Automatic system for the renal and cancer segmentation in CT images

T Les, T Markiewicz, M Dziekiewicz, M Lorent

Corresponding author : Tomasz Les , lest@ee.pw.edu.pl

1. Abstract

This article presents the concept of a complex system for automatic detection of kidneys and kidney tumors in computed tomography images. An effective treatment of cancer depends on a quick and effective diagnosis. Computer support for medical diagnostics is crucial in effective specialists' analysis. Automatic and accurate location, together with precise detection of the kidney and/or tumor contour is a demanding task. In this article, authors present a complex system for automatic detection of kidneys and kidney tumors, based on machine learning techniques, using the U-Net network. Convolutional neural network recognition results are then processed in multiple stages, using morphological processing, 3D model analysis, geometric coefficients analysis and region-growth implementation. The results of the system detection were compared to the reference images marked by an expert. The system presented in the article is characterized by a very high efficiency of recognition and segmentation of kidney and tumor areas.

2. Introduction and problem statement

Diagnosis of neoplastic lesions in the abdominal area requires a computed tomography examination (CTA). This test allows to detect changes in organs located inside the body. The CTA test is used very often in the detection and initial identification of kidney tumors. The result of a tomographic examination for a given patient is a set of digital gray-scale images, representing abdominal sections at different heights. More accurate devices allow an acquisition of more accurate images, and scans with high density. This is an obvious advantage in diagnostics, however at the same time it causes the diagnostics more demanding for humans. One case can be compounded of several hundred slices. In order to accelerate the diagnosis and increase the quality of a final recognition, it is necessary to prepare computer techniques for image analysis support. This article presents a system for automatic detection of kidney and tumor areas. The system developed by us uses machine learning techniques based on the U-Net network. The results of network classification are then processed by post-processing algorithms which increase the accuracy of the final results recognition.

3 Data selection and learning

The analysis process of the CTA data was realized in two steps due to the sharing procedure of the data in the challenge. First step was performed based on the first (learning) part of the data containing 210 cases. The second step was performed using the 90 testing cases, which have been shared before closing the challenge.

The initial examination of the shared CTA data has been made manually by the authors in the aim to recognize the specific cases of renal pathology, predefine the typical size of both the renal and tumor and their location on the CTA image. Moreover, we identified specific problems in the shared data set. We concluded that few CTA series have a lot of quality problems. In 14 of 210 (6.7%) cases, we found an inappropriate contrast phase, which can make the radiological examination of that CTA for renal pathology unreliable. Also, in three cases the artifacts can be noticed (e.g. steel implants after spine surgical intervention), in one case the image is blurred. Therefore, the authors decided to exclude these 18 cases (8.6%) from the experiments. It should be also noted that 28 cases have lower technical problems (mainly low contrast or high noise). Thus, 164 of 210 cases (78%) have a good (sufficient for scientific research) quality of renal CTA imaging. However, such a restricting attitude towards the data, which is the same as radiological diagnostics requirements would significantly limit the study. Finally, we decided that only 18 previously mentioned cases should be excluded from the research and the other 192 cases will be used in the experiments.

To perform the first tests and to optimize the renal and cancer recognition procedure (methods and parameters) the initial data set was divided into the learning and validating subsets. The selection was made randomly and, as a result, 96 cases created the learning group and other 96 cases the validation group. It should be noted that this selection omitted the spacial density of the CTA slices in the specific cases, as well as size and type of renal pathological lesion.

3.1 Requirements for input data in the Deep Learning technique

The U-Net neural network is a multi-layers network, in which the input layers represent the dimensionality of the input data. Therefore, the size of the data in any case must have specific dimensions. If the input data would be the raw CTA image, its dimensions would equal 512 x 512 x 1 (X size, Y size and number of channels). However, inputting the whole size of the CTA image seems to be inadequate due to the range of sizes of renal and pathological lesion as well as oversizing of U-Net network. In this paper we propose the iterative scheme of image classification which uses a moving frame of the specified size. This causes a problem of the proper selection of the frame's size. A single frame should be appropriate to represent the specific structure of renal, cancer (including noncancerous lesions in renal) or other organs treated as a background. A frame which is too small can cause that the region will be unspecified (e.g. interior of renal pyramid), and therefore probably impossible to properly recognize. During the examination of CTA series we concluded that the right size of the frame would be of at least 25 x 25 pixels. Another problem which appears is that the descending dimensions of sub sequential levels in the U-Net structure require that the input data size should be divisible by 2. Therefore, the frame size should be selected from the set of values {32, 48, 64, 80}. In our study the frame size was fixed to 48 x 48 pixels.

Another problem is a choice of proper grey scale representation. The raw CTA dicom images are stored in the integer values from minus 2000 to about 3000-4000. Although the full range is not used, rescaling them to the 8bit range decreases the ability to differentiate specific regions. In this study the original range of CTA gray level values was used as an input to U-Net or other techniques.

3.2 Scheme of moving flame pixel classification

The partition of the single CTA slice into the set of frames requires a proper choice of parameters. In our solution we move the frame horizontally and vertically with the selected displacement value (e.g. 6 pixels). In this case any pixel of image will be classified such times that will be included in the individual frames. Usually the bordering areas of the CTA slice represent the air or table and can be omitted in the sweeping by frame process. Preselection of an area of the expected location of the renal, cancer or other renal lesion allows choosing such a set of frames in which any pixel in this area will be classified a fixed number of times. For the earlier selected 48 x 48 pixel frame and 6 pixel displacement this value will be equal 64.

3.3 Scheme of learning of frame selection

The proper selection of the learning data is a crucial step for a good generalization of a neural network. Here we propose the scheme of frame selection, that will create the learning data. The first theorem that we take into account only the frames in that the specific recognized structure should be sufficient represents. In our case, it is related with the selection of frame location. For example, if a renal structure is covered by only a small corner of the frame, the ability of recognition by neural network of an area such as a renal is very limited. To avoid such data in the learning set of frames, a binary mask of centers of frames with sufficient representation of renal structures was created. This mask is the result of dilation of renal mask by the structuring element of disk shape with the radius equals 4 pixels. The learning frames are selected randomly with theirs centers in within such mask and the condition of the minimal interval between them equals 6 pixels. For each CTA slice with at least 500 pixels renal area unconnected with any cancer regions a six frames were selected if possible (under interval condition). In such an approach a "clear" renal data were created.

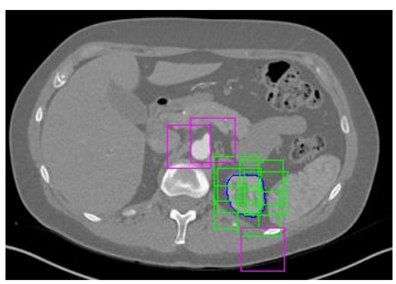

Fig x. An example of a mask frames selection.

The same procedure will be performed for the cancer region. In this case no condition of separation with other regions was applied, therefore these data contain representation of cancers, sometimes parts of renal and sometimes parts of background. However, the dominant of these frames are cancers and other renal lesions.

The last data set represents the other organs treated as a background. The frames are selected in the similar procedure as described for the renal regions, but the mask is created as a ring around the renal and cancer together. This ring is created by a division of dilation of renal/cancer mask by a disk with the radius equal 56 pixels and dilation of the same mask by a disk with the radius equal 26 pixels. In such case the background frames are separable with the renal and cancer regions. Due to the presence of some backgrounds in the renal and cancer frames, only 3 frames were selected from each CTA slice.

4. Kidney and cancer detection system

Kidney recognition is a multi-stage process. The first step is using the previously learned U-Net network. The U-net network returns the result in the form of a gray scale image. An important task is to choose the appropriate threshold value (K) in order to convert the result into a binary mask. For the proper selection of the optimal K value, a receiver operating characteristic (ROC) analysis was performed by examining the dependence of sensitivity and specificity. In figure x a ROC chart has been presented. On the horizontal axis, the threshold values K were determined. The analysis of many cases allowed the selection of the optimal K=32 value. In the further tests, this value was used for kidney and tumor detection.

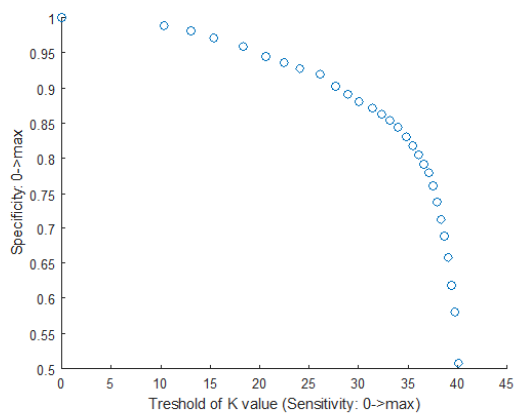

Fig x. ROC chart, allowing the selection of the optimal threshold value K.

An appropriate choice of threshold value K allows to obtain the mask of both kidneys for each image slice. Very often the result returned by the U-NET network contains errors of false positive (FP) and false negative (FN). An example of a kidney detection by the U-Net is shown in figure x, where TP and FN errors are clearly visible.

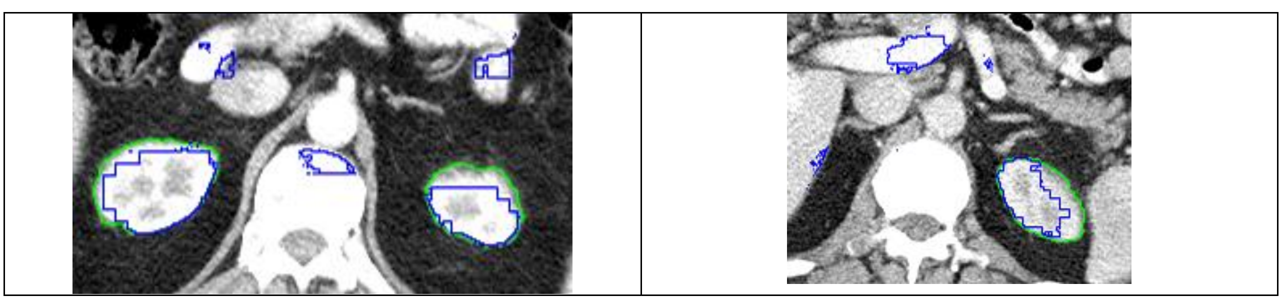

Fig x. The result of the kidney detection by an expert (green border) and by the U-NET network (blue border).

In most cases, the network was able to detect the kidney's boundary, but post-processing was necessary to increase the final result accuracy, as well as to remove the redundant areas.

4.1 Kidney and tumor post-processing method

The first step in post-processing algorithm is to merge all slices of analyzed case to generate a 3D model. Then, in the entire 3D model, the two largest coherent areas are selected. The selected areas define new kidney mask while other areas are removed. Further analysis is performed on each slice separately, only in two dimensions. The next step is to apply segmentation of left and right kidney separately. For this purpose, two points: p1 and p2 (y, x) [[280,117]; [280,395]] were determined. This points are most often located in the kidney area. The next step is to find the largest areas nearest points p1 and p2. The result of this operation are two binary masks, defining the left and right kidneys defined as R1a and R1b. Precise kidney segmentation is performed for both kidneys separately. For each area found, region-growth is performed. The result of this operation are binary masks R2a and R2b. The last step of the algorithm is to calculate the circularity coefficient for areas: R2a and R2b and for the areas detected by the U-Net network: R1a and R1b. Finally, the detected area, determining the exact boundaries of the kidney is selected based on the coefficient of circles. A complete kidney recognition system is shown in Fig x.

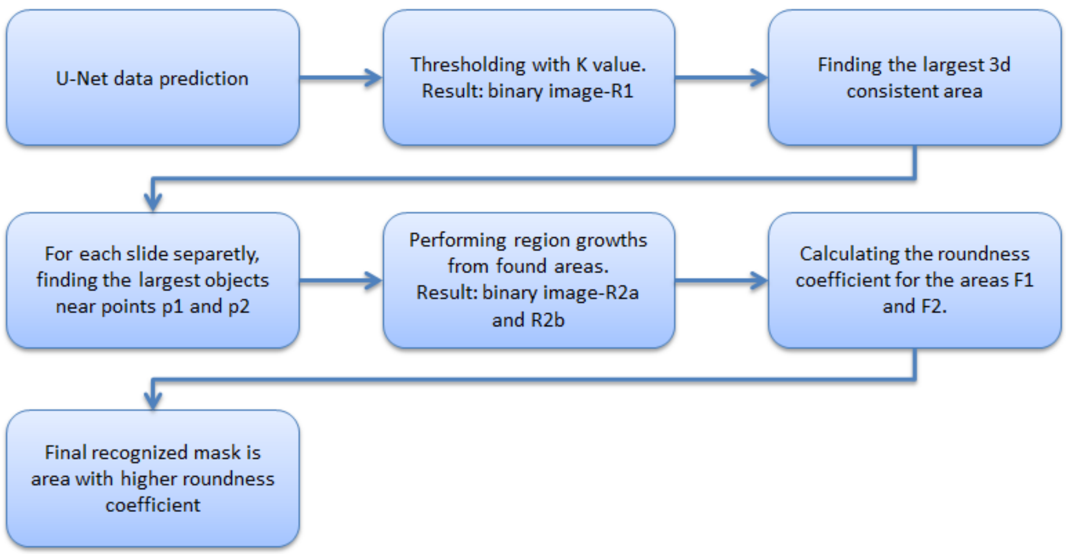

Fig x. Diagram of the complete kidney recognition system

The segmentation result of the example kidneys is presented in Fig x.

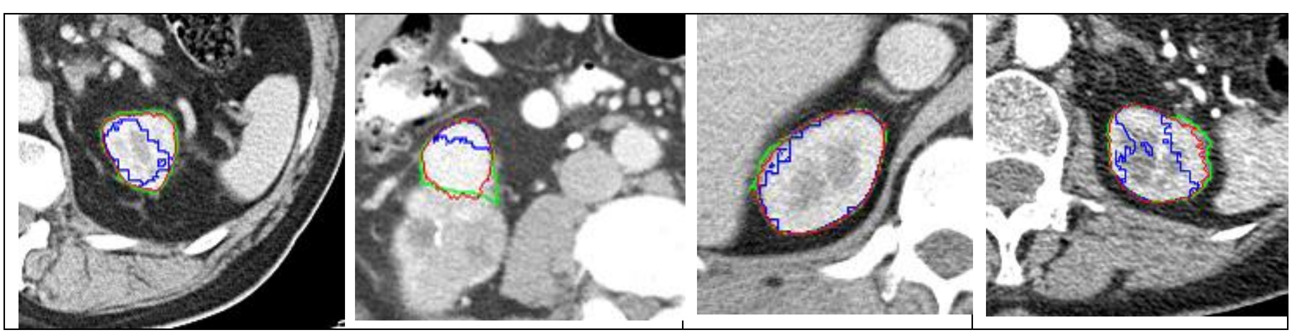

Fig x. The segmentation result of the example kidneys. The green boundary was marked by an expert. Blue boundary is the result of U-net network segmentation and the red color shows the final result of kidney segmentation.

The full process of precise kidney border segmentation is based on the result of classification by the U-net network and complex post-processing.

5. Tests and result

---

6. Conclusion

Machine learning using the U-Net network gives correct results of the location of the kidney and kidney tumor. Unfortunately, very often the detected area does not agree with the exact contour of the kidney, marked by expert. The result of the U-Net network is either overestimated or underestimated. Thus, the U-Net network result was treated as a preliminary processing for the final recognition of the kidney using a complex, multi-stage image processing techniques based on morphological algorithms, 3D model analysis, geometric coefficients analysis and region-growth implementation. This article presents the concept of combining all techniques into one consistent system. This system is characterized by a high recognition accuracy of kidney and tumor. Accuracy has been compared to reference images using ROC analysis. Implementation of the system will allow in practice to increase the speed and effectiveness of diagnoses made by specialist.

7. References