *density* of the lung reflects the total mass of fluid, air, and dry lung tissue per unit volume of the lung. Lung density can be measured by evaluation of attenuation of an electron beam with CT. This technique has been shown to be sufficiently reliable and sensitive to distinguish normal from abnormal lung water [10].

A pulmonary *infiltrate* is a substance denser than air, such as pus, blood, or protein, which lingers within the parenchyma of the lungs. Pulmonary infiltrates are associated with pneumonia, tuberculosis, and sarcoidosis [11].

A *collapsed* lung (pneumothorax) occurs when air escapes from the lung. The air then fills the space outside of the lung between the lung and chest wall. This buildup of air puts pressure on the lung, so it cannot expand as much as it normally does when you take a breath. The medical name of this condition is pneumothorax [12].

Four general categories (classes) of disease can be distinguished from the definitions above:

1) class "Cavity" in the presence of one or more abnormal thick-walled spaces in the lung filled with air (gas);

2) class "Nodule" in the presence of one or more small up to 3 cm opacities;
3) class "Density" or "Infiltrate" in the presence of abnormal lung thickenings;
4) class "Collapse" or pneumothorax, when the lung cannot fully expand when breathe in.

Lesions analysis showed gradation in density, with the exception of the "Collapse" class. The "Cavity" class has the lowest density. Abnormal fluids and other seals are next that fall into the "Density" ("Infiltrate") class. And the densest is the "Nodule" class.

Nodules, abnormal density and cavities can be of different sizes (small, low, medium, large, high, huge), in different numbers (any, multiple, belong to a multi-sextant), in different qualities, especially nodules (calcified, partially-calcified, non-calcified, clustered, stabilized, active).

IV. CXR annotations of overall characteristics

Parameters used to describe the lungs as a whole (without sextants) are in Tab. III.

The seventh parameter "Timika Score" from Tab. III is not present in JSON files. The *Timika CXR score* is a machine learning tool for diagnosing tuberculosis, which was developed in 2010 by investigators at the Menzies School of Health Research in Darwin, Australia. The score in a scale of 1 to 140 was designed for physicians in underserved clinical settings and is based on the overall abnormal percent of volume of the lungs on CXR, plus the presence of cavitation [13], [14]. Most likely, the Timika CXR score is calculated automatically after the radiologist annotation and is not stored in the TB database.

The "Rater" parameter can have three values: "General practitioner", "Radiologist" or "Other", and if it is set, it means that the CXR image has been annotated. In this case, if the image is annotated and the sextants are blank, the patient is free of lung lesions.

V. Description of the top catalogues for further research

At this stage of the project, *the main effort* was focused on building the CXR image databases for further investigations. The catalogue tree is represented as follows.

level0 Original data from TB Portals [1] in DICOM format ("*.dcm") containing CXR, CT, JSON description files and other auxiliary files.

level1 All CXR and CT images have been converted to NIFTI format ("*.nii.gz").

level2 CXR images were manually reviewed and only those images that could be used in further research were selected into this catalogue.

- *cxr* Selected CXR images.
- *cxr_annotations* Excel tables with text annotations, paths in *level1* directory and other auxiliary information for each CXR file.
- *cxr_scripts* Python scripts to prepare *level2* directory.
- *cxr_masks* Lung masks obtained via LungExpert API [15].
- *cxr_thumbnails* Images preview in PNG format and size 512×512 pixels.
- *cxr_data* Preprocessed CXR images. For example, CXR images cropped by lung mask, normalized to the range [-1, +1] and then resized via lanczos method to 256×256 or 512×512 pixels.
- *temp* Intermediate files that will be deleted after *level2* is finished.

level3 Directories with investigations. Each subdirectory here is the separate investigation.

- *cxr_abnormal_volume* Investigation of the lung lesion percentage.
- *cxr_sextants* Investigation of the lung lobes.

All directories named "*cxr_*" mean that CXR files are processed.

All directories named "*ct_*" mean that CT files are processed.

CXR images processing pipeline:

- manual review and screening out defective images with saving data in the "*cxr*" directory;
- verification and correction for orientation, inversion, etc. with modification of the data in the "*cxr*" directory;
- getting mask via LungExpert API [15] with saving data in the "*cxr_masks*" directory;
- normalization and resizing with saving data in the "*cxr_data*" directory;
- slicing into sextants and saving data in the "*cxr_sextants*" directory.

The processing pipeline for the neural network is shown in Fig. 6:

- applying modality, inversion and orientation checks to the input image;
- obtaining lung mask via LungExpert API [15];
- cropping the image along the borders of the lungs, normalization and resizing;
- slice into sextants for lung lobes investigation;
- application of neural networks to determine overall characteristics or to determine sextant lesion characteristics;
- comparing the results obtained with the radiologist's annotations.

VI. Prediction of the parameter "Overall percent of abnormal volume"

A. Task description

The first parameter to research was the "Overall percent of abnormal volume" parameter. It can be an integer between 0 and 100 %. Zero percent means that the lungs are healthy and have no abnormal volume and 100 %

Table III
Parameters to describe the lungs as a whole

№	Tags in **JSON** files "info-all.json"	**CASE BROWSER** [2] website	**TB DEPOT** [3] website
1	Overall percent of abnormal volume	Overall percent of abnormal volume	Overall Percent of Abnormal Volume
2	Pleural Effusion. % of hemithorax involved	Pleural Effusion. % of hemithorax involved	Pleural Effusion Percent of Hemithorax Involved
3	Is Pleural Effusion bilateral?	Is Pleural Effusion bilateral?	Pleural Effusion
4	Other Non-TB abnormalities	Other Non-TB abnormalities	Other Non-TB Abnormalities
5	Are Mediastinal lymph nodes present?	Are Mediastinal lymph nodes present?	Are Mediastinal Lymph Nodes Present
6	Rater	Rater	Rater
7	—	Timika Score	—

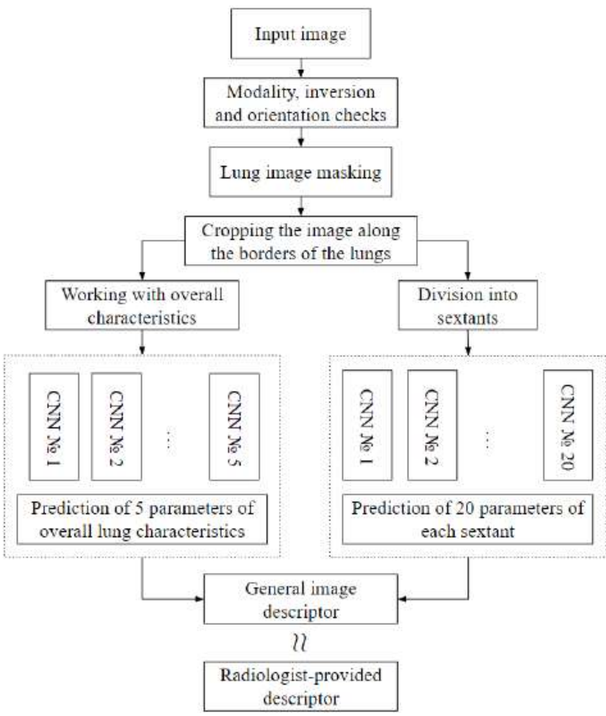

Figure 6. The processing pipeline for the neural network.

means that the entire lung volume is affected. From the TB DEPOT data dictionary [16]: "Overall percent of abnormal volume. Pleural effusion should be excluded. This is a professional judgment number in addition to the volume that can be calculated".

The *InceptionResNet50V2* neural network is used to predict the parameter "Overall percent of abnormal volume" based on the input CXR image. Among several tested architectures, this neural network showed the best results. It has also been suggested that it is not necessary to use a neural network to predict an abnormal percentage of lung volume. It was assumed that the "classical" machine learning method based on regression analysis would suffice.

Thus, three phases have been identified to fulfill this task:

1) preparation of the primary dataset;
2) application of the Support Vector Machine (SVM) regression method;
3) conducting experiments and comparing results.

This study was carried out in order to create a correct pipeline for further investigations of other CXR image parameters.

B. Preparation of the primary dataset

To prepare the dataset, all images were reviewed. Unsuitable images have been excluded. An example of such excluded images is shown in Fig. 7.

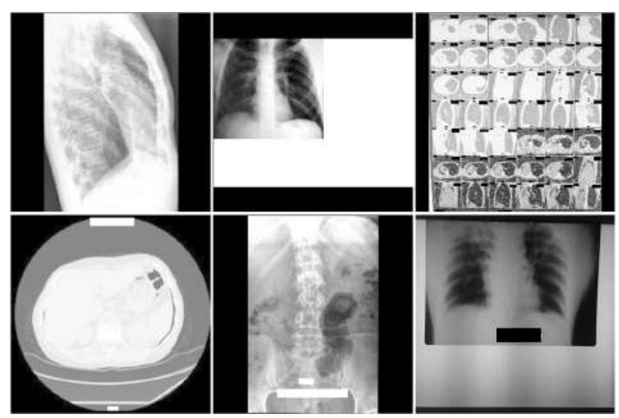

Figure 7. Examples of excluded images.

Excluded images are images with lateral patient orientation, containing large white/black frames, incorrect modality, other body parts, etc.

After review the primary dataset contains in total 8,875 images. It was noticed that some remained images are flipped horizontally (the heart is on the right side). For example, a patient with ID 8620 in Fig. 8 has heart on the right side of the body, the image is flipped.

A simple three-layer convolutional neural network (CNN) was implemented to find such incorrectly oriented images (Fig. 9).

To train this three-layer CNN an additional dataset was used. Each class of the training dataset contains 1,260 flipped (heart on the right) and 1,260 correctly oriented images. Test dataset contains 314 CXR images in each class.

F1-score on the test dataset was 0.9936. The resulting model was enough to optimize the process of preparing a

Figure 8. Incorrectly orientated image, the heart is on the right side.

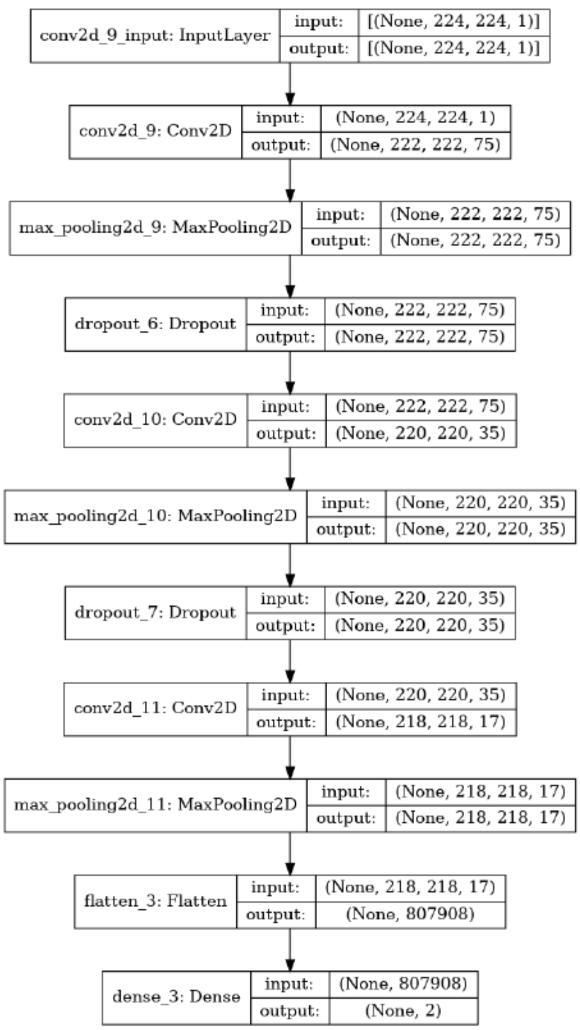

Figure 9. Three-layer CNN architecture to find horizontally flipped images.

general dataset. Solving the image orientation problem is important for proper slicing of CXR images into sextants.

As a result, the primary dataset obtained consists of 6,000 CXR images in the train set and 2,875 CXR images in the test set.

C. Application of the Support Vector Machine regression method

It was hypothesized that a "classical" machine learning method without the use of AI approaches would be sufficient to predict the value of parameter "Overall percent of abnormal volume" from the CXR input image.

To test this hypothesis SVM regression method using image histograms was used and compared with neural network method. A pre-trained on the ImageNet [17] CNN InceptionResNet50V2 was chosen as a neural network method.

Both methods are compared on the same dataset (6,000 CXR images in the train set and 2,875 CXR images in the test set).

The mean absolute error (MAE) was calculated as a metric for the analysis.

D. Conducting experiments and comparing results

MAE for the SVM method was 17.6494.

MAE for the InceptionResNet50V2 was 11.0730.

Undoubtedly, the margin of error is smaller when using InceptionResNet50V2. Accordingly, the SVM method did not perform well and cannot be used to predict "Overall percent of abnormal volume" parameter from the CXR input image.

The Grad-CAM [18] algorithm to visualize class activation maps was used to analyze the performance of the InceptionResNet50V2 neural network.

Examples of correctly predicted CXR images with their prediction heatmaps are shown in Fig. 10.

A comparison of the neural network prediction and the radiologist's annotation showed that prediction heatmaps are partially cover the sextants marked by the radiologist. This is the result for only one of the annotated parameters, for a combination of a group of parameters the results can be significantly improved.

Two examples of incorrectly predicted "Overall percent of abnormal volume" with their prediction heatmaps are shown in Fig. 11.

Both images in Fig. 11 have an "Overall percent of abnormal volume" parameter equal to 5 %, but the neural network predicts values of 25 % and 14 % respectively. On the first example the greatest activation of the heatmap occurred outside the lungs or in the background. On the second example the greatest activation of the heatmap revealed the artifact: protective lead apron.

Some possible errors in CXR textual annotations have also been discovered. For example, the patient with ID 426 in Fig. 12 obviously has a damaged right lung,

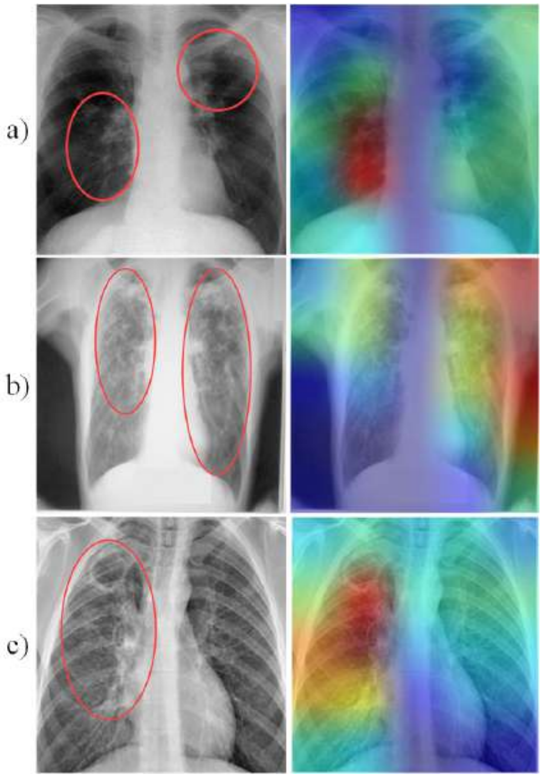

Figure 10. Examples of correctly predicted images with their prediction heatmaps for "Overall percent of abnormal volume" parameter:
a) patient ID 13676, actual value 14 %, predicted value 14 %;
b) patient ID 2641, actual value 30 %, predicted value 28 %;
c) patient ID 19415, actual value 34 %, predicted value 33 %.

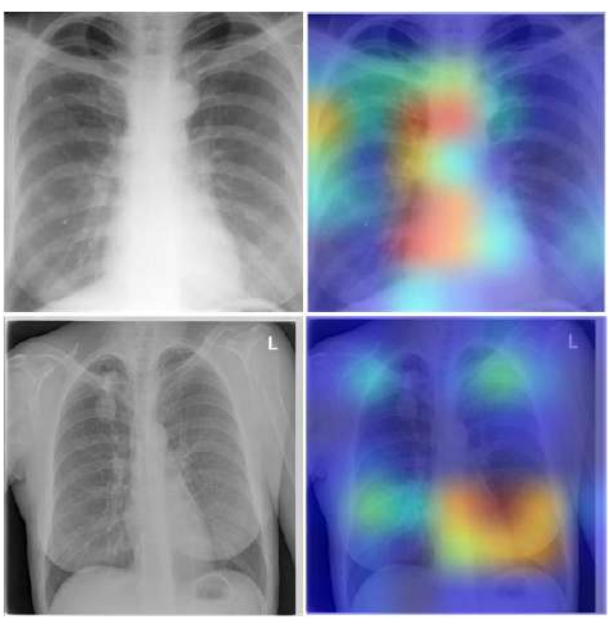

Figure 11. Examples of incorrectly predicted "Overall percent of abnormal volume" with corresponding prediction heatmaps.

but "Overall percent of abnormal volume" parameter is set to zero.

Figure 12. The patient's right lung has been damaged, but "Overall percent of abnormal volume" parameter is set to zero.

A study of patient ID 426 in the CASE BROWSER [2] found that the other parameter "Pleural Effusion. % of hemithorax involved" is equal to 50 %. Also, four sextants were annotated and have some lesions: Upper, Middle, Lower Right and Lower Left Sextants. The right lung is particularly badly damaged.

At least 44 such images with markup inconsistencies were found. It is necessary to exclude them from the training samples.

Simple consistency checks between different CXR annotation markup parameters should be developed for future research. It is necessary to pay attention not only to the images, but also to the various textual descriptions provided.

As a result, it was concluded that the high MAE value is due to the following reasons:

- background noise (outside the lungs) that requires mask cropping to remove it;
- the presence of artifacts in the lungs that requires more data to train the neural network to correctly distinguish these artifacts;
- some errors in CXR image annotations that should be excluded from the dataset.

VII. Discussion of the application of semantic technologies in the context of the tasks considered in this paper

In 1982, Japanese scientists developed a program for the fifth generation of electronic computing machines. Despite the fact that more than 42 years have passed, the fifth generation computers have not been fully realized. The main difficulty in creating fifth-generation computers or future computers is to create a machine with artificial intelligence (AI) that will be able to draw logical conclusions from the facts presented.

To interact with a fifth-generation computer, a person (user) will not need to develop software for the machine to solve the task at hand. In addition, fifth generation computers will solve the problem of data formalization in the interaction between computer and computer and human and computer. Commands for the machine can be formulated in ordinary spoken language without knowledge of formal programming languages as well as input and output data formats.

Thus, fifth-generation computers are intelligent semantic systems with extremely high interoperability. Interoperability is the ability of a product or system, whose interfaces are completely open, to interact and function with other products or systems without any access or implementation restrictions.

Although fifth generation computers have not yet been realized, active work is being done to remove the barriers between humans and computers, and between computers and computers.

One of the directions for the development of interoperable intelligent computer systems is automatic data transformation between different modalities: image to text, text to image, speech to text, text to speech, image to speech, speech to image, image to music, 3D reconstruction of an object from a set of images, CT scan to X-ray image, satellite image to geographical map, geographical map to satellite image and other modality transformations [19], [20].

This research project is developing an intelligent semantic system that converts a textual description into an image, namely the textual markup of a radiologist into a heatmap of lesion foci in the corresponding CXR medical image.

A neural network trained to solve such a problem will be a decision support system for population screening. The trained neural network model will produce a textual description of lung lesions based on the patient's chest radiograph, as well as a heatmap corresponding to these lesions in the form of a graphical representation Fig. 13.

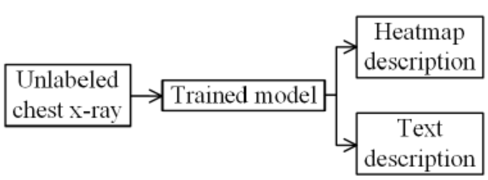

Figure 13. Operation schematic of the CXR-based decision support system under development.

The problem of generating a heatmap from a textual description and the subsequent problem of generating a heatmap and corresponding textual description from an input CXR image is one of the challenges of semantic image segmentation.

Semantic image segmentation is the task of dividing parts of an image into subgroups of pixels belonging to corresponding objects, with subsequent classification of these objects. Unlike classification and object detection tasks, the task of semantic segmentation is more complex both in terms of solution methods and computational resources [21].

Analyzing the literature to identify different semantic methods for medical image processing revealed the following approaches:

- probabilistic latent semantic analysis (PLSA), which, in conjunction with neural network, is able to mining the hidden semantics of an image [22];
- implementation of Semantic Similarity Graph Embedding (SSGE) framework, which explicitly explores the semantic similarities among images [23];
- investigation and development of the concept of a personal intellectual assistant (secretary, referent) [24].

Unfortunately, this paper does not apply the found semantic methods to medical image processing. However, the application of such semantic techniques in the context of the tasks considered in this paper is very relevant in future investigations.

As correctly noted in [25], "Currently, decision support systems in radiation mammology focus on the detection and classification of neoplasms, despite the fact that *the real work of a radiologist does not imply a diagnosing*. Computer vision systems use *a black box model* and do not explain the results of work, which is unacceptable in medicine".

This study also focuses on a deeper evaluation of the behavior of such a "black box" (neural network model) by studying the activation heatmaps on different convolutional layers of the neural network, which are obtained using the Gradient-weighted Class Activation Mapping (Grad-CAM) method [18]. The investigation of heatmaps on different convolutional layers provides a better understanding of the decision-making logic of the neural network based on the input data.

Fig. 5 shows the result of the semantic analysis of lesion names. As shown in Fig. 5, there are four classes of lesions: "Cavity", "Density", "Nodule" and "Collapse". Meanwhile, the two classes "Density" and "Nodule" have meaning overlap in the three lesion names. In the future, semantic analysis methods will be applied to better understand the disease asymmetry of different lung lobes (see Tab. II).

After the decision support system for population screening is completed, it is planned to be implemented into the already existing software "AI-based software for computer-assisted diagnosis of lung diseases using chest X-Ray and CT images" (LungExpert, https://lungs.org.by).

VIII. Conclusions

In this article, text annotations to CXR images were analyzed. At this stage of the project, the main efforts of the authors were focused on the formation of databases for further research. A catalogue tree with new datasets is described and constantly developing.

Two main tasks based on radiologists' textual annotations were planned and the corresponding pipeline for the neural networks training was described: pulmonary disease study using sextants and the pulmonary disease study using overall characteristics.

The task of predicting the parameter "Overall percent of abnormal volume" showed that it is necessary to develop additional simple consistency checks between different CXR textual annotations. Also using lung masks is a good idea to improve the quality of neural networks.

Further research is planned to create neural network attention heatmaps based on the textual descriptions of radiologists.

Acknowledgment

This work was carried out with the financial support of the ISTC-PR150 "Belarus TB Database and TB Portal" project.

References

[1] TB Portals. Available at: https://tbportals.niaid.nih.gov (accessed 2024, March)

[2] CASE BROWSER. Available at: https://data.tbportals.niaid.nih.gov (accessed 2024, March)

[3] TB DEPOT. Available at: https://depot.tbportals.niaid.nih.gov/#/cohort-creation?tab=4 (accessed 2024, March)

[4] Tuberculosis (pulmonary manifestations). Available at: https://radiopaedia.org/articles/tuberculosis-pulmonary-manifestations-1 (accessed 2024, March)

[5] Statistical Atlas of Lung Lesions. Available at: https://image.org.by/lesionAtlas (accessed 2024, March)

[6] V. Kovalev, V. Liauchuk, A. Gabrielian, A. Rosenthal, "Towards Statistical Atlas of Lung Lesions," *International Journal of Computer Assisted Radiology and Surgery*, 21-25 June, Munich, Germany, 2020, Vol. 15, Suppl. 1, pp. s31–s32. [Online]. Available: https://www.researchgate.net/publication/344217581_Towards_Statistical_Atlas_of_Lung_Lesions

[7] Lung cavity. Available at: https://en.wikipedia.org/wiki/Lung_cavity (accessed 2024, March)

[8] D.M. Hansell, A.A. Bankier, H. MacMahon, T.C. McLoud, N.L. Müller, J. Remy, et al., "Fleischner Society: Glossary of terms for thoracic imaging," *Radiology*, 2008, Vol. 246, No. 3, pp. 697–722. [Online]. Available: https://doi.org/10.1148/radiol.2462070712

[9] K. Loverdos, A. Fotiadis, C. Kontogianni, M. Iliopoulou, M. Gaga, "Lung nodules: A comprehensive review on current approach and management," *Annals of Thoracic Medicine*, 2019, Vol. 14, Issue 4, pp. 226–238. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6784443

[10] G. Metry, G. Wegenius, B. Wikström, V. Källskog, P. Hansell, P.G. Lindgren, H. Hedenström, B.G. Danielson, "Lung density for assessment of hydration status in hemodialysis patients using the computed tomographic densitometry technique," *Kidney international*, 1997, Vol. 52, Issue 6, pp. 1635–1644. [Online]. Available: https://doi.org/10.1038/ki.1997.496

[11] Pulmonary infiltrate. Available at: https://en.wikipedia.org/wiki/Pulmonary_infiltrate (accessed 2024, March)

[12] Collapsed Lung (Pneumothorax). Available at: https://www.pennmedicine.org/for-patients-and-visitors/patient-information/conditions-treated-a-to-z/collapsed-lung-pneumothorax (accessed 2024, March).

[13] 'Timika score' on x-rays may help identify complex TB cases. Available at: https://www.auntminnie.com/imaging-informatics/artificial-intelligence/article/15635618/timika-score-on-xrays-may-help-identify-complex-tb-cases (accessed 2024, March)

[14] A. Chakraborthy, A.J. Shivananjaiah, S. Ramaswamy, N. Chikkavenkatappa, "Chest X ray score (Timika score): an useful adjunct to predict treatment outcome in tuberculosis," *Advances in Respiratory Medicine*, 2018, Vol. 86, No. 5, pp. 205–210. [Online]. Available: https://doi.org/10.5603/ARM.2018.0032

[15] LungExpert. Available at: https://lungs.org.by (accessed 2024, March).

[16] TB DEPOT Data Dictionary. Available at: https://depot.tbportals.niaid.nih.gov/#/data-dictionary (accessed 2024, March)

[17] ImageNet. Available at: https://www.image-net.org (accessed 2024, March)

[18] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam., D. Parikh, D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *International Journal of Computer Vision*, 2020, Vol. 128, No. 2, pp. 336–359. [Online]. Available: https://doi.org/10.1007/s11263-019-01228-7

[19] Top 7 text-to-image generative AI models. Available at: https://byby.dev/ai-text-to-image-models (accessed 2024, Apr)

[20] Rohit Kundu. Image Processing: Techniques, Types, & Applications [2023]. Available at: https://www.v7labs.com/blog/image-processing-guide (accessed 2024, Apr).

[21] M. Arsalan, M. Owais, T. Mahmood, J. Choi, K.R. Park, "Artificial Intelligence-Based Diagnosis of Cardiac and Related Diseases," *Journal of Clinical Medicine*, 2020, 9(3):871, pp 1–27. [Online]. Available: https://doi.org/10.3390/jcm9030871

[22] M.R. Zare, M. Mehtarizadeh, "An Ensemble of Deep Semantic Representation for Medical X-ray Image Classification," *2021 55th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, USA, 2021, pp. 1–6. [Online]. Available: https://www.doi.org/10.1109/CISS50987.2021.9400268

[23] B. Chen, Z. Zhang, Y. Li, G. Lu, D. Zhang, "Multi-Label Chest X-Ray Image Classification via Semantic Similarity Graph Embedding," in *IEEE Transactions on Circuits and Systems for Video Technology*, April 2022, vol. 32, no. 4, pp. 2455–2468. [Online]. Available: https://www.doi.org/10.1109/TCSVT.2021.3079900

[24] V. Rostovtsev, "Intelligent health monitoring systems," *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]*, 2023, pp. 237–240. [Online]. Available: https://libeldoc.bsuir.by/handle/123456789/51289

[25] A. Kayeshko, A. Efimova, "Decision support system for breast cancer screening," *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]*, 2021, pp. 229–232. [Online]. Available: https://libeldoc.bsuir.by/handle/123456789/45424

ОБРАБОТКА РЕНТГЕНОВСКИХ ИЗОБРАЖЕНИЙ ГРУДНОЙ КЛЕТКИ НА ОСНОВЕ ТЕКСТОВЫХ АННОТАЦИЙ РАДИОЛОГОВ

Косарева А. А., Павленко Д. А.,Снежко Э. В.

Проанализировано более 11 000 рентгеновских снимков грудной клетки и соответствующих им текстовых аннотаций, а также проведены первые пилотные исследования по обработке изображений с учетом текстовых аннотаций специалистов-рентгенологов. Разработан конвейер обработки изображений для базы данных и нейронной сети. Проведено прогнозирование параметра «Общий процент аномального объема», для которого средняя абсолютная ошибка составила 11,073 при использовании нейросетевой модели InceptionResNet50V2.

Received 12.03.2024