Automated analysis of x-ray images using transparency masks

A R Dabagov¹,², R A Tomakova²,⁴, V A Alekseev²,⁵ and D S Kondrashov²,⁶

¹JSC "Medical technologies Ltd", 31 Ibragimov str., Moscow, 105318, Russia
²Southwest State University, 94, 50 Let Oktyabrya str., Kursk, 305040, Russia
³E-mail: dar@mti.ru
⁴E-mail: rtomakova@mail.ru
⁵E-mail: vladislaw.al2015@yandex.ru
⁶E-mail: kondrashov012@mail.ru

Abstract. The method for classification of x-ray images, which is implemented in this article, improves the differential diagnosis of socially considerable diseases. The novelty of this method is that the input digital x-ray image is supplemented by a transparency mask which is found by image segmentation, and the classified sign vector is formed by pixels which aren’t masked by the transparency mask. The original x-ray image is divided into segments to classify morphological structures with pathological formings. The classification of the selected segment is implemented by a modified convolutional neural network which uses pooling layers and layers of a fully-connected neural network. The proposed classifier allows dividing patients into two groups: patients without found pathology and patients with disorders of health condition. The found morphological pathological formings help clinicians to find the most promising way for prevention of the disease development at the patients, saving money and time.

1. Introduction
X-ray image classifiers are widely used in intelligent medical systems and are related to digital image processing. Artificial neural networks (ANN) are used to construct intelligent systems for differential diagnosis of diseases [1-6].

2. Materials and methods
The novelty of the method is in that the input digital x-ray image is supplemented by a transparency mask which is found by image segmentation. The classified sign vector is formed by pixels which aren’t masked by the transparency mask.

A block diagram of convolutional neural network (CNN) that implements a classifier with the transparency mask is shown in figure 1.

A block diagram of device that is used by construction an automatic classifier is shown in figure 2.

This device form digital x-ray images to a matrix of object optical densities. The original image is sent to the classifier input. Deep image layers are created by processing of the original digital image with local filters. The local filters that are unique for each layer are implemented as identity operators.
Figure 1. Block diagram of a convolutional neural network that implements a classifier with a transparency mask.

Figure 2. Block diagram of device for obtaining of x-ray images.

The deep layer image size is reduced by pooling technology (subsampling). A space of informative signs is formed for the fully-connected neural network which is trained by scanning of the deep layers. Fully-connected neural network training is performed on the x-ray image with specified morphological pathological formings [7-9].
Pooling technology is designed to construct of three-dimensional tensors by two for each deep layer. The processing of the deep layer is from two differential operators. The number of the elements is determined by the depth layer scale with binary outputs and a threshold activation function. As a result, each scale is defined by three-dimensional megapixels.

Transparency mask and single-layer perceptrons are used as subdescriptors. The transparency mask cuts through three-dimensional tensors of informative signs, at the same time only informative signs, which are pertained to a single segment, are remained in output.

The results of the classification measurement arrive at smart analyzer.

The smart analyzer detects outliers of each single-layer perceptron and then it aggregates the produced classification results.

The aggregated results form inputs of the fully-connected neural network. This neural network is configured to classify a segment with a given pathology.

3. Results
The algorithm diagram of the proposed method is shown in figure 3.

![Algorithm diagram of the method for classification of x-ray images.](image)

**Figure 3.** Algorithm diagram of the method for classification of x-ray images.
This algorithm provides 4 steps of execution.

At the first step, a transparency matrix is formed. The transparency matrix size corresponds to the size of the x-ray image matrix. Transparency matrix elements take on the value 0 or 1. The transparency matrix pixels that are to the interest area of the x-ray image (segments) take on the value 1.

The transparency matrix generator is a two-alternative classifier.

Transparency masks are stored in files. The transparency mask for the interest pneumonia zone is shown in figure 4(a), and the transparency mask for x-ray mammogram with the interest cancer zone cancer is shown in figure 4(b).

![Figure 4](image)

**Figure 4.** Example of transparency masks which are determined in the first step of the classification of x-ray images.

At the second step, deep layers of original image are formed in the system. This method uses identity operators. The identity operators have a mask structure that brightness of pixel with \((i, j)\) coordinates is presented as

\[
y'_{i,j} = f \left( \sum_{g=0}^{M_1-1} \sum_{q=0}^{M_2-1} x_{(i-1)+g, (j-1)+q} \cdot w'_{g,q} \right),
\]

where \(u\) is convolution identifier (scale identifier), \(M_1 \times M_2\) is convolution matrix size, \(w'_{g,q}\) are weight coefficients of the convolution with \(u\) identifier.

Peculiarity of the identity operator is that only one weight coefficient, that coordinate coincides with the coordinate of active pixel in the image, take on the value 1.

The identity operator translates original image to the deep layer with assigning of a defined scale. This requirement is responded to the following product type rule

\[
IF ((g=M_1-1) AND (q=M_2-1)), THEN w'_{g,q} = 1, ELSE w'_{g,q} = 0.
\]

The original image is marked up in megapixels. The image size in the deep layer is reduced and takes on the value \((N_1 - M_1/2) \times (N_2 - M_2/2)\), where \(N_1 \times N_2\) is size of the original image, and \(M_1 \times M_2\) is size of the megapixel.

At the third step, the size of the deep layers is reduced. This process of CNN technology is called subsampling or pooling.

The task of the pooling technology is to decrease cards of the signs, that is, to reduce the sign number in the deep layers. As in the implementation of a convolutional layer, a matrix of weight coefficients is needed to the classic pooling technology, which is used in the CNN. Example of 2×2 matrix classic pooling that reduces deep layer size from 24×24 to 12×12 is shown in figure 5.
Figure 5. Example of the classic pooling technology.

Two Haar wavelets, vertical and horizontal, are used for each scaling window in the pooling technology.

The size of a vertical wavelet is equal to $(M_2/2) \times M_1$, and the size of a horizontal wavelet is equal to $(M_1/2) \times M_2$. The horizontal Haar wavelet is moved along the mask in a vertical direction with a $\Delta V$ step, and the vertical Haar wavelet is moved along the mask in a horizontal direction with a $\Delta G$ step.

Difference for each $i$-th position of the wavelet is calculated to

$$Z_i = W_i - B_i,$$

where $W_i$ is sum of brightness of the pixels that is located under “white” part of the Haar wavelet, $B_i$ is sum of brightness of the pixels that is located under “black” part of the Haar wavelet.

Requirements for space of the informative signs are as follows:

I. The informative sign vector shouldn’t depend on the deep layer scale and number of scale windows in the segment;

II. Invariance of the informative signs is to the dynamic range of pixel brightness on the x-ray image.

A single-layer perceptron, which is trained for a known class with a known scale mask, is used to perform of the first requirement. Therefore, its structure doesn’t depend on the number of scale masks (megapixels) which are included on the classifying segment. This single-layer perceptron is a “weak” classifier. The “weak” classifiers form an input vector for the fully-connected neural network.

A threshold parameter $\eta$ is entered to implement of the second requirement.

Result of interaction to masked pixels with the Haar wavelet is transformed as

$$F_i = \begin{cases} 
-1, & \text{if } Z_i < -\eta; \\
0, & \text{if } -\eta \leq Z_i \leq \eta; \\
1, & \text{if } Z_i > \eta.
\end{cases}$$

Counts that are accorded to (3) are got to result of moving the Haar wavelet along the scale mask.
4. Discussion of results
X-ray images and clinical observation results of patients, who were examined in the Kursk City clinical hospital of emergency medical services, were used to form of the training and control samples. The 84 annotated thorax x-ray images with focal darkening in the lungs were selected for research.

The control sample consists of 50 patients who passed a medical examination for employment. Fluorograms which did not show any pulmonary pathology were used by system training too.

Diagnostic sensitivity (DSe), diagnostic specificity (DSP) and diagnostic efficiency of the decision rule (DE) were used to calculate indicators of the quality of diagnostic decision rules.

It is considered that the first class includes x-ray images are without pathology; the second class includes x-ray images with pathology.

The results of the convolutional neural network work on the control sample are shown in table 1.

Table 1. Results of control tests at the diagnostic x-ray examination of the patients.

| Examined patients | Results of the rule operations | \( \Sigma \) |
|-------------------|-------------------------------|-------------|
| \( n_{\text{off}}(50) \) | Positive: 46 | Negative: 4 | 50 |
| \( n_{\text{off}}(34) \) | 6 | 28 | 34 |
| \( \Sigma \) | 52 | 32 | 84 |

Analysis of the experiment results showed that DSe = 82%, DSP = 94%, DE = 89%. These parameters exceed the quality indicators of x-ray image diagnostics from convolutional neural network structures by 10-15%.

5. Conclusion
The original x-ray image is divided into segments in order to solve the task of classifying morphological structures on the grayscale x-ray images. The interest free-form zones of the x-ray image are selected by segmenting of this image. The transparency mask, whose size is as the original image, is attached to this image in order to convert a non-raster segment to a raster one. Classification of the selected segment is implemented by a modified convolutional neural network which uses pooling layers and layers of a fully-connected neural network. A layer of the "weak" classifiers which are constructed on the single-layer perceptrons is after the pooling layer in the convolutional neural network. The number of "weak" classifiers doesn't depend on the structure and size of the megasegment (scale of the deep layer). Each "weak" classifier is configured for a specific class of the segment and specific scale mask size. Each weak classifier makes a result about belonging of a megapixel on the selected segment to a given class, but the fully-connected neural network gets average information by all segments.

The analysis of the results by the automatic classifier showed that DSe = 82%, DSP = 94%, DE = 89%. These parameters exceed the quality indicators of x-ray image diagnostics from the convolutional neural network structures by 10-15%.

The proposed classifier allows dividing patients into two groups: patients without found pathology and patients with disorders of health condition. The found morphological pathological formings help clinicians to find the most promising way for prevention of the disease development at the patients, saving money and time.

References
[1] Dyudin M V, Filist S A and Kudryavtsev P S 2014 Science Intensive Technologies 15(12) 25–30
[2] Filist S A, Dabagov A R, Tomakova R A, Malyutina I A and Kondrashov D S Proc. of the Southwest State University. Series: Control, Computer Engineering, Information Science. Medical Instruments Engineering 9(1) 49–61
[3] Tomakova R A, Filist S A and Durakov I V 2018 Human Ecology 6 59–64
[4] Filist S A, Tomakova R A, Degtyarev S V and Rybochkin A F 2017 Biomedical Engineering 5 41–45
[5] Hayat M, Mai M, Mabrouk S and Amr S 2014 Computer methods and programs in biomedicine 116 226–235
[6] Tomakova R A, Filist S A, Gorbatenko S A and Shevtsova N A 2010 Biotechnosphere 3(9) 54–60
[7] Tomakova R A, Filist S A and Rudenko V V 2011 Belgorod State University Scientific Bulletin Economics Information technologies 1(96) 188–195
[8] Filist S A, Tomakova R A, Brezhneva A N, Malyutina I A and Alekseev V A 2019 Radio Industry (Russia) 29(1) 45–52
[9] Kudryavtsev P S, Kuzmin A A and Filist S A 2016 Biomedicine Radioengineering 9 10–15