

U-Net based frames partitioning and volumetric analysis for kidney detection in tomographic images

Tomasz LES*

Faculty of Electrical Engineering, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warszawa, Poland

Abstract. This work presents an automatic system for generating kidney boundaries in computed tomography (CT) images. This paper presents the main points of medical image processing, which are the parts of the developed system. The U-Net network was used for image segmentation, which is now widely used as a standard solution for many medical image processing tasks. An innovative solution for framing the input data has been implemented to improve the quality of the learning data as well as to reduce the size of the data. Precision-recall analysis was performed to calculate the optimal image threshold value. To eliminate false-positive errors, which are a common issue in segmentation based on neural networks, the volumetric analysis of coherent areas was applied. The developed system facilitates a fully automatic generation of kidney boundaries as well as the generation of a three-dimensional kidney model. The system can be helpful for people who deal with the analysis of medical images, medical specialists in medical centers, especially for those who perform the descriptions of CT examination. The system works fully automatically and can help to increase the accuracy of the performed medical diagnosis and reduce the time of preparing medical descriptions.

Key words: kidney detection; medical image processing; U-Net; frames partitioning; volumetric analysis.

1. Introduction

This paper presents the possibilities of U-Net application as a classifier task in the problem of kidney diagnosis based on images obtained from a CT examination. Convolutional neural networks (CNNs) were defined in the 1990s by Le Cun [1]. Solutions based on the modeling of mathematical structures that perform signal processing have been the focus of attention over the last decade as a major effective tool for machine learning. The development of the new GPU graphics cards has contributed to the possibility of wide application in prediction and classification tasks in a wide variety of scientific and industrial fields. Today, neural networks are the subject of intense research, especially in the field of medical image processing [2–5]. The U-Net was designed for medical image processing in the Computer Science Department at the University of Freiburg, Germany [6].

Nowadays, various types of pathomorphological kidney lesions constitute one of the most serious disease groups. Renal neoplastic lesions pose a great challenge in diagnostics and treatment. Kidney cancer is characterized by a very high mortality rate. Kidney cancer is detected in about 400,000 new cases worldwide, according to the data from 2018, with an estimated mortality rate of 44% [7]. According to the World Health Organization (WHO) classification, kidney cancer includes the most common bright cell kidney cancer (65–70% of all kidney cancers) as well as multicellular cystic kidney cancer, papillary kidney cancer, chromophobic kidney cancer, and other relatively rare types of cancer [8]. Also, oncocytoma, angiomyolipoma,

and other non-cancerous lesions such as cysts are often diagnosed in the kidneys. Kidney diagnostics is based on insightful and time-consuming analysis of CT images. Some help in the diagnostic process and can be obtained by using computer image analysis methods. Over the past few years, many solutions for internal organs segmentation have been developed. Automatic kidney detection is one of the most important challenges for medical systems. Results of segmentation based on deformable models are presented in [9]. Solutions based on deformable models are available [10]. Solutions based on fuzzy segmentation are still being developed [11].

Among the systems based on machine learning, solutions based on U-Net [3, 12–17] should be considered. By using a convolutional network such as U-Net, it is possible to omit the feature extraction phase. The disadvantage of this type of solution is the limited possibility of parametrization of the learning data and high calculation requirements. Especially for medical data, the size of the network input data may be too large for a typical computer unit if the whole image is being loaded. This article presents a method of image partitioning into frames, thereby providing only a fragment of the image to the network input. The advantage of frame creation is the reduction in the amount of the learning data. The parts of the image for learning the network have been selected to achieve the best possible network generalization possibilities. The volumetric coherent area analysis technique is also presented in this work, which facilitates removing errors in the form of redundant areas generated by the neural network. The manuscript presents 2D and 3D volumetric analysis techniques. The results of numerical experiments of the analyzed solutions based on U-Net, Dense U-Net, and 3D volumetric analysis support are presented. The U-Net-based frames partitioning supported by volumetric analysis for kidney detection in tomographic images proposed in this paper is a novel solution, which can be considered in the

*e-mail: Tomasz.Les@pw.edu.pl

Manuscript submitted 2020-10-22, revised 2021-02-18, initially accepted for publication 2021-02-18, published in June 2021

construction of the modern medical diagnostic computerized systems.

Currently, complete, and automatic diagnosis of kidneys, as well as further classification of neoplastic lesions based on CT scans, is a major research challenge.

2. Materials

The study presented in this paper was conducted with the use of data provided by the Military Institute of Medicine in Warsaw. The CT slices used in the study are very different. The data include images representing many renal pathologies such as kidney tumors and others, cystic fibrosis, hydronephrosis, etc. The effectiveness of the diagnosis of a pathomorphological lesion depends on the type of lesion, its location, and image quality. In Fig. 1 we can see different slices of CT, showing a large variety of changes in the kidneys.

The images used in the study were prepared in the process of CT examination with a contrast medium. As the contrast can absorb X-rays, it facilitates shading the image obtained in the examination, which significantly improves its readability. Thanks to the use of a contrast medium, it is possible to visualize the details of the structure, both of whole tissues and individual organs. Despite the use of a contrasting medium, the automatic task is still difficult. Figure 1 shows six selected, different examples of slices, showing particular problems occurring in the analysis of this type of image. Example A in Fig. 1

shows the problem of the kidney (indicated by an orange arrow) touching the liver and highlights related difficulties in segmentation. The appropriate contrast adjustment is important in this case. Example B shows an unusual shape of the kidney; example C shows the final section of a small kidney; examples D and E show that the kidney is invisible or displaced. Example F shows both kidneys in the middle section with a regular shape and typical position.

The presented examples also show the differences in contrast intensity. It is also very difficult to precisely define the location of neoplastic lesions as well as the position of the kidneys in the whole CT image. Following the above observations, a scheme based on U-Net was proposed.

The study was performed with the collection of image data of patients from the Military Institute of Medicine in Warsaw. The data are anonymized and it is not possible to identify the individual patient based on the provided data. The image data is characterized by great diversity. Some cases are characterized by a different phase of contrast. The task of kidney detection is also difficult due to the major anatomical differences among the patients. The position of the kidneys may move within a certain range in the x, y, and z axes. The kidneys may also have different positions from each other. Therefore, it is not possible to establish a single fixed kidney search area. To assess the effectiveness of the kidney detection system proposed in this paper, 30 different cases were analyzed. Every single case consists of a series of single slices as a raw CT image. The number of slices for one case may vary significantly.

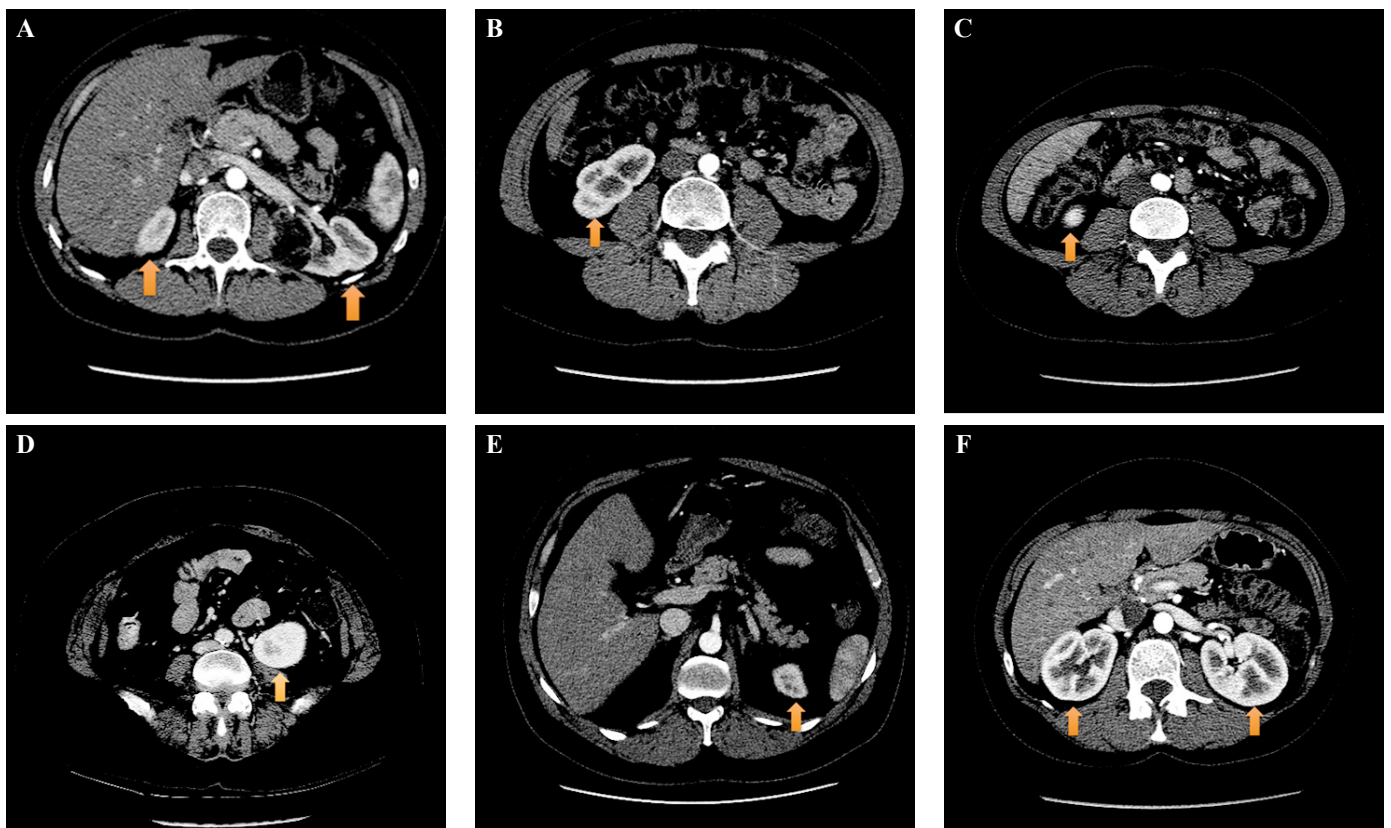


Fig. 1. Example of a large variety of kidneys; the brightness of the images has been automatically adjusted

The resolution in the patient's Z-axis depends on the slice thickness parameter, which determines how many mm of the section is represented. This parameter can vary from 1.25 mm to 5 mm depending on the examination. Also, differences may arise since some of the kidneys may have tumors, which increases the number of necessary scans. Sometimes the differences occur because of the anatomical structure, i.e. kidneys are never located at the same height.

3. Processing methods

This section presents the individual main stages of image processing in the developed system for kidney detection. The main stages of which the system consists include initial brightness correction, preparation of the U-Net model applying frames partitioning, U-Net classification, and volumetric analysis of coherent areas.

3.1. Image brightness correction. A raw, single CT slice, analyzed on a computer, is a grayscale image with a brightness range of about 4000 pixels. This is a number that cannot be visualized in the form of a typical image with a range of 256 pixels. For data processing by using mathematical morphology and U-Net techniques, this range is not narrowed because it would reduce the valuable information within the image. Only for the visualization of the output image for the presentation, the image is scaled to a range of 256. Typical automatic adjustment techniques fail in the case of CT images because they rely on the analysis of the image intensity in the entire field of view. In the case of CT image analysis, it was decided to adjust the brightness of the image based on pixel intensity analysis in 80×80 pixel frames, whose upper left corner is located at points $[x_1, y_1] = [110, 264]$ and $[x_2, y_2] = [319, 282]$.

These points were selected after analyzing 300 different slices to select a single frame that covers the largest part of the kidney. In the next step, the pixel values from both frames are averaged and scaled to a range of 1–256. In this way, a grayscale image is obtained, which emphasizes as much as possible the differences in contrast between the kidney and other organs (Fig. 2).

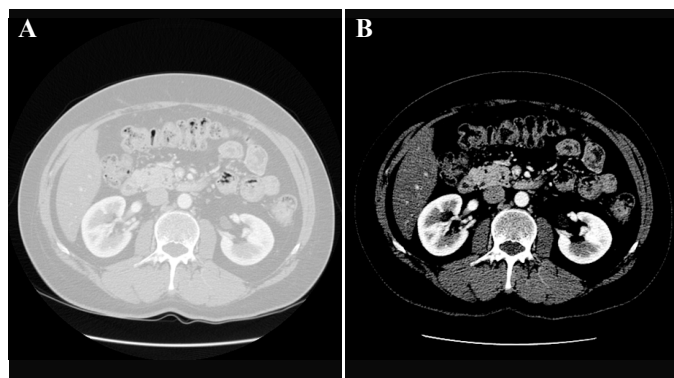


Fig. 2. A) Raw image scaled to a range of 0–255; B) Image after applying brightness correction

3.2. U-Net in medical image segmentation. Kidney detection was performed using the U-Net. The U-Net is a multi-layer network. In this network, the input layer represents the dimensions of the input data [18]. A typical CT image size is 512×512 pixels. Trying to input the full image size into a U-Net network is not optimal for a typical computing unit because of the kidney size. Because the CT image is represented by the dimensions $512 \times 512 \times 1$ (X size, Y size, and several channels) it was decided to divide the image based on patches that cover the selected kidney area to reduce the size of the learning data.

It was decided to select a set of 48×48 pixel frames from each slice belonging to the learning dataset. A mask was generated as a result of dilatation operation, using the structuring element of 4 px, from a mask created by an expert. Then, up to six frames were randomly selected, assuming that the distance between frames is 6 px and which region overlaps the mask.

Each layer is shown as a blue block and corresponds to the multi-channel representation of the output of the previous layer. The size (i.e. height \times width – $H \times W$) of the image representation is marked on the left side of the block, while the number of D channels is above the block. The white blocks correspond to the copied representation of the image from the deeper layer. Colored arrows mark the operations performed on the images and are described in the legend in the lower right corner of the diagram. The violet color indicates dropout operation, i.e. resetting randomly selected weights in the layer on which it was applied with 0.5 probability. The constructed model consists of:

- The input layer (single-channel image)
- Convolutional layers with ReLU activation function, filter size 3×3
- Sampling layers with a maximum 2×2 filter size operation
- Two layers of dropout
- Operations of copying output images from deeper layers
- Layers with sigmoid activation function, filter size 1×1
- Output layer (single-channel greyscale image)

U-Net was configured based on validation data according to the following parameters:

- The number of epochs: 20
- The learning algorithm: the stochastic gradient descent with momentum
- The minimum batch size: 256
- The gradient threshold: 0.05
- The initial learning rate: 0.05 (was changed between 0.05 and 0.06)
- The L2 regularization: 0.0002 (was examined from the values of 0.0001, 0.0002, and 0.0003),
- The momentum 0.95 (was selected from 0.92, 0.95 and 0.98)

3.3. Frame pixel classification. The possibilities of generalization of the neural network depend on the correct selection of learning data. In the learning process, data sets have been prepared in which the structure (kidney) is sufficiently representable. For this purpose, a binary mask was created, obtained in dilatation transformation using a 4-pixel radius disk structural element.

Pre-defined 48×48 pixel patches are randomly generated in a way that their centers are within the mask, with a minimum

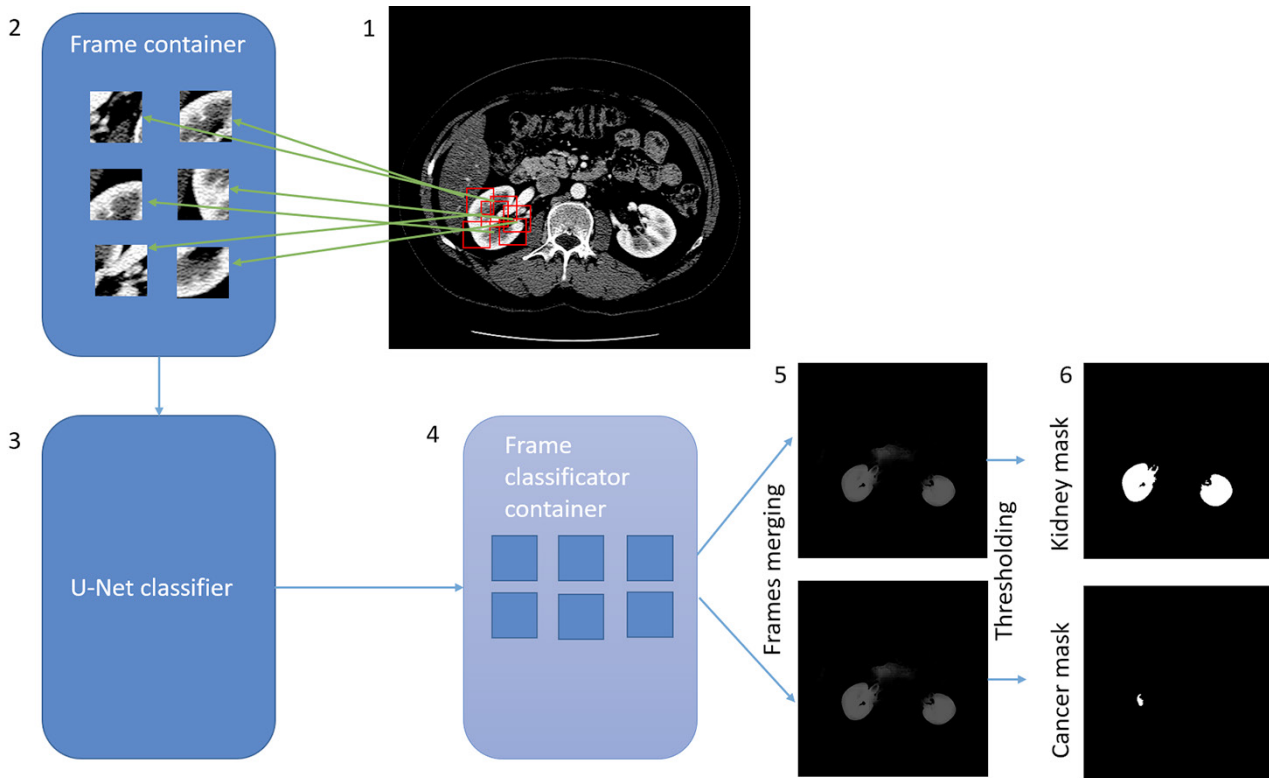


Fig. 3. The process of generating patches for the kidney; generated kidney patches are passed to the classifier input (1–3), then, the frames (4–5) are merged, the threshold is calculated and the final result is visualized (6)

distance between the patches of 6 pixels. Then, for each slice, six patches with an area of kidney greater than 500 pixels were selected. The final segmentation is a combination of patches representing the kidney. The algorithm is based on moving the frames horizontally and vertically every 6 px. For each pixel, which is a potential kidney fragment, the segmentation is performed 64 times, which is a result of a specific selection of frame size and offset.

A graphical example of the selection of patches is shown in Fig. 3.

The next step of the algorithm is to perform a threshold. The result of the neural network is in the form of an image in grayscale. Each pixel represents the number of votes cast for the fact that there is a kidney at a given point. To select the appropriate threshold, a Precision-Recall analysis was conducted.

The equilibrium point between a true positive ratio TPR (sensitivity) (1) and a positive predictive value PPV (precision) (2) was equal to $Th = 38$, based on 10 cases of validation set data.

$$TPR = \frac{TP}{TP + FN}, \quad (1)$$

$$PPV = \frac{TP}{TP + FP}, \quad (2)$$

where,

- TP is the number of pixels defined as a kidney by the system as well as by the expert.

- FP is the number of pixels defined by the system as background and by the expert as kidney.
- FN is the number of pixels defined by the system as a kidney and by the expert as a background.

Figure 4 shows the Precision-Recall curve of PPV and TPR. Based on the curve analysis, the threshold point was determined.

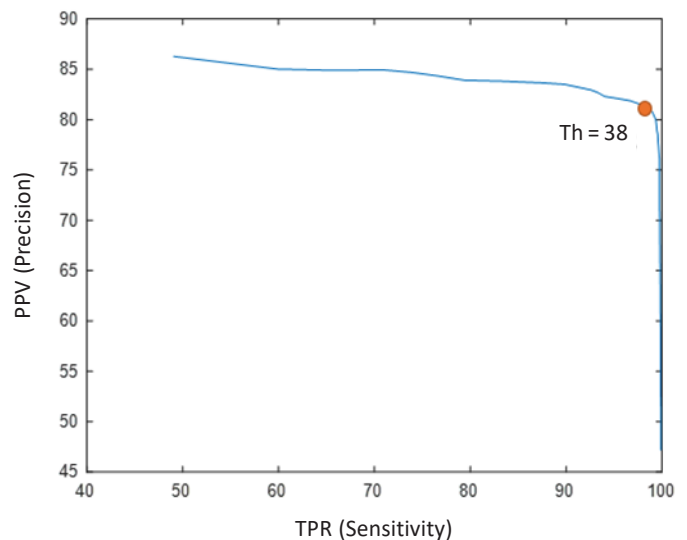


Fig. 4. Precision-Recall curve of TPR (Sensitivity) and PPV (Precision); based on the curve analysis the threshold point (Th) was determined

After the threshold is calculated, a binary image is obtained (Fig. 6). The result has errors of type I (false positive). To remove this type of error it is necessary to perform a volumetric analysis of coherent areas, described in Section 3.4.

3.4. Volumetric analysis of coherent areas. To eliminate FP-type errors, all coherent areas should be removed, leaving only the two largest ones (kidneys). The removal of coherent areas can be done in two dimensions (for each slice separately) or three dimensions (for all slices together).

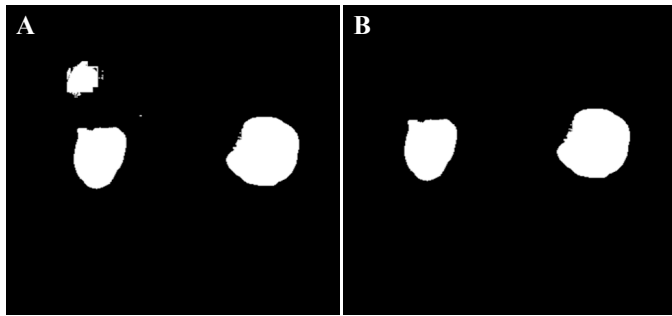


Fig. 5. A) Example result returned by U-Net after binarization with threshold value 38; B) Result of removing coherent areas

This type of analysis can also be performed in three-dimensional space. The same case is presented in Fig. 6 in 3D space.

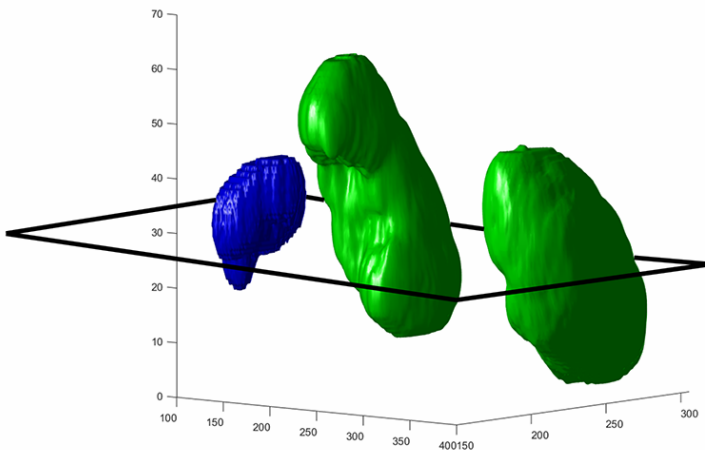


Fig. 6. A full 3D model using binary masks was obtained as a result of U-Net segmentation with threshold 38. The two largest areas are marked in green. The blue color shows the area to be removed. The black surface marks the slice visible in Fig. 5

In the case of two-dimensional analysis, it is necessary to calculate the area of all consistent areas and remove all, leaving the two areas with the largest area. The result of this operation is shown in Fig. 5.

Performing an analysis of consistent areas in three dimensions is a more complex operation. However, this type of anal-

ysis allows us to detect areas that are connected to the kidney in 2D space but are disconnected from the kidney in 3D space. The disadvantage of this solution is the need to calculate the results for all slices in advance.

3.5. The complete system of kidney detection. This section presents the concept of a complete kidney detection system in CT images. The first step is to construct frames covering the areas of kidney interest. Next, the U-Net model should be generated based on the frames. The next step is to perform a threshold, using a previously determined value. The last step is to analyze the consistent areas to remove redundant areas. A diagram of the major steps of the system is presented in Fig. 7.

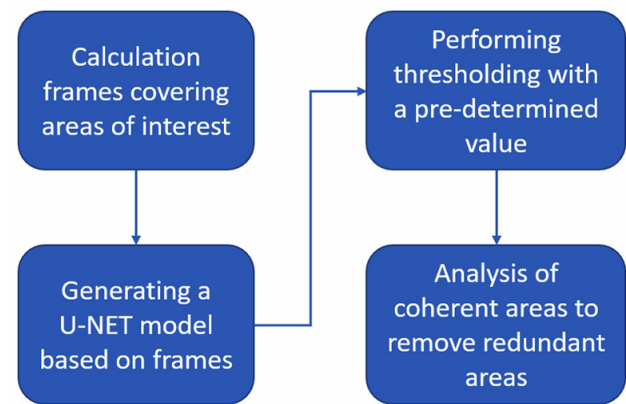


Fig. 7. Diagram showing the four main stages of the presented kidney detection system

4. Results

To assess the quality of the system, numerical tests were performed. The coefficients were calculated for each slice: TP, FP, FN, and F1-score (F1) (3).

$$F1 = \frac{2 * TP}{2 * TP + FP + FN} \cdot 100\%, \quad (3)$$

F1-score is a measure widely accepted and used in publications for the evaluation of the possibilities of the detection and segmentation of medical images. Also, the Hausdorff metric (d_H) was calculated for each pair of expert and system-generated masks (4).

$$d_H(X, Y) = \max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\} \quad (4)$$

Both F1-score and d_H were calculated in a pixel-wise way by comparing the kidney mask created by the expert (X) and the mask generated by the system (Y).

In the experiments, 30 cases were used for cross-validation tests. Cases were divided into 10 groups in which 27 cases were selected as the training data and 3 cases were selected as the testing data. Each case is composed of 23 to 92 slices. In total, 1,447 slices were analyzed. The numerical results of the calculated F1-score measure of the tests are presented in Table 1.

Table 1

Numerical results of automatic kidney detection based on F1-score, standard deviation (SD), and confidence interval (CI)

Group	Kidney detection		
	Average F1-score SD (CI)	CI range	d_H
1	89.22 ± 14.74 (85.42–93.03)	7.61	10.57
2	90.44 ± 6.78 (88.34–92.54)	4.20	16.83
3	85.99 ± 13.57 (83.16–88.81)	5.65	9.39
4	89.90 ± 11.50 (87.60–92.21)	4.61	11.95
5	92.32 ± 6.66 (90.98–93.65)	2.67	14.93
6	90.34 ± 11.54 (88.03–92.66)	4.63	8.48
7	90.27 ± 11.61 (87.95–92.60)	4.65	10.74
8	87.93 ± 13.82 (85.16–90.70)	5.54	16.53
9	88.34 ± 15.31 (85.15–91.53)	6.38	14.94
10	76.42 ± 9.23 (70.56–82.28)	11.72	16.85
average	88.12 ± 11.48 (85.24–91.00)	5.77	13.12

The numerical results presented in Table 1 show that the average F1-score in kidney detection is 88.12. The standard deviation has also been calculated for each group, which varies from 6.66 to 15.31. Large values of the standard deviation result from the fact that the kidney area varies according to the Z-axis. The surface area counted in the middle section pixels is usually the largest, while the surface area of the end sections (e.g. first and last) is the smallest. When calculating the F1-score measure, even a small error (about a few pixels) results in a large decrease in the F1-score. To illustrate the results better, a confidence interval measure (CI 95%) was added, rejecting outliers. The average CI range is 5.77 and most F1-score values are between 85.24 and 91.00.

The experiments were performed on a computer equipped with a MATLAB environment.

Tests were performed on a computer with Intel® Core™ i7–7820HQ, 2.90 GHz, 4 Cores, 16 GB installed physical memory, and graphics card: NVIDIA Quadro M1200.

Figure 8 shows examples of kidney detection results. The red color indicates a kidney boundary generated fully automatically by the system. The blue color indicates a kidney boundary generated by a human expert.

The possibilities of the system in kidney detection with 3D reconstruction are presented in Fig. 9. Image results were generated after removing all consistent areas and leaving the two largest ones. The results were compared with the kidney reconstruction generated using masks created by a human expert.

The evaluation of the influence of volumetric analysis of coherent areas on kidney detection capabilities was examined by calculating the F1-score for both systems. Additionally, the

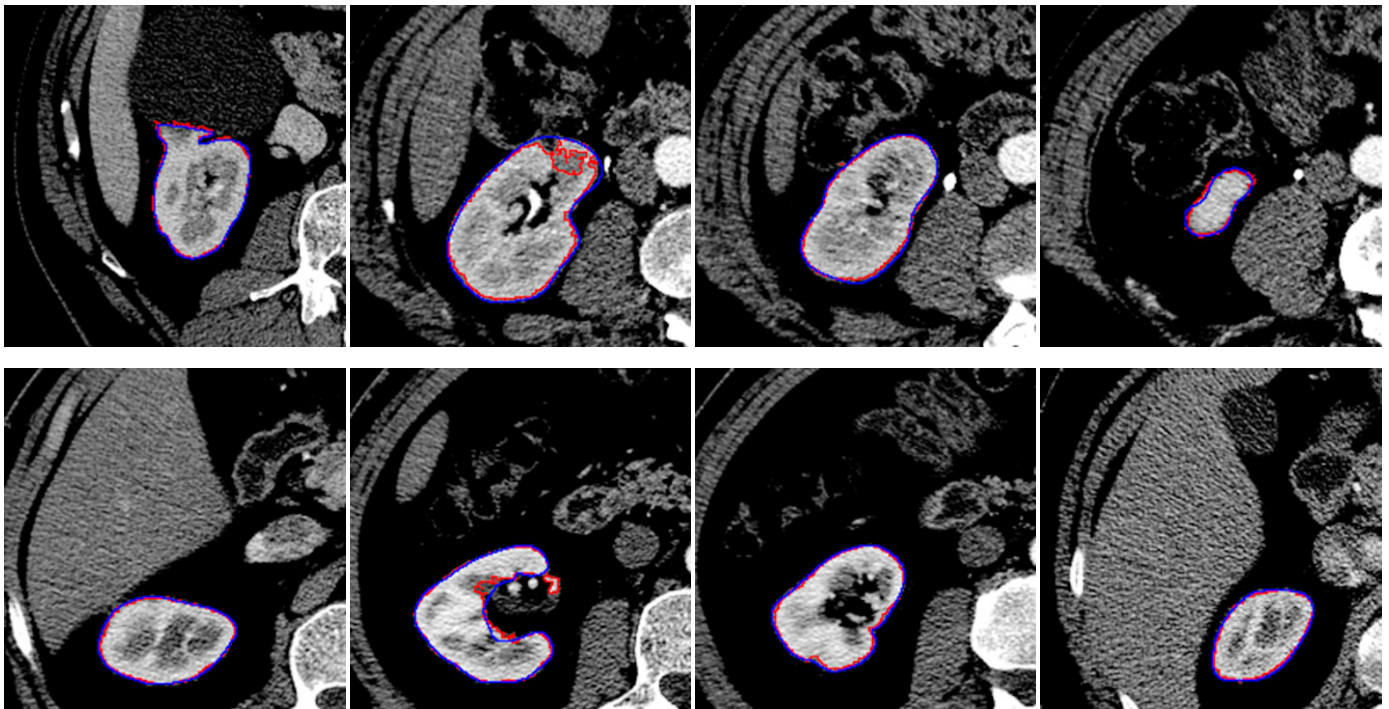


Fig. 8. Selected results showing the possibilities of kidney detection. The red color indicates a kidney boundary generated fully automatically by the system. The blue color indicates a kidney boundary generated by a human expert

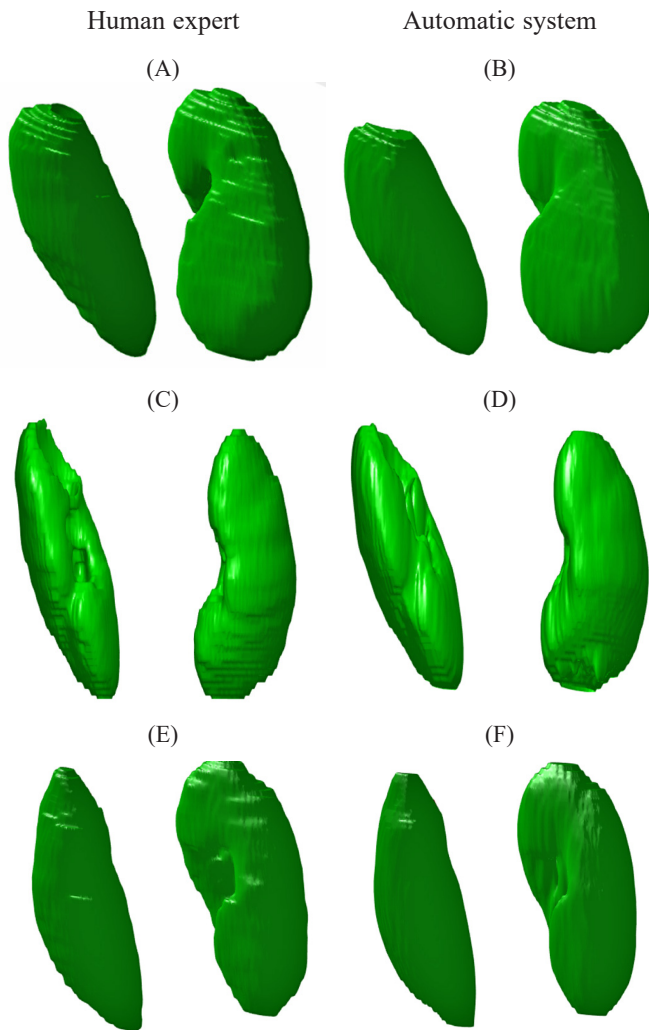


Fig. 9. Selected results showing the possibilities of kidney detection. The kidney area is marked in green. Examples A, C, E are generated using human expert data. Examples B, D, F are generated by an automatic system

results were compared to an alternative architecture: Dense U-Net. The results of kidney detection for 10 cases are presented in Table 2.

Table 2

Numerical results showing the impact of volumetric analysis of coherent areas on automatic kidney detection possibilities

System	F1-score	
	3D volumetric analysis	2D analysis
U-Net: 3D Volumetric analysis	90.84	87.85
Dense U-Net: 3D Volumetric analysis	88.48	84.73
U-Net: Without coherent areas analysis	85.39	
Dense U-Net: Without coherent areas analysis	84.94	

The implementation of volumetric analysis of coherent areas significantly improves the efficiency of kidney detection in both systems. Sometimes, however, areas connected by a “thin line” occur, as shown in Fig. 10. In such cases, the 3D coherent analysis fails. This error is not significant in the 2D analysis where each slice is treated separately.

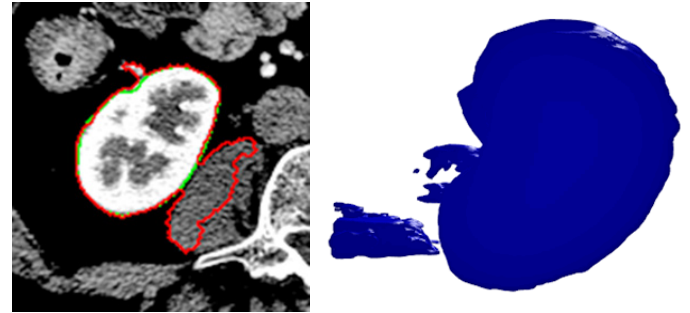


Fig. 10. Examples of areas connected by a “thin line” where 3D coherent analysis fails

5. Discussion

This paper presents a kidney detection system for CT images. The system uses U-Net and morphological operations to generate a kidney boundary fully automatically. Processing all CT slices also facilitates the generation of a 3D model of the kidney. The work presents four main stages of data processing: frame construction, U-Net model generation, thresholding, and volumetric analysis of coherent areas. The advantage of frame creation is the reduction of the size of learning data; however, it increases the complexity of the final detection process due to the necessity of a combination of patches representing the kidney. Another advantage of the proposed system is the ability to parallel the analysis of many images at the same time. Data processing can be performed in any order without the need for analysis of other slices. This is a significant difference from the earlier presented approach [19], in which some incorrectly recognized previous slices might result in deterioration of the analysis process. The proposed method facilitates avoiding this problem resulting from inappropriate data acquisition or transmission failures.

Additionally, it was observed that slices, even for the same patient, could differ significantly in brightness and sharpness. An independent analysis of slices is an advantage in opposition to the system proposed in [19], which required pre-calculation of the previous slice to generate a mask response for the actually analyzed slice.

In the proposed system, instead of the whole image, only a fragment is passed to the network model. An important stage of systems based on neural networks is to choose the appropriate threshold value. In this work, a fixed threshold has been chosen based on the precision-recall analysis. Though the determined value is optimal from a global point of view by analyzing multiple slices, the calculated value is not always

optimal for a single slice. This is reflected in the FP-type errors. Elimination of this error has been realized by implementing the volumetric algorithm of a coherent area analysis. As a result, all consistent areas were removed from the image, leaving the two largest ones, i.e. two kidneys. The analysis of a coherent area can be done in 2D for every slice separately or in 3D. Numerical results confirm that a 3D volumetric coherent area analysis is more effective.

The limitations of the system are errors when using a 3D coherent analysis including “thin lines”. Such errors are not significant when removing the disconnected areas for each slice separately. In the future, consideration should be given to fusion combining the advantages of a 3D volumetric analysis and a 2D coherent area analysis.

Further work is also necessary, especially on the dynamic selection of threshold values, individual for each slice, which will reduce the FP errors. Computer image segmentation methods are currently one of the biggest challenges in medical technology [3, 9, 20]. Neural network solutions are a valuable direction for development. U-Net-based diagnostic support systems are characterized by high accuracy of 85–95% [12, 21, 22].

The numerical results of the presented system confirm the high efficiency of kidney boundary recognition at F1-score: $88.12 \pm 11.48\%$.

It should be emphasized that the best results of medical image segmentation can be obtained only by combining neural network techniques with morphological image processing. Therefore, there is still a need to develop new techniques for medical image processing.

6. Conclusions

This study presents a system for automatic kidney detection for a set of CT images. The system can automatically generate a contour of the kidney and build a 3D model. The main idea of the system is based on semantic U-Net segmentation, frame partitioning, and volumetric coherent area analysis. The performed numerical experiments confirm the high efficiency of the proposed solution. The presented solution can be applied in medical centers and can support radiologists and specialists involved in the image description.

Acknowledgments. This work has been supported by the National Science Centre (2016/23/B/ST6/00621 grant), Poland.

Special thanks to Doctors from the Military Institute of Medicine in Warsaw, especially Wojciech Kozłowski, Mirosław Dziekiewicz, Małgorzata Lorent, Marzena Cichowicz for sharing their knowledge, experience, and support in medical consultations.

REFERENCES

- [1] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition”, *Proc. IEEE* 86(11), 2278–2324 (1998), doi: 10.1109/5.726791.
- [2] F. Isensee, “An attempt at beating the 3D U-Net”, ed. K.H. Maier-Hein, 2019.
- [3] Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, T. Brox, and O. Ronneberger, “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation”, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, 424–432, Springer International Publishing, 2016.
- [4] C. Li, W. Chen, and Y. Tan, “Render U-Net: A Unique Perspective on Render to Explore Accurate Medical Image Segmentation”, *Appl. Sci.* 10(18), 6439 (2020), doi: 10.3390/app10186439.
- [5] Z. Fatemeh, S. Nicola, K. Satheesh, and U. Eranga, “Ensemble U-net-based method for fully automated detection and segmentation of renal masses on computed tomography images”, *Med. Phys.* 47(9), 4032–4044 (2020), doi: 10.1002/mp.14193.
- [6] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, ArXiv, abs/1505.04597, 2015.
- [7] M.E.J. Ferlay, F. Lam, M. Colombet, L. and Mery. “Global Cancer Observatory: Cancer Today.” [Online] Available: <https://gco.iarc.fr/today>, accessed (accessed).
- [8] P.A. Humphrey, H. Moch, A.L. Cubilla, T. M. Ulbright, and V.E. Reuter, “The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part B: Prostate and Bladder Tumours”, *Eur. Urol.* 70(1), 106–119 (2016), doi: 10.1016/j.eururo.2016.02.028.
- [9] D.L. Pham, C. Xu, and J.L. Prince, “Current Methods in Medical Image Segmentation”, *Ann. Rev. Biomed. Eng.* 2(1), 315–337 (2000), doi: 10.1146/annurev.bioeng.2.1.315.
- [10] B. Tsagaan, A. Shimizu, H. Kobatake, and K. Miyakawa, “An Automated Segmentation Method of Kidney Using Statistical Information”, in *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2002*, pp. 556–563, Springer Berlin Heidelberg, 2002.
- [11] J.C. Bezdek, “Objective Function Clustering”, in *Pattern Recognition with Fuzzy Objective Function Algorithms*, pp. 43–93, Boston: Springer US, 1981.
- [12] K. Sharma *et al.*, “Automatic Segmentation of Kidneys using Deep Learning for Total Kidney Volume Quantification in Autosomal Dominant Polycystic Kidney Disease”, *Sci. Rep.* 7(1), 2049 (2017), doi: 10.1038/s41598-017-01779-0.
- [13] P. Jackson, N. Hardcastle, N. Dawe, T. Kron, M.S. Hofman, and R. J. Hicks, “Deep Learning Renal Segmentation for Fully Automated Radiation Dose Estimation in Unsealed Source Therapy”, *Front. Oncol.* 14(8), 215, (2018), doi: 10.3389/fonc.2018.00215.
- [14] C. Li, W. Chen, and Y. Tan, “Point-Sampling Method Based on 3D U-Net Architecture to Reduce the Influence of False Positive and Solve Boundary Blur Problem in 3D CT Image Segmentation”, *Appl. Sci.* 10(19), 6838 (2020).
- [15] A. Myronenko and A. Hatamizadeh, “3d kidneys and kidney tumor semantic segmentation using boundary-aware networks”, arXiv preprint arXiv:1909.06684, 2019.
- [16] W. Zhao, D. Jiang, J. P. Queraltá, and T. Westerlund, “Multi-Scale Supervised 3D U-Net for Kidneys and Kidney Tumor Segmentation”, arXiv preprint arXiv:2004.08108, 2020.
- [17] W. Zhao, D. Jiang, J. Peña Queraltá, and T. Westerlund, “MSS U-Net: 3D segmentation of kidneys and tumors from CT images with a multi-scale supervised U-Net”, *Inform. Med. Unlocked* 19, 100357 (2020), doi: 10.1016/j.imu.2020.100357.
- [18] Y. LeCun and Y. Bengio, “Convolutional networks for images, speech, and time series”, in *The handbook of brain theory and neural networks*, pp. 255–258, MIT Press, 1998.

U-Net based frames partitioning and volumetric analysis for kidney detection in tomographic images

- [19] T. Les, T. Markiewicz, M. Dziekiewicz, and M. Lorent, “Kidney Boundary Detection Algorithm Based on Extended Maxima Transformations for Computed Tomography Diagnosis”, *Appl. Sci.* 10(21), 7512 (2020), doi: 10.3390/app10217512.
- [20] Z. Swiderska-Chadaj, T. Markiewicz, J. Gallego, G. Bueno, B. Grala, and M. Lorent, “Deep learning for damaged tissue detection and segmentation in Ki-67 brain tumor specimens based on the U-net model”, *Bull. Pol. Acad. Sci. Tech. Sci.* 66(6), 849–856 (2018).
- [21] W. Wieclawek, “3D marker-controlled watershed for kidney segmentation in clinical CT exams”, *Biomed. Eng. Online* 17(1), 26 (2018), doi: 10.1186/s12938-018-0456-x.
- [22] T. Les, “Patch-based renal CTA image segmentation with U-Net”, in *2020 IEEE 21st International Conference on Computational Problems of Electrical Engineering (CPEE)*, Poland, 2020, pp. 1–4, doi: 10.1109/CPEE50798.2020.9238735.