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 multisextant), 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 1: 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 in- volved	Pleural Effusion. % of hemithorax in-volved	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	–

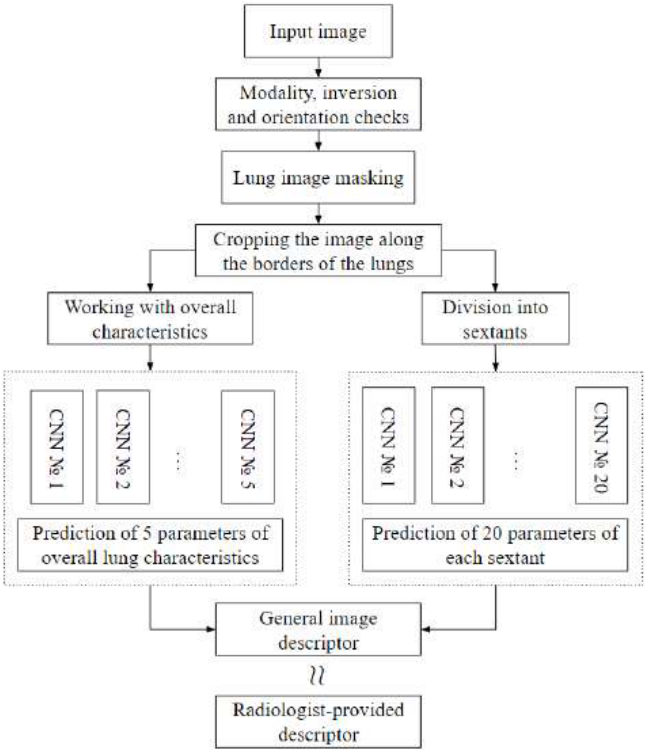
Input image
Modality, inversion and orientation checks
Lung image masking
Cropping the image along the borders of the lungs
Working with overall characteristics
Division into sextants
CNN № 1
CNN № 2
...
CNN № 5
Prediction of 5 parameters of overall lung characteristics
CNN № 1
CNN № 2
...
CNN № 20
Prediction of 20 parameters of each sextant
General image descriptor
≈
Radiologist-provided descriptor

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 vol- ume” 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.

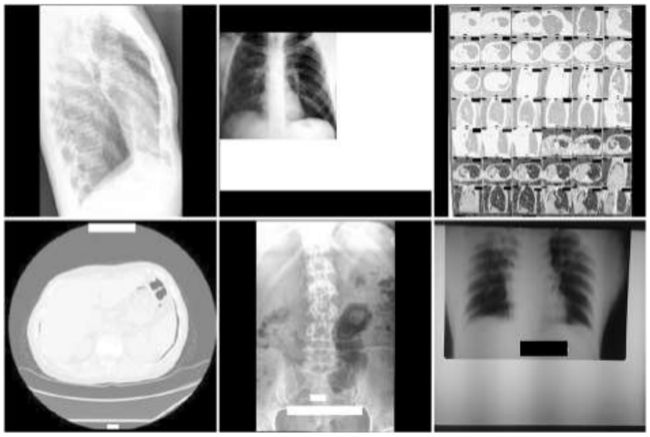

Figure 7: Examples of excluded images.

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

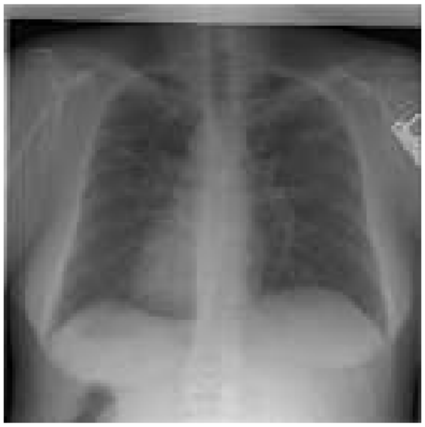

Figure 8: Incorrectly orientated image, the heart is on the right side.

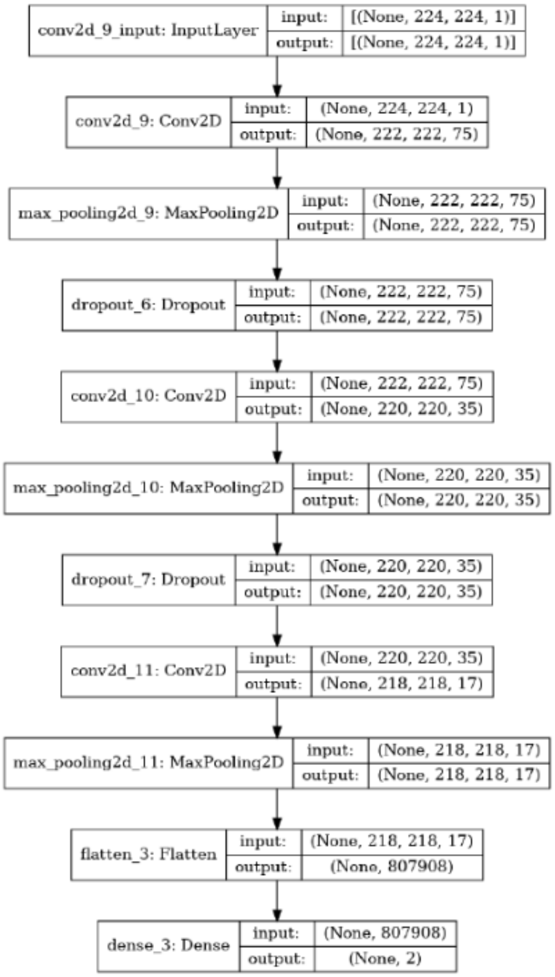
conv2d_9_input: InputLayer | input: | [(None, 224, 224, 1)] | output: | [(None, 224, 224, 1)]
conv2d_9: Conv2D | input: | (None, 224, 224, 1) | output: | (None, 222, 222, 75)
max_pooling2d_9: MaxPooling2D | input: | (None, 222, 222, 75) | output: | (None, 222, 222, 75)
dropout_6: Dropout | input: | (None, 222, 222, 75) | output: | (None, 222, 222, 75)
conv2d_10: Conv2D | input: | (None, 222, 222, 75) | output: | (None, 220, 220, 35)
max_pooling2d_10: MaxPooling2D | input: | (None, 220, 220, 35) | output: | (None, 220, 220, 35)
dropout_7: Dropout | input: | (None, 220, 220, 35) | output: | (None, 220, 220, 35)
conv2d_11: Conv2D | input: | (None, 220, 220, 35) | output: | (None, 218, 218, 17)
max_pooling2d_11: MaxPooling2D | input: | (None, 218, 218, 17) | output: | (None, 218, 218, 17)
flatten_3: Flatten | input: | (None, 218, 218, 17) | output: | (None, 807908)
dense_3: Dense | input: | (None, 807908) | output: | (None, 2)

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 acti- vation 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 param- eters, 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,