

AUTOMATIC QUALITY CONTROL IN LUNG X-RAY IMAGING WITH DEEP LEARNING

A. A. Dovganich,¹ A. V. Khvostikov,^{1,2} A. S. Krylov,^{1,2} and L. E. Parolina³ UDC 519.6+004.891.3

The development of deep learning and its growing application in medical diagnosis have focused the attention on automatic control of image quality for neural-network medical image analysis algorithms. This article presents a method for automatic determination of the hardness (penetration) of lung X-ray images using standard criteria from chest X-ray diagnosis. The proposed method can be applied to automatically filter images by hardness (penetration) level and to detect low-quality images, thus facilitating the creation of high-quality data sets and increasing the efficiency of neural-network approaches to the analysis of lung X-ray images.

Keywords: lung X-ray imaging, deep learning, quality control, convolutional neural networks, X-ray hardness (penetration).

Introduction

Analysis of medical data using deep neural networks is becoming increasingly developed and popular. Particularly relevant are methods based on convolutional neural networks in image analysis, including classification, segmentation, detection, anomaly identification, etc. Neural-network methods are convenient and effective in image analysis because they do not require specialized algorithms for detection and description of imaged structures that are difficult for medical diagnosis; they furthermore have excellent generalizing capacity.

However, the principle of direct feature construction from images, which is the basis of current convolutional neural networks, makes neural-network approaches highly demanding with regard to image quality and homogeneity of the model-training sample.

The quality of the images fed into the neural-network model for both training and validation is therefore of the utmost importance. Due to the specifics and the wide range of medical equipment settings, quality control of input images is relevant for various image modalities, e.g., magnetic-resonance tomography [1], computer tomography [2], and radiography [3].

In this article, we consider “hardness” (penetration) control for chest X-ray imaging. Radiography is used to diagnose various lung diseases, in particular tuberculosis [4, 5] and COVID-19 related lung damage [6, 7, 8]. The “hardness” of an X-ray image indicates the optimality of the X-ray dose chosen for the particular patient. An adequate X-ray dose reveals various structural features of the chest organs in the image. When analyzing and describing an X-ray image, the radiologist always considers its hardness, because underestimation of this factor may lead to incorrect interpretation. During the learning stage, the neural network is tuned to the hardness of the images from a training set. A neural network trained on “hard” images will produce a poor result when

¹ Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow, Russia; e-mail: a.dovganich@yandex.ru.

² Moscow Center of Pure and Applied Mathematics, Moscow, Russia; e-mail: khvostikov@cs.msu.ru; kryl@cs.msu.ru.

³ National Medical Research Center of Phthisiopulmonology and Infectious Diseases, Russian Ministry of Health, Russia; e-mail: parolinale@nmrc.ru.

shown “soft” images and vice versa. To resolve this difficulty, we propose an automated method for assessing the hardness of chest X-ray images. In the first stage, this allows to filter the input data and omit X-ray images that a priori fall outside the competency range of the classifier. Subsequently, it opens possibilities for training several classifier models designed to work with X-ray images of different hardness levels and allows input data to be distributed between them

To develop an automated hardness determination procedure for X-ray images, we have decided to formalize the empirical method applied by radiologists. A criterion of optimal X-ray hardness (penetration) for a particular patient is clear visibility of the three-four top thoracic (not any other!) vertebrae. In an enhanced penetration (“hard”) image, thoracic vertebrae below the fourth are clearly imaged. If a chest X-ray shows clearly fewer than three top thoracic vertebrae, this is an indication of reduced penetration (a “soft” X-ray image) [9, 10].

To determine what vertebrae are clearly visible, we adapted an automated method of spine detection and vertebrae visualization on X-ray images used for diagnosing scoliosis [11]. Spine and chest X-ray images have different formats, different overall brightness levels, and different contrast. Therefore, some steps in the spine image analysis algorithm had to be modified or replaced. The method of ridges was used to detect and count the vertebrae on an already detected spine.

The Dataset

We used three public datasets for the development and testing of the method: Montgomery [12], Shenzhen [12], and Open-I [13]. The images were acquired on different machines with different settings and visually differed by “hardness”.

The Montgomery data set consists of 138 grayscale images with 8-bit color depth and $\sim 4,000 \times 4,900$ pixel resolution. The Shenzhen data set consists of 662 grayscale images with 8-bit color depth and $3,000 \times 2,900$ resolution. All images are accompanied by a clinical description of the diagnosed disease. The Open-I data set is somewhat different: it has 7,470 grayscale images with 8-bit color depth and 512×512 resolution. The radiologist’s interpretation is included for 3,955 images.

Preprocessing of X-Ray Images

The distortion of the brightness range in X-ray imaging is not uniform over the image, changing from region to region. This is primarily due to the fact that X-rays passing through the observed object are attenuated differently in different parts of the organ. The degree of attenuation depends on the chemical composition of the organ, its density and thickness. Furthermore, the X-ray penetration (hardness) increases as the X-ray tube voltage increases. Harder radiation leads to higher scattering, and scattered radiation is one of the causes of the contrast-reducing diffuse veil in X-ray images.

There are three main approaches to equalizing brightness on X-ray images: normalization, classical histogram equalization, and adaptive histogram equalization. Normalization works on images where a part of the range is totally missing (there are no very bright or conversely very dark pixels), which does not apply to the datasets used in this study. The standard histogram equalization method works well for images with a very low overall contrast. However, histograms of lung X-ray images span the entire range of intensities and low contrast is observed only in localized regions (Fig. 1). To resolve the difficulty with brightness equalization in localized regions, we applied adaptive histogram equalization. The results produced by the three algorithms applied to the lung X-ray from Fig. 1 are compared in Fig. 2.

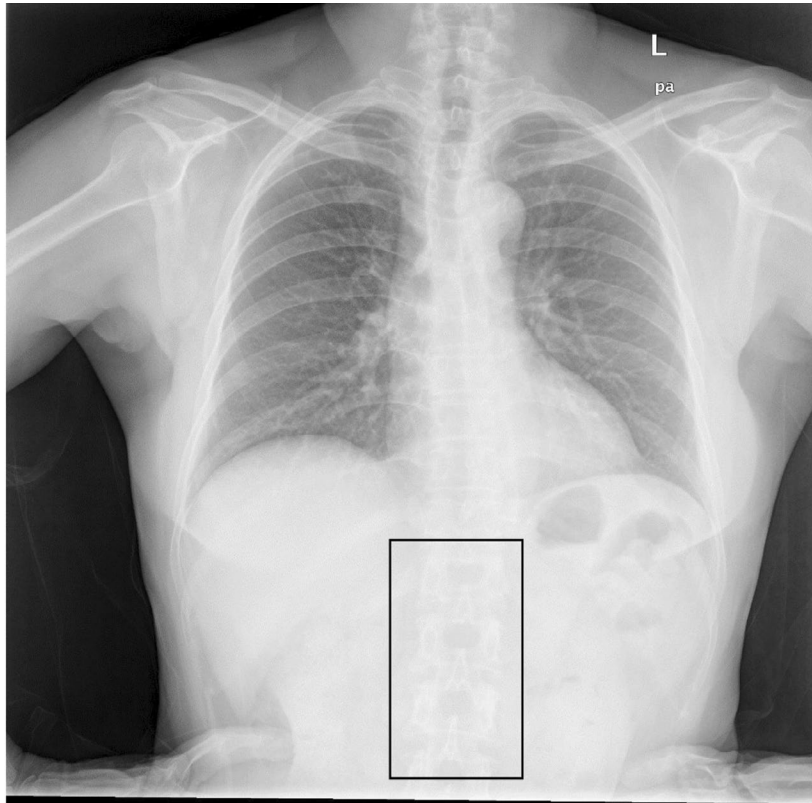


Fig. 1. A lung X-ray image. The rectangle encloses a low-contrast region.

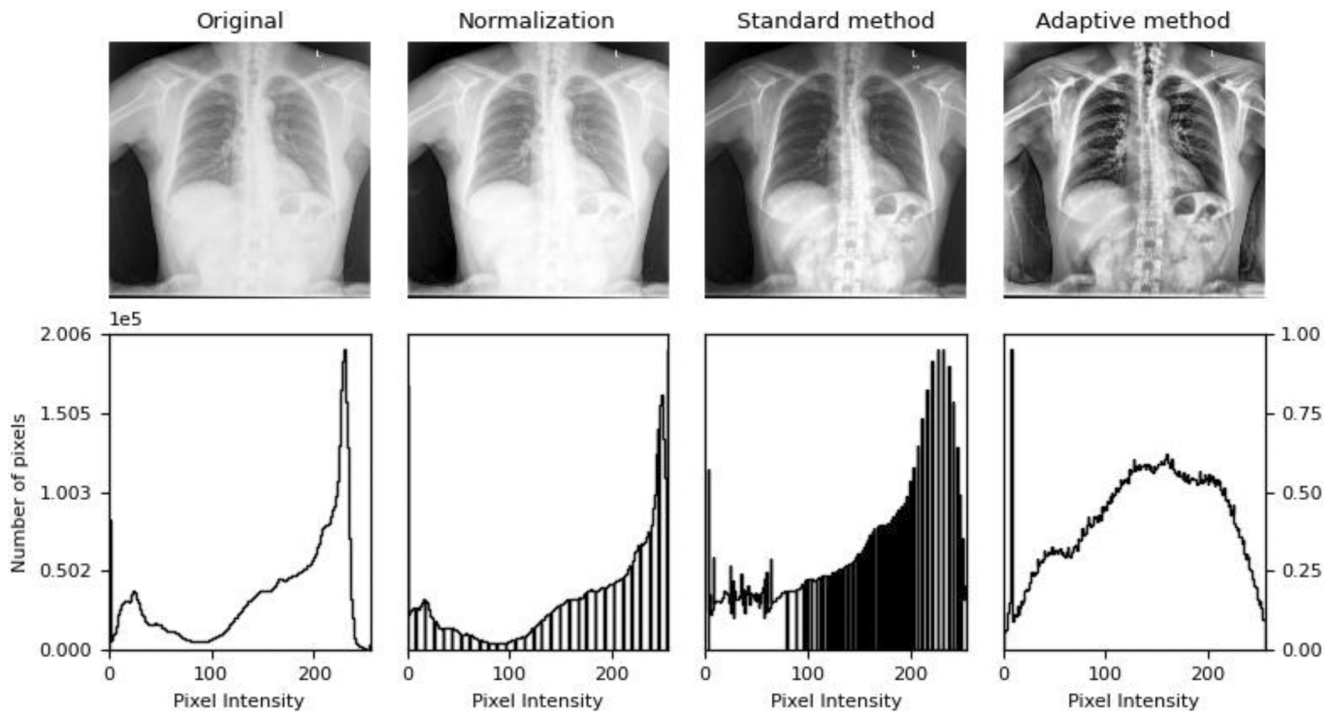


Fig. 2. Application of different brightness equalization algorithms with the resulting histograms shown.

Table 1
Results Produced by the Spine Detection Algorithm

| Epoch | L1-loss training sample | L1-loss test sample | IoU training sample | IoU test sample |
|-------|-------------------------|---------------------|---------------------|-----------------|
| 2 | 0.534 | 0.257 | 0.001 | 0.012 |
| 3 | 0.136 | 0.064 | 0.470 | 0.627 |
| 5 | 0.053 | 0.053 | 0.741 | 0.676 |
| 10 | 0.049 | 0.050 | 0.755 | 0.698 |

Spine Detection

The first stage of our method is spine detection by the neural-network approach. To ensure greater processing effectiveness, we first reduced the size of all images to one-quarter of the initial size. Then we applied a neural-network method that cut out the regions of interest in the form of rectangles. The Amazon SageMaker Ground Truth service tagged the images from the original datasets. The tagging results were acquired as a json file, with the relative coordinates of the rectangular region shown for each image. Then the combined dataset was mixed and split into a training sample and a test sample in a ratio of 80% to 20%. Then, the deep-learning framework pytorch was applied to complete the training of the model resnet-34 [14] locally on the graphics card during 10 epochs. L1-loss was used as the loss function during training. To assess the quality of the results produced by the neural-network model, we additionally calculated after each epoch the Intersection over Union (IOU) quality measure on the training and the test samples. The results have shown that the spine region is localized quite successfully on the lung X-ray image. The spine detection results are presented in Table 1.

Detection of the Spine Central Line and Spine Boundaries

The next step after spine detection is the detection of separate vertebrae on the X-ray image. The detection of vertebrae consists of three steps:

1. Detection of the spine central line (Figs. 3a, 3c);
2. Detection of spine boundaries (Figs. 3b, 3c);
3. Detection of individual vertebrae

Below we describe the algorithm for the detection of the spine central line. We use $H \times W$ rectangular windows moving horizontally by one pixel. Inside each window, we evaluate the sum of pixel intensities and find the window with the maximum total intensity. The upper midpoint of this window provides the first reference point for spine central line detection:

$$s(x, y) = \sum_{i=0}^H \sum_{j=-W/2}^{W/2} I(x + j, y + i),$$

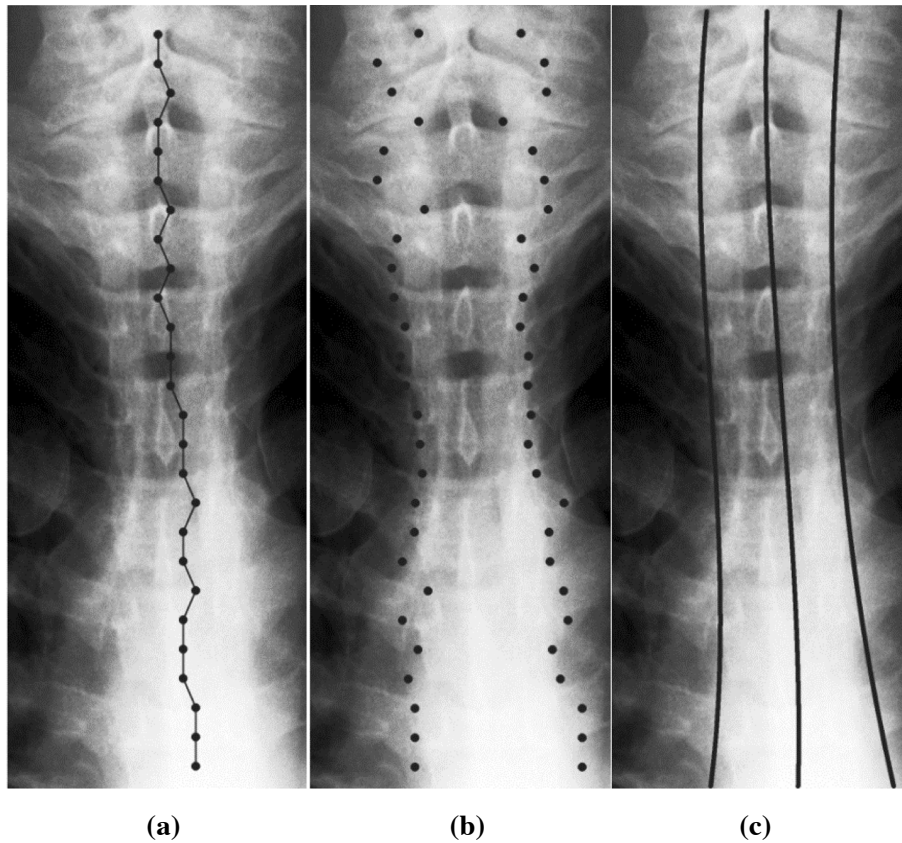


Fig. 3. Detection of the spine central line and boundaries: (a) spine central line, (b) spine boundaries, (c) spine central line and boundaries after interpolation.

$$(x_{\max}, y_{\max}) = \operatorname{argmax}_{(x,y)} (s),$$

where I is the $H_I \times W_I$ region containing the spine image, $H = \frac{H_I}{8}$, $W = \frac{W_I}{4}$, (x, y) are the window coordinates (x is the coordinate of the window midline, y is the coordinate of the lower window boundary), $s(x, y)$ is the sum of pixel intensities in the window, (x_{\max}, y_{\max}) are the coordinates of the acquired point.

Then the current rectangular window is moved p pixels down and the next control point is sought in a q -pixel interval on both side with a one-pixel step. The window upper midpoint is taken as the second control point for the spine central line. The calculations are repeated until we have found n control points (x_{c_i}, y_{c_i}) , $i = 1, \dots, n$. Central line detection is concluded with polynomial interpolation over the n control points.

Then we look for spine boundaries along the x -axis. The left and the right boundaries of the spine are defined by points at the vertebrae edges with the coordinates (x_{l_i}, y_{l_i}) and (x_{r_i}, y_{r_i}) with $i = 1, \dots, n$, where $y_{l_i}, y_{r_i} = y_{c_i}$ and $x_{l_i} < x_{r_i}$. During the detection of the left and right spine boundaries, two adjacent windows of equal size are used; these windows are moved by at most r pixels along the normal to the spine central line in each direction from the center. The upper midpoint of the adjacent windows with maximum intensity difference is identified as a point on the spine boundary. Boundary detection continues until all the control points of the spine central line have been examined.

The left and right spine boundary detection concludes with polynomial interpolation over the n detected points on each side of the spine. Third-degree polynomials are used here and in spine central line detection.

The parameters of the method were related to the image dimensions H_I and W_I ,

$$H = \frac{H_I}{8}, \quad W = \frac{W_I}{4}, \quad p = \frac{H}{4}, \quad q = \frac{W}{2}, \quad r = \frac{W_I}{2},$$

where H_I is the image height in pixels, and W_I is the width. An example of the results produced by the spine detection algorithm is shown in Fig. 3.

Vertebrae Detection and Counting

The vertebrae counting procedure proposed in [11] has proved itself satisfactorily on spine X-ray images but has not produced acceptable results on chest X-ray images. This is so because lung X-ray imaging uses smaller radiation doses than spine imaging. We have accordingly decided to replace the method of [11] with the construction of a modified chart of ridge structures. The chart was constructed by the following algorithm:

- 1) finding second derivatives of the images;
- 2) forming the pixel-wise matrix $Q = \begin{pmatrix} L_{xx} & L_{xy} \\ L_{xy} & L_{yy} \end{pmatrix}$, where L is the intensity of the original image;
- 3) finding the eigenvalues λ_1 and λ_2 of the matrix Q , where $\lambda_1 > \lambda_2$;
- 4) constructing second ridge-structure charts I_1 и I_2 from the eigenvalues λ_1 and λ_2 , where $I_i(x, y) = \frac{\lambda_i(x, y)}{\max(\lambda_i)} * 255$, $i = 1, 2$ (Figs. 4d, 4e);
- 5) binarizing each chart by the thresholds θ_1 and θ_2 (Figs. 4f, 4g);
- 6) finding the intersection I_R of $I_{1\theta}$ and $I_{2\theta}$ (Fig. 4h).

The algorithm produces an intensity matrix, which is a modified ridge structure chart.

Then we count the independent connected regions with at least one point in the spine central line. The end result is the number of detected vertebrae.

A general scheme of the algorithm is shown in Fig. 5.

In this study, we validated the vertebrae detection and count algorithm using vertebrae tagging by medical experts. We compared the number of vertebrae counted by the algorithm on the X-ray image with the number of vertebrae detected by an expert. For 800 images we then calculated the difference $r = |k_e - k_a|$, where k_e is the expert vertebrae count and k_a is the count produced by the algorithm, r is the vertebrae-count algorithm error. The value of r did not exceed 2 ($r_i \leq 2$, $\forall i \in [1, 800]$, where i is the image serial number).

Hardness-Determination Results for X-Ray Images from Montgomery and Shenzhen Datasets

To analyze the hardness matching of the X-ray images from the Montgomery and Shenzhen datasets, which are often used simultaneously for training and testing of deep-learning systems for COVID-19 diagnosis, we constructed the histograms of the number of vertebrae in the images in the two datasets.

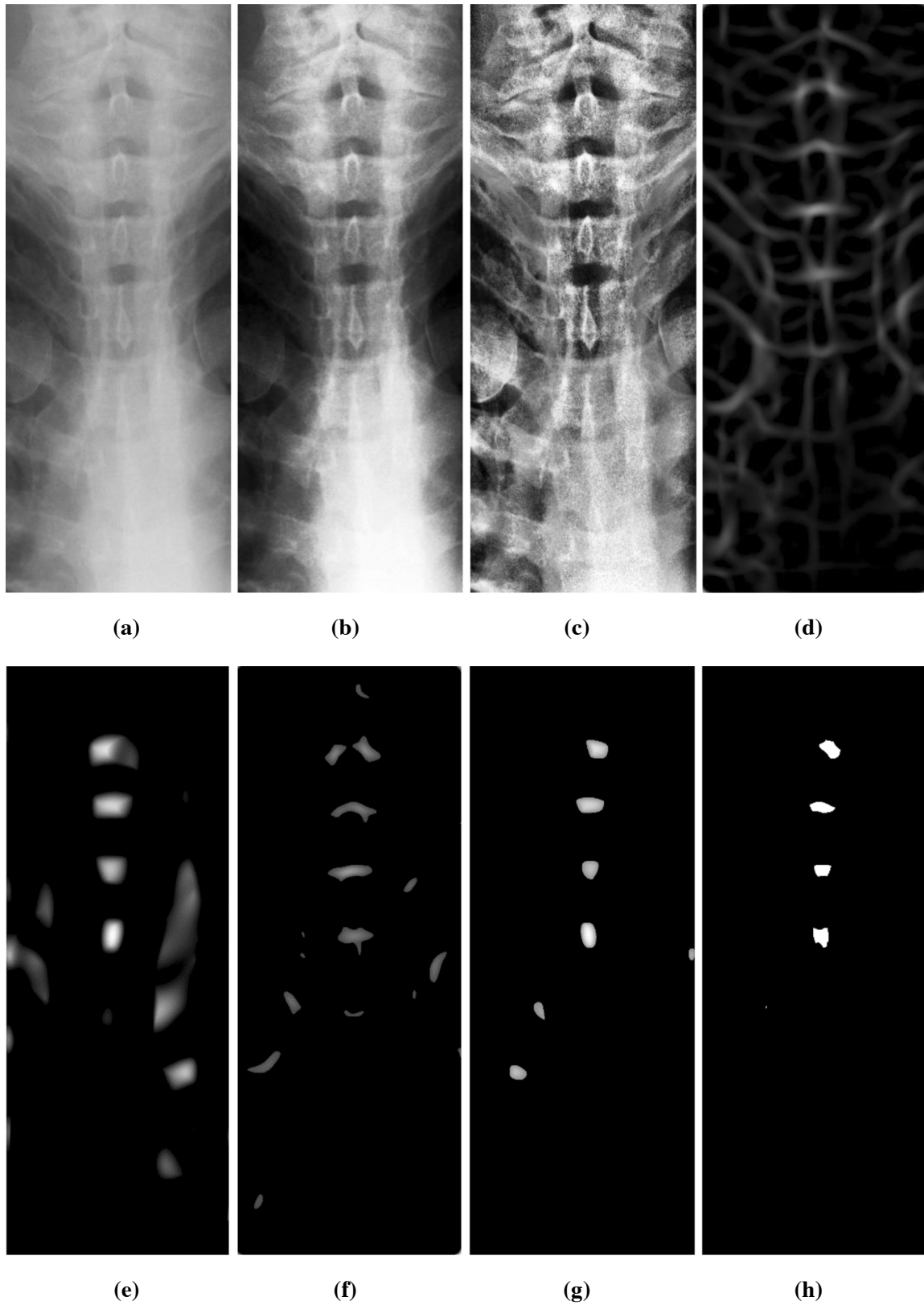


Fig. 4. (a) the original image, (b) standard histogram equalization, (c) adaptive histogram equalization, (d) I_1 , (e) I_2 , (f) I_{1_0} , (g) I_{2_0} , (h) I_R .

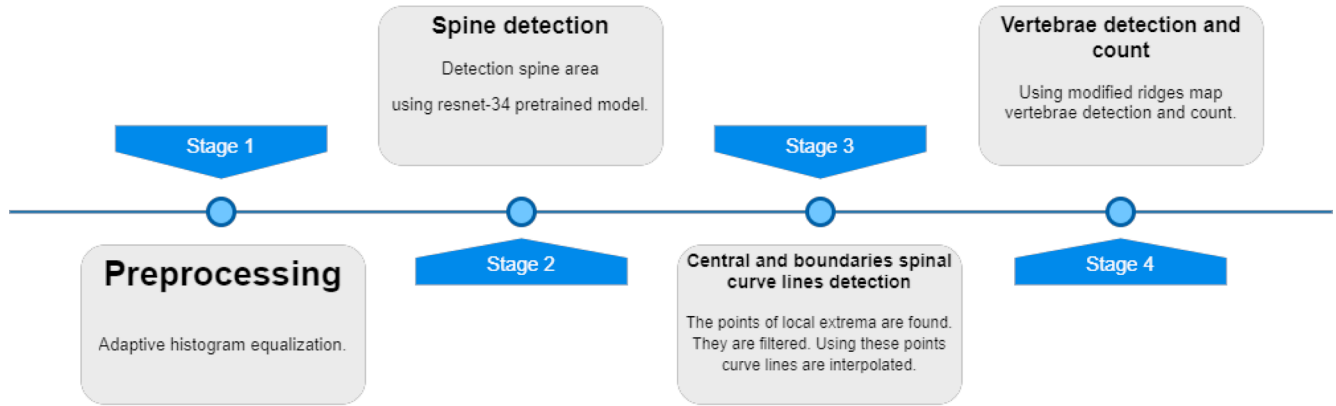


Fig. 5. General scheme of the X-ray image hardness determination algorithm.

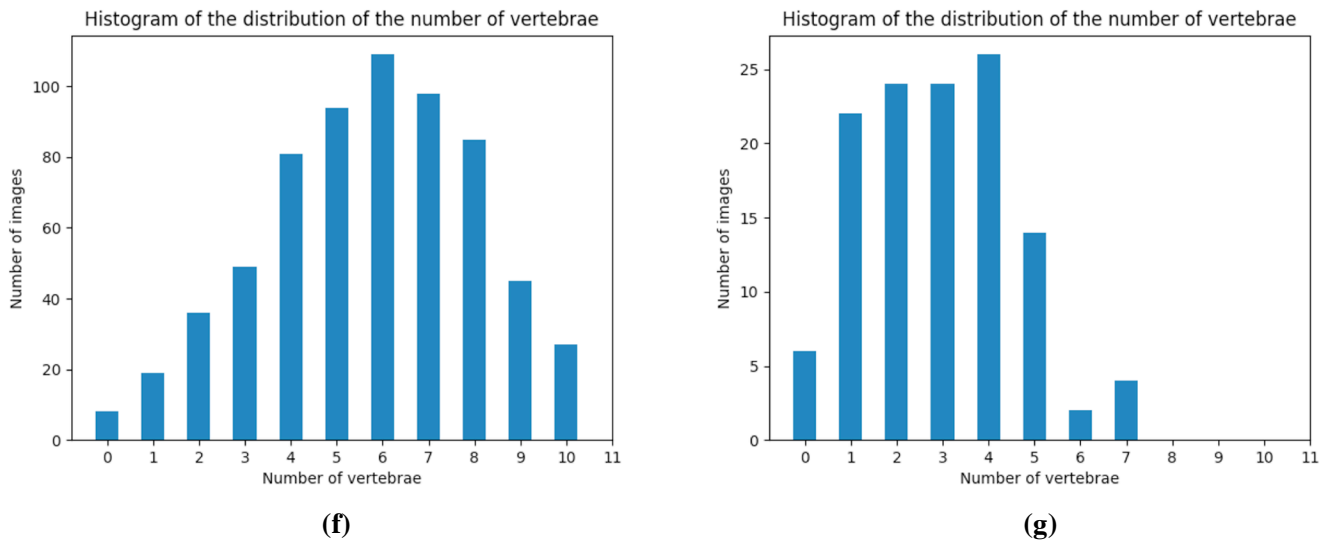


Fig. 6. (a) Histogram of the distribution of the number of vertebrae in the Shenzhen dataset, (b) Histogram of the distribution of the number of vertebrae in the Montgomery dataset.

The algorithm has detected a difference by the hardness of images in the two datasets (Fig. 6). The Shenzhen images displayed higher hardness than the Montgomery images, which is also corroborated by a visual analysis of the images. We also identified low-quality images (images on which vertebrae could not be detected). In most cases, these were images that suffered from digitization errors (Fig. 7).

Implementation Details

The program for the proposed method was implemented in Python 3. The machine learning models were implemented in the deep-learning framework pytorch 1.8.1. The Pillow 8.2.0 library was used for image manipulation. Graphs were constructed with matplotlib. We also used the library scipy 1.6.3. Model training was performed on Intel Core i7-9750H CPU 2.60GHz and 16 GB RAM, with NVIDIA GeForce RTX 2060 6 GB graphic accelerator.

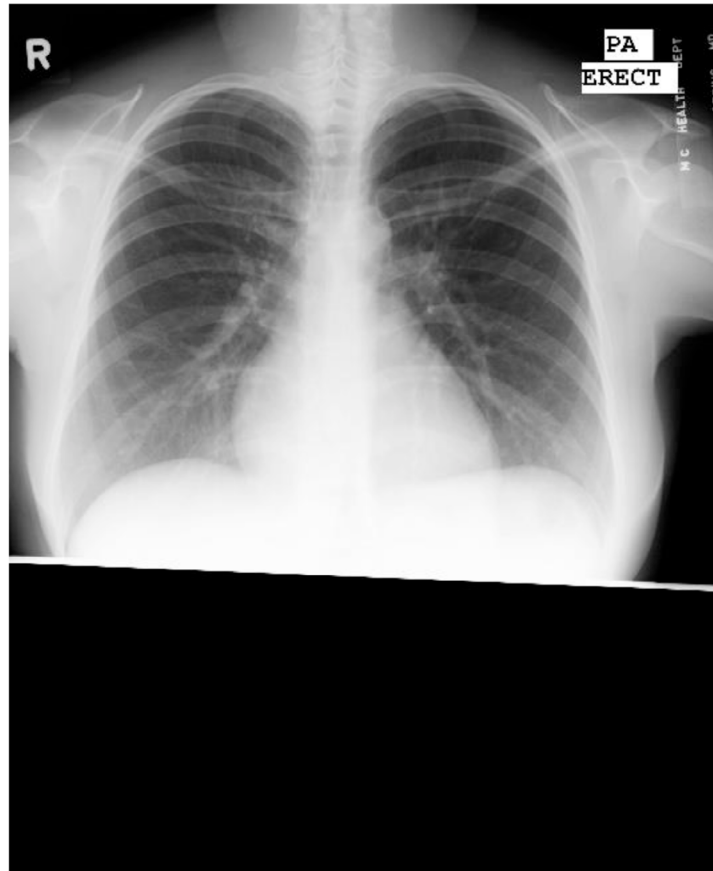


Fig. 7. Example of a low-quality image from the Montgomery dataset. The lung X-ray image shows a large region without useful information.

CONCLUSION

In this article, we have developed, implemented, and tested a method for automatic determination of the hardness of X-ray images by counting the thoracic vertebrae. This method successfully separates “soft” from “hard” X-ray images. The application of this method coupled with neural-network classifiers should increase the effectiveness of neural-network classifiers by producing a more uniform hardness distribution of the images in the training sample.

Research supported by Russian Foundation for Basic Research grant 19-57-80014 (BRICS2019-394).

REFERENCES

1. I. Fantini, et al., “Automatic MR image quality evaluation using a Deep CNN: A reference-free method to rate motion artifacts in neuroimaging,” *Computerized Medical Imaging and Graphics*, **90**, 101897 (2021).
2. Q. Zhang, et al., “Deep learning with attention supervision for automated motion artefact detection in quality control of cardiac T1-mapping,” *Artificial Intelligence in Medicine*, **110**, 101955 (2020).
3. K. Nousiainen, et al., “Automating chest radiograph imaging quality control,” *Physica Medica*, **83**, 138–145 (2021).
4. R. O. Panicker, et al., “A review of automatic methods based on image processing techniques for tuberculosis detection from microscopic sputum smear images,” *J. Medical Systems*, **40**, No. 1, 1–13 (2016).

5. R. D. J. Samuel and B. R. Kanna, "Tuberculosis (TB) detection system using deep neural networks," *Neural Computing and Applications*, **31**, No. 5, 1533–1545 (2019).
6. A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," in: *Pattern Analysis and Applications* (2021), pp. 1–14.
7. M. E. H. Chowdhury et al., "Can AI help in screening viral and COVID-19 pneumonia?" *IEEE Access*, **8**, 132665–132676 (2020).
8. I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, **43**, No. 2, 635–640 (2020).
9. L. A. Timofeeva, T. N. Aleshina, and A. V. Bykova, *Syndromes of Lung Tissue Pathology. A Textbook* [in Russian], Izd. Chuvashia University, Cheboksary (2013).
10. L. S. Rosenstrauch and M. G. Vinner, *X-ray Diagnosis of Diseases of Respiratory Organs* [in Russian], Book on Demand, Moscow (2013).
11. M. H. Horng, et al., "Cobb angle measurement of spine from X-ray images using convolutional neural network," *Computational and Mathematical Methods in Medicine*, **2019** (2019).
12. S. Jaeger, et al., "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases," *Quantitative Imaging in Medicine and Surgery*, **4**, No. 6, 475 (2014).
13. <http://archive.nlm.nih.gov/repos/chestImages.php>.
14. V. Subramanian, *Deep Learning with PyTorch: A Practical Approach to Building Neural Network Models Using PyTorch*, Packt Publishing (2018).