Hindawi
Scientific Programming
Volume 2021, Article ID 1885512, 13 pages
https://doi.org/10.1155/2021/1885512

Research Article

An Autocontouring Method for the Kidneys Using an Adaptive Weighted Multikernel Support Vector Machines

Yi Gu[1] **and Bo Li**[1,2]

[1]*School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, Jiangsu, China*
[2]*Jiangsu Key Laboratory of Media Design and Software Technology, Jiangnan University, Wuxi 214122, Jiangsu, China*

Correspondence should be addressed to Yi Gu; 8202101437@jiangnan.edu.cn

Received 19 April 2021; Accepted 5 July 2021; Published 14 July 2021

Academic Editor: Chenxi Huang

In radiotherapy, the location of the target area is very important. If the target area is small, the treatment is not complete, so the location of the target area is generally larger than the actual cancerous site. However, the damage of radiotherapy to normal cells is the same. In order to reduce the damage to the body as much as possible, we need to complete the most suitable target area. This paper uses an adaptive weighted multikernel support vector machine, which solves the parameter problem in the traditional multikernel support vector machine. The new AW-SVM can adjust the kernel weights adaptively. We completed our experiment on the abdominal MR dataset, using DSI as an evaluation indicator, and the experimental results showed its excellent classification performance. The minimum value of DSI in all results is 0.9654 (more than 0.7 is acceptable).

1. Introduction

In 2020, it is estimated that there will be 73750 new cases of kidney and renal pelvis cancer patients in the United States, of which 45520 are males, 28230 are females, and 14830 are expected to die, including 9860 males and 4970 females. In the five years from 2012 to 2016, the incidence of kidney and renal pelvis cancer was 16.6 per 100,000 people, 22.5 for men and 11.5 for women. In 2013–2017, kidney and renal pelvis cancer has a mortality rate of 3.7 per 100,000 people, 5.4 for men and 2.7 for women [1].

Adaptive radiation therapy (ART) is still the main method to solve cancer, especially for patients with advanced cancer [2]. Broadly speaking, any technology that adjusts the treatment process through feedback can be included in the scope of ART, such as image-guided radiation therapy (IGRT) [3], dose-guided radiation therapy (DGRT) [4], and structure-guided radiation therapy (SGRT) [5]. IGRT can be described as the primary stage of ART. It adds the concept of time factor on the basis of three-dimensional radiotherapy and fully considers the movement of anatomical tissue during the treatment process, and the displacement error of the divided treatment time, such as breathing and peristaltic movement, daily pendulum position errors, and target area contraction, causes changes in the radiation dose distribution and imaging of the treatment plan. Before the patient undergoes treatment, various types of advanced imaging equipment are used to monitor the tumor and surrounding risk areas in real time and can adjust the treatment conditions according to the changes of the organs so that the irradiation field only follows the target area so that it can achieve precise treatment in the true sense. DGRT is proposed on the basis of IGRT. In addition to comparing image data, DGRT also compares the actual absorbed dose of the tumor and surrounding normal tissue during treatment with the dose in the treatment plan and adjusts the treatment plan in time. In general, adaptive ART is a self-responsive, self-correcting dynamic closed-loop system from diagnosis location, plan design, and treatment implementation to verification. Compared with computed tomography-guided ART (CT-ART), the main feature of MR-ART is that there is no radiation damage and no bone artifacts and it can perform multifaceted and multiparameter imaging, has a high degree of soft tissue resolution, can display vascular structures without the use

of contrast agents, and have multisequence biological function images [6, 7].

With the development of three-dimensional conformal radiotherapy and intensity-modulated radiotherapy in recent years, more and more researchers pay attention to the accuracy of tumor target definition to minimize the damage to organs at risk (OARs) [8]. The traditional circle drawing of the target area completely relies on manpower to complete it manually. Because manual contour drawing involves knowledge of multiple departments and engineering technology, it is extremely difficult for general doctors to draw. It is done by an experienced imaging surgeon or radiation oncologist. In most cases, it needs to be jointly discussed by multiple experts in the imaging department and the oncology department. However, there is also a problem. The manual contour is almost impossible to reproduce, and the target area contour depends entirely on the knowledge level of the experts and the image quality. There are also deviations between different experts. Manual contour drawing is also very time-consuming and labor-intensive. In the face of high-dimensional multimode images, as the images increase, the working pressure increases, which may cause contour deviations [9]. In addition, the accuracy of manual contours needs to be increased. It is unrealistic to use manual contours for online MR-ART, so it is necessary to develop an online automatic contour algorithm for MR-ART.

Broadly speaking, the current image segmentation algorithms mainly include segmentation methods based on edge detection, threshold, region growth, statistics, graph theory, information theory, fuzzy set theory, knowledge, and so on and segmentation methods based on convolutional neural networks.

Early image segmentation methods were mainly based on point, line, and edge detection and segmentation methods, using Robert operator [10], Canny operator [11], Sobel edge detection operator [12], and so on. Statistics-based segmentation methods include unsupervised k-means clustering, FCM clustering, and Markov random field, and the other is supervised support vector machine (SVM) [13], naive bayes (NB) [14], and random forest model [15]. Compared with the unsupervised segmentation algorithm, the supervised one obtains certain prior knowledge through training and performs better in image segmentation. But unfortunately, in the face of complicated and ambiguous organ boundaries, these algorithms have limited success in judging the boundaries.

With the continuous maturity and improvement of deep learning technology, the experimental results gradually exceed the previous machine learning, relying on the strong feature extraction ability of convolution neural network. The larger the amount of data given in the training process, the stronger the generalization performance of the network. In terms of image segmentation, FCN (Fully Convolutional Networks) [16] and U-net [17] are outstanding. FCN classifies images at the pixel level, thus solving the problem of image segmentation at the semantic level. U-net performs particularly well in medical image segmentation. U-net can make full use of the low-level and high-level information of medical images to provide a basis for physical category recognition and precise segmentation and positioning.

The second chapter will introduce the dataset, data preprocessing, multiple kernel learning, and fixed-weight multikernel SVM to solve the problem of single-kernel inflexibility of SVM in image segmentation. Finally, we will introduce an adaptive weighted multikernel SVM to solve the weight problem of multikernel SVM. The third chapter will introduce the choice of kernel function, the evaluation index of the segmentation result, the comparison experiment setting, and the experiment process. The fourth chapter analyzes the experimental results, and the fifth chapter summarizes the full paper.

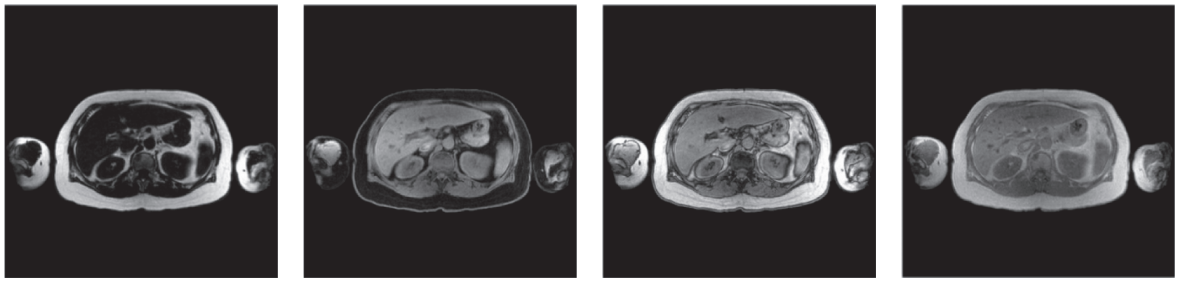
2. Related Work

2.1. Dataset. This experiment uses a dataset of eight patients with unresectable malignant tumors of the lower abdomen. We mark the eight patients as Sub1~8. Each patient's dataset consists of 16 pictures, and the first 15 pictures are in the process of treatment. The last one is the MRI image of the treatment day, the resolution is 370×370, the axial pixel spacing is $1.5\,\text{mm} \times 1.5\,\text{mm}$, and the plane spacing is 3 mm. The total dataset consists of a total of 126 images of eight patients. The kidney and skin contours in the dataset are all drawn by professional radiation oncologists as our automatic contour assessment.

In terms of feature extraction, this article only uses the most basic voxel features and three-dimensional space coordinates. The local texture features (voxels) corresponding to each sequence are extracted from the four sequences shown in Figure 1. In terms of three-dimensional coordinates, since the coordinates of each sequence are the same, the voxel coordinates of the IP sequence are selected in this experiment.

2.2. Data Preprocessing. In order to further improve the separability of the data, consider using a filtering algorithm to process the data. The filtering algorithm filters the image information, removes the pixels we do not want, and enhances the information we need. Because the lower abdomen organs/body fluids of the human body are also constantly in motion, it is considered to use Kalman filter [18] to increase the dimension of the data. After experimental verification and analysis, it is completely appropriate to use Kalman filter to process this MR dataset.

Kalman filter is often used in uncertain systems (robots, real-time systems, etc.). It is relatively rare in classification algorithms. Kalman filter can infer the state of the next step based on the state of the previous step. In our MR data, the voxels are all adjacent, which is associated with this feature of Kalman filtering. Similar to the idea of KNN algorithm, adjacent voxels have the same label. Through experimental comparison, after using Kalman filter for feature extraction, the separability of the data has been greatly improved. The experimental results can be seen in Table 1.



(a) (b) (c) (d)

Figure 1: (a) Dixon-FAT sequence, (b) Dixon-WATER sequence, (c) Dixon-OP sequence, and (d) Dixon-IP sequence. All four sequences are reconstructed from Dixon sequence.

Table 1: Performance comparison of classifiers using Kalman filter.

Part A: without a Kalman filter				
Datasets	AW-SVM	SVM	S4VM	MKL
Sub1	0.9621 ± 0.0148	0.9643 ± 0.0014	0.9348 ± 0.0050	0.9652 ± 0.0139
Sub2	0.9811 ± 0.0102	0.9770 ± 0.0032	0.9524 ± 0.0046	0.9768 ± 0.0091
Sub3	0.9731 ± 0.093	0.9712 ± 0.0013	0.9380 ± 0.0103	0.9725 ± 0.0124
Sub4	0.9875 ± 0.0115	0.9845 ± 0.0011	0.9525 ± 0.009	0.9794 ± 0.0121
Sub5	0.9848 ± 0.0122	0.9812 ± 0.0037	0.9553 ± 0.0033	0.9810 ± 0.0116
Sub6	0.9756 ± 0.0126	0.9746 ± 0.0021	0.9655 ± 0.0069	0.9727 ± 0.0112
Sub7	0.9818 ± 0.0116	0.9732 ± 0.0025	0.9610 ± 0.0026	0.9754 ± 0.0123
Sub8	0.9895 ± 0.088	0.9875 ± 0.0038	0.9507 ± 0.0113	0.9862 ± 0.0082
Datasets	RSM	TSVM	NB	RBF
Sub1	0.5549 ± 0.0036	0.7385 ± 0.0093	0.8646 ± 0.0070	0.6321 ± 0.0026
Sub2	0.5049 ± 0.0067	0.7886 ± 0.0028	0.8848 ± 0.0085	0.5894 ± 0.0037
Sub3	0.7060 ± 0.0082	0.7425 ± 0.0008	0.8667 ± 0.0071	0.5723 ± 0.004
Sub4	0.6415 ± 0.0058	0.7625 ± 0.0132	0.8949 ± 0.0039	0.6415 ± 0.0033
Sub5	0.6235 ± 0.0064	0.7625 ± 0.0116	0.9012 ± 0.0077	0.6149 ± 0.0037
Sub6	0.5936 ± 0.0081	0.8043 ± 0.0142	0.8980 ± 0.008	0.6512 ± 0.004
Sub7	0.5395 ± 0.0074	0.7389 ± 0.0092	0.8841 ± 0.0059	0.6035 ± 0.0034
Sub8	0.6043 ± 0.0098	0.7410 ± 0.0086	0.8468 ± 0.0073	0.9817 ± 0.0042
Part B: with a Kalman filter				
Datasets	AW-SVM	SVM	S4VM	MKL
Sub1	0.997 ± 0.0025	0.9966 ± 0.0015	0.9863 ± 0.0041	0.996 ± 0.0023
Sub2	0.9999 ± 0.0001	1	0.9888 ± 0.0048	0.9998 ± 0.0001
Sub3	0.9998 ± 0.0001	0.9995 ± 0.0003	0.9953 ± 0.0013	0.9995 ± 0.0004
Sub4	1	0.9999 ± 0.0001	0.9837 ± 0.0013	0.9999 ± 0.0001
Sub5	1	0.9999 ± 0.0001	0.9918 ± 0.021	1
Sub6	0.9999 ± 0.0001	1	0.9923 ± 0.0021	0.9997 ± 0.0002
Sub7	1	0.9999 ± 0.0001	0.9962 ± 0.0023	0.9994 ± 0.0003
Sub8	1	1	0.9835 ± 0.0026	1
Datasets	RSM	TSVM	NB	RBF
Sub1	0.8105 ± 0.0080	0.9646 ± 0.0075	0.9953 ± 0.0017	0.9817 ± 0.0034
Sub2	0.8641 ± 0.0031	0.9621 ± 0.0083	0.9267 ± 0.0084	0.9836 ± 0.0012
Sub3	0.9078 ± 0.0046	0.9784 ± 0.0071	0.9784 ± 0.0027	0.9951 ± 0.0034
Sub4	0.9033 ± 0.0069	0.9697 ± 0.0059	0.9172 ± 0.0077	0.9980 ± 0.0016
Sub5	0.8737 ± 0.074	0.9790 ± 0.0021	0.9556 ± 0.0112	0.9949 ± 0.0017
Sub6	0.8565 ± 0.0061	0.9641 ± 0.0043	0.9289 ± 0.0075	0.9823 ± 0.0025
Sub7	0.9159 ± 0.0034	0.9759 ± 0.0043	0.9589 ± 0.0045	0.9971 ± 0.0026
Sub8	0.8226 ± 0.0026	0.9135 ± 0.0095	0.8963 ± 0.0073	0.9941 ± 0.0032

2.3. Multiple Kernel Learning. We all know that SVM is a supervised learning method of machine learning. It is often used to solve classification problems. Its basic principle is to find a hyperplane in the feature space and separate positive and negative samples with the highest accuracy. However, the previous SVMs are all single-kernel, which is based on a single feature space. In the actual classification task, we need our experience to select the appropriate kernel function (Gaussian kernel function, polynomial kernel function, etc.) and specify different parameters. This is not only

inconvenient, but also when our dataset features are heterogeneous, the effect is not so good.

The multiple kernel learning (MKL) [19] model is born because of this application. The multikernel model is more flexible than the single-kernel model and can give full play to the mapping capabilities of different kernel functions in a combined space composed of multiple feature spaces.

2.4. Multikernel SVM. According to the Mercer theorem and its properties [20], if K_1 and K_2 are kernels on $X \times X$ and X belongs to R, then the following are also kernel functions:

$$K(x,z) = K_1(x,z) + K_2(x,z), \tag{1}$$

$$K(x,z) = \alpha K_1(x,z). \tag{2}$$

If there is more than one kernel on $X \times X$, but multiple kernels coexist, then according to (1) and (2), we know that the following is also a kernel function:

$$\begin{aligned} &K(x,z) = \sum_{m=1}^{M} d_m K_m(x,z), \quad d_m \geq 0, \\ &\sum_{m=1}^{M} d_m = 1. \end{aligned} \tag{3}$$

Equation (3) is the most common form of multiple kernel function combination, which is a weighted combination of single-kernel functions, is the basis kernel function, is the number of basis kernel functions, and is the weight coefficient corresponding to the first basis kernel function.

In general, the original problem of MKL can be described as follows:

$$\begin{cases} \min\limits_{w_{m,b,\xi,d}} & \dfrac{1}{2}\sum\limits_{m=1}^{M}\dfrac{1}{d_m}\|w_m\|_{H_m}^2 + C\sum\limits_{i=1}^{n}\xi_i \\ \text{s.t.} & y_i\left(\sum\limits_{m=1}^{M} w_m \cdot \varphi(x_i) + b\right) \geq 1 - \xi_i, \quad \xi_i \geq 0 \\ & \sum\limits_{m=1}^{M} d_m = 1, \quad d_m \geq 0, \end{cases} \tag{4}$$

where d_m is the kernel weight, w_m is the normal of the separating hyperplane corresponding to the M-th kernel, b is the deviation term, ξ is the slack variable, and C is the misclassification penalty coefficient. The above formula can be transformed into an optimization problem with d_m as a variable:

$$J(d) = \begin{cases} \min\limits_{d} & J(d) \\ \text{s.t.} & \sum\limits_{m=1}^{M} d_m = 1, \quad d_m \geq 0, \end{cases}$$

$$J(d) = \begin{cases} \min\limits_{w_m,b,\xi} & \dfrac{1}{2}\sum\limits_{m=1}^{M}\dfrac{1}{d_m}\|w_m\|_{H_m}^2 + C\sum\limits_{i=1}^{n}\xi_i \\ \text{s.t.} & y_i\left(\sum\limits_{m=1}^{M} w_m \cdot \varphi(x_i) + b\right) \geq 1 - \xi_i, \quad (i = 1, 2, \ldots, n) \\ & \xi_i \geq 0. \end{cases} \tag{5}$$

The above objective function $J(d)$ is a standard SVM problem, which is a convex optimization problem that includes d_m. We can introduce a generalized Lagrangian function to convert the original problem $J(d)$ into a dual problem. The dual problem is a minimax problem; then,

$$L(w_m, b, \xi, \alpha, \nu) = \frac{1}{2}\sum_{m=1}^{M}\frac{1}{d_m}\|w_m\|_{H_m}^2 + C\sum_{i=1}^{n}\xi_i + \sum_{i=1}^{n}\alpha_i\left(1 - \xi_i - y_i\left(\sum_{m=1}^{M} w_m \cdot \varphi(x_i) + b\right)\right) - \sum_{i=1}^{n}\nu_i\xi_i, \tag{6}$$

where α and ν is the Lagrangian multiplier, and the Langrangian function $L(w_m, b, \xi, \alpha, \nu)$ takes the partial derivative of and makes it equal to zero, we can get

$$\begin{aligned}\nabla_{w_m} L(w_m, b, \xi, \alpha, \nu) &= \frac{1}{d_m} w_m - \sum_{i=1}^{n} \alpha_i y_i \varphi_m(x_i) = 0,\\ \nabla_b L(w_m, b, \xi, \alpha, \nu) &= -\sum_{i=1}^{n} \alpha_i y_i = 0,\\ \nabla_\xi L(w_m, b, \xi, \alpha, \nu) &= C - \alpha_i - \nu_i = 0.\end{aligned} \tag{7}$$

Through (8), we can get

$$\begin{aligned} w_m &= d_m \sum_{i=1}^{n} \alpha_i y_i \varphi_m(x_i),\\ \sum_{i=1}^{n} \alpha_i y_i &= 0,\\ C - \alpha_i + \nu &= 0.\end{aligned} \tag{8}$$

Substituting (9) into the Langrangian function (7), we can get

$$\begin{cases} \min\limits_{w_m, b, \xi} \quad L(w_m, b, \xi, \alpha, \nu) = -\dfrac{1}{2} \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} \alpha_i \alpha_j y_i y_j \sum\limits_{m=1}^{M} d_m K_m(x_i, x_j) + \sum\limits_{i=1}^{n} \alpha_i \\ \text{s.t.} \quad \sum\limits_{i=1}^{n} \alpha_i y_i = 0 \\ 0 \le \alpha_i \le C, \quad i = 1, 2, \ldots, n. \end{cases} \tag{9}$$

Find the maximum of $\min_{w_m, b, \xi}$ to α, and convert the maximum to the minimum to obtain the equivalent dual optimization problem:

$$\begin{cases} \min\limits_{\alpha} \quad L(w_m, b, \xi, \alpha, \nu) = \dfrac{1}{2} \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} \alpha_i \alpha_j y_i y_j \sum\limits_{m=1}^{M} d_m K_m(x_i, x_j) - \sum\limits_{i=1}^{n} \alpha_i \\ \text{s.t.} \quad \sum\limits_{i=1}^{n} \alpha_i y_i = 0 \\ 0 \le \alpha_i \le C, \quad i = 1, 2, \ldots, n. \end{cases} \tag{10}$$

Formula (10) is a dual form of the standard SVM problem including the combined kernel $K(x_i, x_j)$, assuming that $\alpha^* = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_l^*)^{\mathrm{T}}$ is the optimal solution of the dual optimization problem, then for the expression

$$w_m^* = d_m \sum_{i=1}^{n} \alpha_i^* y_i^* \varphi_m(x_i), \tag{11}$$

$$b_m^* = y_j - \sum_{m=1}^{M} d_m \sum_{i=1}^{N} \alpha_i^* y_i K(x_i \cdot x_j). \tag{12}$$

Through (11) and (12), the final decision function is

$$f(x) = \operatorname{sign}\left(\sum_{m=1}^{M} \left(d_m \sum_{i=1}^{n} \alpha_i^* y_i^* K(x \cdot x_i) + b_m^* \right) \right). \tag{13}$$

In summary, the description of the MKL learning algorithm can be found in Algorithm 1.

In this section, we introduced the multiple kernel learning. You can see that the kernel weights in MKL are fixed. In the next section, we will improve a new multiple kernel learning based on this point.

2.5. Adaptive Weighted Multikernel SVM. The kernel weight d_m in multikernel SVM is not adaptive, and it is more troublesome to use the gradient descent method to solve it. It takes multiple iterations to converge when it is close to the minimum. In this part we propose an adaptive weighted multikernel SVM model (AW-SVM).

In general, the original problem of AW-SVM is as follows:

$$\begin{cases} \min\limits_{w_m, b, \xi, d} \quad \dfrac{1}{2} \sum\limits_{m=1}^{M} d_m^r \left\| w_m \right\|_{H_m}^2 + C \sum\limits_{i=1}^{n} \xi_i \\ \text{s.t.} \quad y_i \left(\sum\limits_{m=1}^{M} w_m \cdot \varphi(x_i) + b \right) \ge 1 - \xi_i, \quad \xi_i \ge 0 \\ \sum\limits_{m=1}^{M} d_m = 1, \quad d_m \ge 0, \end{cases} \tag{14}$$

where d_m^r is the kernel weight, the index r is similar to the fuzzy index in FCM clustering, which is a relaxation of

Input: training set: $T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$ where $x_i \in \chi \in R^n$, $y_i \in \{-1, +1\}, i = 1, 2, \ldots, N$
Output: classification decision function $f(x)$
Step 1: choose the appropriate kernel function $K(x, z)$, the appropriate parameter C, and the loss parameter ξ
Kernel weight matrix d_m initialization: $d_m^1 = 1/M$, where M is the number of kernelss
Construct and solve optimization problems (10), and find the optimal solution $\alpha^* = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_l^*)^{\mathrm{T}}$
Step 2: choose $0 < \alpha^* < C$ positive component a from α^*, and calculate formula (11) and formula (12)
Step 3: construct decision function (13)

Algorithm 1: Learning algorithm for MKL.

the weight, and is used to realize the self-adaptation of the weight, which w_m is the normal of the separating hyperplane corresponding to the M-th kernel, b is the deviation term, ξ is the slack variable, and C is the misclassification penalty coefficient. Formula (14) can be transformed into an optimization problem with d_m as a variable

$$\begin{cases} \min\limits_{d} \quad J(d) \\ \text{s.t.} \quad \sum\limits_{m=1}^{M} d_m = 1, \quad d_m \geq 0, \end{cases}$$

$$J(d) = \begin{cases} \min\limits_{w_m, b, \xi} \quad \frac{1}{2} \sum\limits_{m=1}^{M} d_m^r \left\| w_m \right\|_{H_m}^2 + C \sum\limits_{i=1}^{n} \xi_i \\ \text{s.t.} \quad \left(\sum\limits_{m=1}^{M} w_m \cdot \varphi(x_i) + b \right) \geq 1 - \xi_i, \quad (i = 1, 2, \ldots, n) \\ \xi_i \geq 0. \end{cases} \tag{15}$$

The above objective function $J(d)$ is a standard SVM problem, which is a convex optimization problem involving d_m. Using the same method, we can get

$$L(w_m, b, \xi, \alpha, \nu) = \frac{1}{2} \sum_{m=1}^{M} d_m^r \left\| w_m \right\|_{H_m}^2 + C \sum_{i=1}^{n} \xi_i + \sum_{i=1}^{n} \alpha_i \left(1 - \xi_i - y_i \left(\sum_{m=1}^{M} w_m \cdot \varphi(x_i) + b \right) \right) - \sum_{i=1}^{n} \nu_i \xi_i. \tag{16}$$

Use the same method in the previous section (7), (8), and (10) to get the final dual optimization problem:

$$\begin{cases} \min\limits_{\alpha} \quad L(w_m, b, \xi, \alpha, \nu) = \frac{1}{2} \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} \alpha_i \alpha_j y_i y_j \sum\limits_{m=1}^{M} \frac{1}{d_m^r} K_m(x_i, x_j) - \sum\limits_{i=1}^{n} \alpha_i \\ \text{s.t.} \quad \sum\limits_{i=1}^{n} \alpha_i y_i = 0 \\ 0 \leq \alpha_i \leq C, \quad i = 1, 2, \ldots, n. \end{cases} \tag{17}$$

Assuming that $\alpha^* = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_l^*)^{\mathrm{T}}$ is the optimal solution of the dual optimization problem, then for the expression

$$w_m^* = \frac{1}{d_m^r} \sum_{i=1}^{n} \alpha_i^* y_i^* \varphi_m(x_i), \tag{18}$$

$$b_m^* = y_j - \sum_{m=1}^{M} \frac{1}{d_m^r} \sum_{i=1}^{N} \alpha_i^* y_i K\left(x_i \cdot x_j\right). \tag{19}$$

After getting the parameter w_m, d_m can be solved by iteration.

To construct a Lagrangian function with d_m as a variable for the original problem, we get

$$J(d, \alpha, \nu) = \frac{1}{2} \sum_{m=1}^{M} d_m^r \left\| w_m \right\|_{H_m}^2 + C \sum_{i=1}^{n} \xi_i + \sum_{i=1}^{n} \lambda_i \left(1 - \xi_i - y_i \left(\sum_{m=1}^{M} w_m \cdot \varphi(x_i) + b \right) \right) - \sum_{i=1}^{n} \nu_i \xi_i + \lambda_1 \left(\sum_{m=1}^{M} d_m - 1 \right). \tag{20}$$

Taking the partial derivative of the parameter d_m, λ_1 and making it equal to zero, we can get

$$\frac{\partial J(d)}{\partial (d_m)} = \frac{r}{2} d_m^{r-1} \left\| w_m \right\|^2 + \lambda_1 = 0, \tag{21}$$

$$\frac{\partial J(d)}{\partial (\lambda_1)} = \sum_{m=1}^{M} d_m = 1. \tag{22}$$

From the simultaneous formulas (21) and (22), using a similar solution method of U_{ij}^m in FCM [21], the final solution d_m is equal to

$$d_m = \frac{\left(1 / \left\| w_m \right\|^2 \right)^{(1/r-1)}}{\sum_{k=1}^{M} \left(1 / \left\| w_m \right\|^2 \right)^{(1/r-1)}}. \tag{23}$$

In summary, the two classification problem of AW-SVM can be described as follows:

Step 1: initialization of kernel weight matrix d_m, $\sum_{m=1}^{M} d_m = 1$, M is the number of kernels.

Step 2: in each iteration, use the combined kernel $K = \sum_{m=1}^{M} 1/d_m^r K_m$ to calculate the final dual optimization problem (17).

Step 3: calculate the optimal solution of (17). Calculate the separation hyperplane (w_m, b) corresponding to each kernel by formulas (18) and (19).

Step 4: update dm by formula (23).

Step 5: if the iteration termination condition is not met, return to Step 2 and repeat Steps 2–4. If the iteration condition is satisfied, the calculation is ended. Then, output the final decision function:

$$f(x) = \operatorname{sign}\left(\sum_{m=1}^{M} \frac{1}{d_m^r} \left(\sum_{i=1}^{n} \alpha_i^* y_i^* K\left(x \cdot x_i\right) + b_m^* \right) \right). \tag{24}$$

Iterative stop condition is duality gap (DG) [22]. Karush–Kuhn–Tucker (KKT) condition, $\Delta d = d^{t+1} - d^t \le \varepsilon_1$, reaches the maximum number of iterations.

The DG expression is

$$\max_m \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i^* \alpha_j^* y_i y_j K_m\left(x_i \cdot x_j\right) - \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i^* \alpha_j^* y_i y_j \sum_{m=1}^{M} \frac{1}{d_m^r} K_m\left(x_i, x_j\right) \le \varepsilon. \tag{25}$$

Both ε and ε_1 are thresholds.

3. Experiment

After getting the artificial contour image of all the data drawn by the expert, we start to train the AW-SVM learning algorithm. The algorithm runs on a computer with Intel(R) Core(TM) i5-8500CPU 3.00 GHZ 12 GB memory and 64 bit Windows10 operating system. The algorithm iterates 10 times on average, and the training time is about 30 minutes. Given a trained classifier model, the average segmentation time of images on treatment days is about 3 minutes.

3.1. Kernel Function. In this experiment, we choose Gaussian kernel function (26), Laplacian kernel function (27), and logarithmic kernel function (28). The formulas of the three are as follows:

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right), \quad (26)$$

$$K(x, y) = \exp\left(-\frac{\|x - y\|}{\sigma}\right), \quad (27)$$

$$K(x, y) = -\log\left(1 + \|x - y\|^d\right). \quad (28)$$

The Gaussian kernel is a classic robust radial basis kernel. It has good anti-interference ability against the noise in the data, which has been widely used, but the performance of the Gaussian kernel function is very sensitive to the parameters. The Laplacian kernel function is a variant of the Gaussian kernel function. The main change is to adjust the 2-norm to 1-norm, which reduces the sensitivity to parameters. Logarithmic kernel is often used in image segmentation.

3.2. Evaluation Index. In addition to using classification accuracy to evaluate the accuracy of the algorithm, this article also uses Dice Loss [23], which is very common in segmentation algorithms, to evaluate our segmentation accuracy, because segmentation usually classifies each pixel:

$$\text{Dice Loss} = 1 - \text{Dice Coefficient}. \quad (29)$$

Dice Coefficient, named after Lee Raymond Dice, is a set similarity measure function, which is usually used to calculate the similarity between two samples (value $\in [0, 1]$). A value of 1 indicates complete coincidence. The formula is as follows:

$$\text{Dice Coefficient} = \frac{2|X \cap Y|}{|X| + |Y|}, \quad (30)$$

where $|X|$ and $|Y|$, respectively, represent the number of elements in the set, as shown in Figure 2. In the segmentation task, the two, respectively, represent the real result and the predicted result.

Combining (29) and (29), Dice Loss expression is as follows:

$$\text{Dice Loss} = 1 - \frac{2|X \cap Y|}{|X| + |Y|}. \quad (31)$$

3.3. Comparative Test. In order to compare the performance of the AW-SVM algorithm on this dataset, this paper arranges other seven algorithms as an experimental comparison. The comparison algorithm is shown in Table 2.

The experimental training set consists of daily MRI data before treatment, and the treatment day dataset is used as the test set. The organs and skin except the kidney in the dataset are labeled as background label 1, and the kidney as the circle target, and the label is −1.

3.4. Experiment Process. For each patient (Sub1~8), complete the preliminary data processing (local texture feature extraction, three-dimensional coordinate extraction, Kalman filter dimension enhancement, etc.). The MRI image during the treatment will be used as training data, denoted as I_t; the MRI image of the treatment day is used as the test data, denoted as I_s. The experiment is carried out as shown in Figure 3:

(1) Randomly select the $L = 4000$ group of examples in the labeled I_t (label: {−1, 1}); that is, select 4000 voxels from the kidney and background in the MRI image during the treatment process to form the training set G.

(2) Using the AW-SVM method on the training set G, through parameter tuning, we can get the best classifier C_r for this patient data.

(3) In the treatment day MRI dataset I_t, select a complete treatment day MRI image as the test input (including 36,000+ voxels), the label corresponding to each voxel is obtained through the best classifier C_r trained in Step 2, and the result is recorded as U.

(4) Select all MRI images of the treatment day on I_t as the test input, and obtain the labels corresponding to all voxels through the KNN algorithm (label in U is used as the training label).

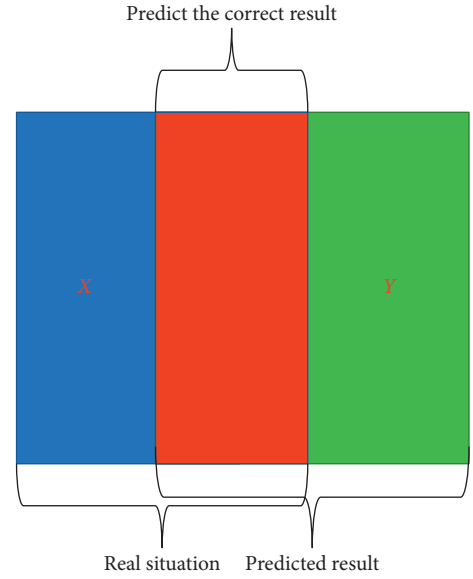


Figure 2: The description of the Dice Coefficient.

Table 2: Comparison algorithm used.

Our algorithm	Comparison algorithm
Adaptive weighted multikernel SVM (AW-SVM)	Support vector machine (SVM) Multiple kernel learning (MKL) Transductive support vector machine (TSVM) [24] Semisupervised SVM (S4VM) [25] Random subspace method (RSM) [26] Naive bayes (NB) Radial basis function (RBF) [27]

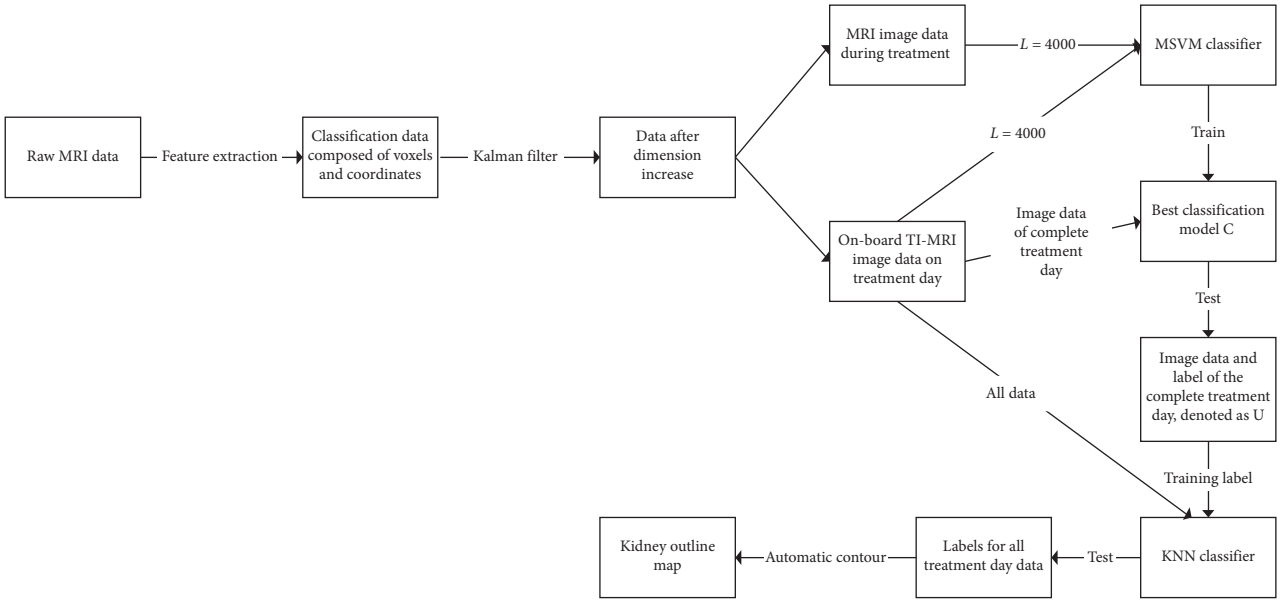


Figure 3: Experiment process.

(5) Obtain the outline of the kidney through the obtained label.

3.5. Results and Discussion. In order to verify that the data processed by Kalman filter performs better in classification performance, we set up the first set of comparative experiments, we have prepared two sets of patient data; the one that is not processed by Kalman filtering is recorded as the first group, and the one that is processed by Kalman filtering is recorded as the second group. The eight algorithms have been experimented on these two sets of data, and the experimental results are shown in Table 1, from which we can get the following information:

(1) It can be clearly seen that compared with the first set of results, each classifier performs better on the second set of data, and the classification accuracy has been improved, indicating that the use of Kalman filtering can improve the classification accuracy and enhance interpretable.

(2) In each set of results, basically the classification results of AW-SVM are the best. Classification accuracy is much higher than TSVM, S4VM, NB, RBF, and RSM, and the accuracy of MKL and SVM is close to AW-SVM, but for this sample, AW-SVM's performance is still the best. In general, the experimental results show that our algorithm is effective and excellent for classification on the MRI data of the lower abdomen.

In order to compare whether there are significant differences between AW-SVM and the other seven algorithms in statistics, Friedman test [28, 29] is used to analyze the classification accuracy of the eight algorithms. In the Friedman test, each individual has the same sample size, and each individual has a corresponding relationship with the corresponding individual in other samples, so the Friedman test can make full use of the information in the original data. If $p \leq 0.025$, the null hypothesis was rejected. At the same time, in order to further verify that the method to obtain the best Friedman ranking and other methods are statistically significant differences, we also conducted Holm post hoc test [28, 29].

As shown in Tables 3 and 4, the Friedman test shows that the classification performance of the AW-SVM method in this dataset is greater than that of the other seven algorithms. Holm post hoc test shows that the performance of AW-SVM algorithm on this dataset is significantly better than RSM, NB, S4VM, TSVM, and RBF, and for S4VM and MKL, it is not obvious. In summary, the AW-SVM algorithm can be used for lower abdomen MRI image segmentation, and its classification performance is better than the comparison of these seven classification algorithms. The experimental results show that its accuracy rate is close to 1 on each dataset, and the subsequent DSI values also show its excellent classification performance.

In the field of medical segmentation, Dice similarity index (DSI) value greater than 0.7 is an acceptable value, which means that the predicted contour and the real contour have a good overlap [30]. The results of DSI value are shown in Table 5. The performance of AW-SVM, SVM, and MKL models is as good as accuracy; DSI value reached 1 in multiple patients, which indicates that the automatic contour and manual contour are completely coincident. The lowest DSI values of the three are generated on Sub1, respectively, 0.9654 ± 0.0121, 0.9423 ± 0.0127, and 0.9324 ± 0.0162, which are also acceptable values, but it also shows that the MRI image of Sub1 is the most complicated (there may be overlaps, missing, shadows, etc.) [31]. The performance of several other algorithms is uneven, but the DSI value of the RSM algorithm is around 0.2, which is a completely

Table 3: Result of Holm's post hoc test ($\alpha = 0.05$).

Holm's post hoc test for AW-SVM				
Algorithms	$z = (R_o - R_i)/\mathrm{SE}$	P	holm = α/i	Hypothesis
RSM	5.358259	0	0.007143	Rejected
NB	4.184545	0.000029	0.008333	Rejected
TSVM	3.878359	0.000105	0.01	Rejected
S4VM	2.602583	0.009252	0.0125	Rejected
RBF	2.602583	0.009252	0.016667	Rejected
MKL	0.918559	0.358326	0.025	Not rejected
SVM	0.459279	0.646034	0.05	Not rejected

Table 4: Result of Friedman test ($\alpha = 0.05$).

Friedman test for AW-SVM			
Algorithm	Friedman rank	P value	Hypothesis
RSM	8	0	Rejected
NB	6.5625		
TSVM	6.1875		
S4VM	4.625		
RBF	4.625		
MKL	2.5625		
SVM	2		
AW-SVM	1.4375		

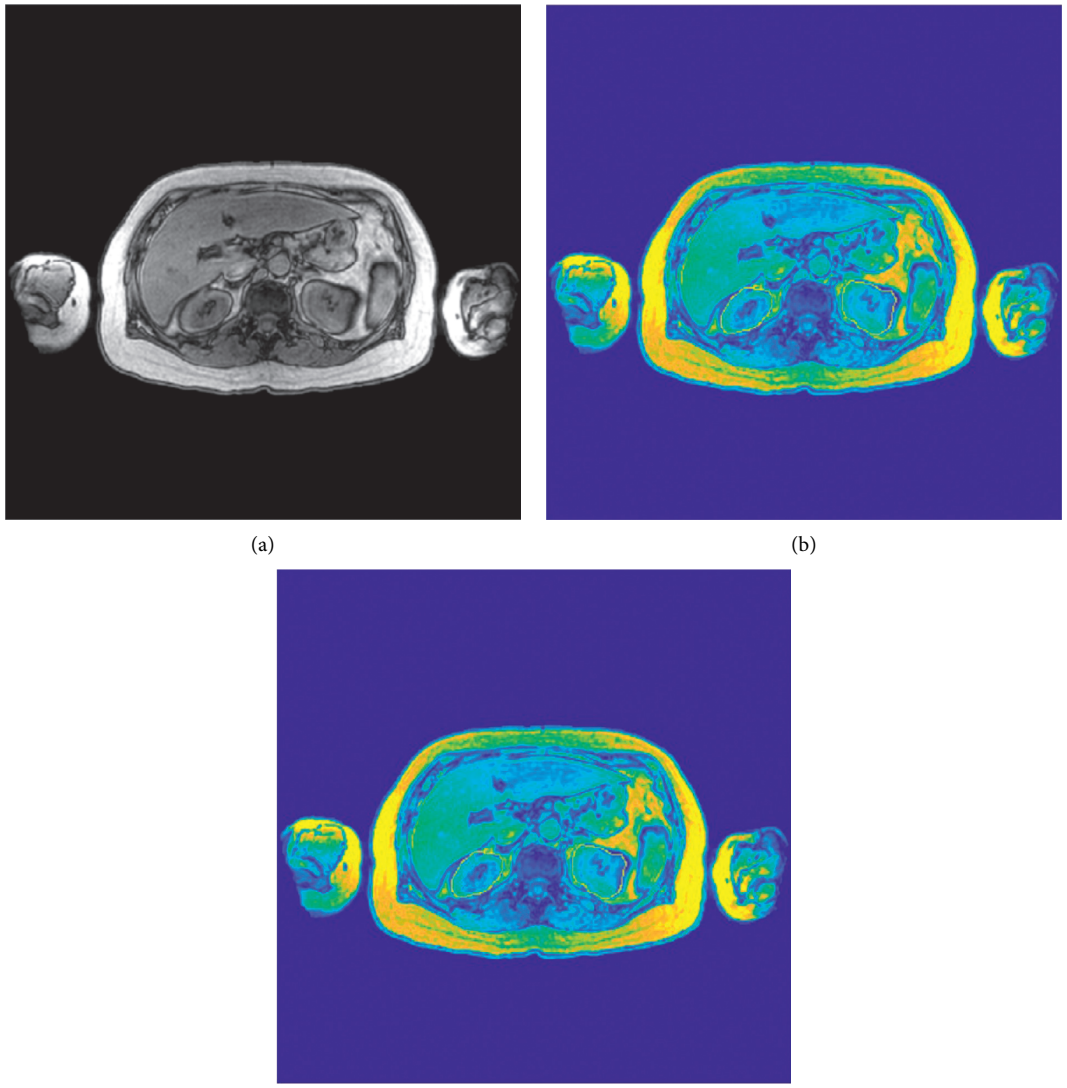
Table 5: Average dice similarity index (DSI) for the kidneys using the classification algorithm.

Part A: enter as part of a complete treatment day image				
Datasets	AW-SVM	SVM	S4VM	MKL
Sub1	0.9654 ± 0.0121	0.9423 ± 0.0127	0.7176 ± 0.0211	0.9324 ± 0.0162
Sub2	1	1	0.8355 ± 0.0156	1
Sub3	0.9882 ± 0.0094	0.9749 ± 0.0114	0.8431 ± 0.0257	0.9782 ± 0.0076
Sub4	1	1	0.9100 ± 0.0336	0.9995 ± 0.0002
Sub5	1	1	0.8800 ± 0.0546	1
Sub6	1	1	0.8436 ± 0.0235	0.9988 ± 0.0011
Sub7	1	1	0.8852 ± 0.0303	1
Sub8	1	1	0.7181 ± 0.0489	1
Datasets	RSM	TSVM	NB	RBF
Sub1	0.2157 ± 0.0021	0.4267 ± 0.0153	0.7510 ± 0.0189	0.5976 ± 0.0213
Sub2	0.3078 ± 0.0012	0.6127 ± 0.0483	0.8298 ± 0.0220	0.6624 ± 0.0176
Sub3	0.1724 ± 0.0031	0.6925 ± 0.0221	0.8528 ± 0.0176	0.6768 ± 0.0229
Sub4	0.2226 ± 0.0012	0.8986 ± 0.0318	0.9220 ± 0.0304	0.7633 ± 0.0318
Sub5	0.1862 ± 0.0016	0.8318 ± 0.0430	0.9111 ± 0.019	0.7796 ± 0.0174
Sub6	0.2543 ± 0.0021	0.6811 ± 0.0194	0.8363 ± 0.0343	0.6717 ± 0.0312
Sub7	0.2784 ± 0.0011	0.5653 ± 0.0236	0.8968 ± 0.0308	0.6602 ± 0.0158
Sub8	0.1748 ± 0.0016	0.7168 ± 0.0297	0.7113 ± 0.0352	0.8681 ± 0.0153

unacceptable value. Indicating that RSM is not applicable on this dataset.

Figure 4 shows the original cross-sectional image of the lower abdomen generated by the OP sequence (the clearest OP sequence) in MRI, the manual contour drawn by a professional doctor, and the automatic contour generated based on the AW-SVM algorithm. The manual contour is used as the evaluation criterion.

The running time of the automatic contour method is also a key point for clinical application. The faster the treatment speed, the shorter the time the patient will be fixed on the treatment bed and the less pain they will suffer. We select the three models with the best performance of DSI and calculate their prediction time; the prediction time refers to the time it takes to input the complete treatment day data and get the automatic contour after obtaining the corresponding best classification model. The prediction time is shown in Table 6, affected by programming; our AW-SVM algorithm is not superior in predicting time, but it is also a completely acceptable value, with an average time of 2.9263 seconds. The automatic contour algorithm based on AW-SVM greatly shortens the drawing of the contour of the target area on the treatment day and greatly reduces the pain of the patient.



(a) (b) (c)

Figure 4: Abdominal MRI cross-sectional image showing the automatic contour for kidneys obtained by the AW-SVM method. (a) Original image, (b) manual contour, and (c) automatic contour. Red: manual contours used for ground truth; green: automatic contours of the kidneys.

Table 6: Average prediction time of the AW-SVM and MKL and SVM methods.

Datasets	AW-SVM	MKL	SVM
Sub1	2.53	1.18	0.23
Sub2	3.01	1.81	0.24
Sub3	2.87	1.56	0.22
Sub4	2.81	1.79	0.26
Sub5	2.73	1.27	0.25
Sub6	3.12	1.22	0.27
Sub7	3.02	1.31	0.25
Sub8	3.32	1.20	0.29

4. Conclusion

A mature automatic contour algorithm is the key to the clinical application of MR-ART. The experimental results show that the adaptive weighted AW-SVM algorithm proposed in this paper shows excellent performance in abdominal MRI image segmentation, which is comparable to many excellent classification algorithms. The features

used in this article are only simple voxel and coordinate features. In the future, deep neural networks can be used to process data to extract deep features to further improve the classification performance of the AW-SVM algorithm. In the future, it will not only be drawn for a single kidney outline but also is hoped that all organs in the abdomen can be drawn.

Data Availability

The labeled datasets used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

Yi Gu and Bo Li contributed equally to this work.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61772241 and in part by the 2018 Six Talent Peaks Project of Jiangsu Province under Grant XYDXX-127.

References

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA: A Cancer Journal for Clinicians*, vol. 70, no. 1, pp. 7–30, 2020.

[2] A. Brandmaier and S. C. Formenti, "The impact of radiation therapy on innate and adaptive tumor immunity," *Seminars in Radiation Oncology*, vol. 30, no. 2, pp. 139–144, 2020.

[3] R. K. Funk, A. L. Stockham, and N. N. I. Laack, "Basics of radiation therapy," *Clinical Cardio-Oncology*, Elsevier, Amsterdam, Netherland, 2016.

[4] J. Chen, O. Morin, M. Aubin, M. K. Bucci, C. F. Chuang, and J. Pouliot, "Dose-guided radiation therapy with megavoltage cone-beam CT," *The British Journal of Radiology*, vol. 79, no. 1, pp. 87–98, 2006.

[5] S. Kuznetsova, P. Grendarova, S. Roy, R. Sinha, K. Thind, and N. Ploquin, "Structure guided deformable image registration for treatment planning CT and post stereotactic body radiation therapy (SBRT) Primovist (Gd-EOB-DTPA) enhanced MRI," *Journal of Applied Clinical Medical Physics*, vol. 20, no. 12, pp. 109–118, 2019.

[6] B. W. Raaymakers, J. J. W. Lagendijk, J. Overweg et al., "Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept," *Physics in Medicine and Biology*, vol. 54, no. 12, pp. N229–N237, 2009.

[7] S. Devic, "MRI simulation for radiotherapy treatment planning," *Medical Physics*, vol. 39, no. 11, pp. 6701–6711, 2012.

[8] S. Acharya, B. W. Fischer-Valuck, R. Kashani et al., "Online magnetic resonance image guided adaptive radiation therapy: first clinical applications," *International Journal of Radiation Oncology∗Biology∗Physics*, vol. 94, no. 2, pp. 394–403, 2016.

[9] J. Yang, H. Veeraraghavan, W. Elmpt, A. Dekker, M. Gooding, and G. Sharp, "CT images with expert manual contours of thoracic cancer for benchmarking auto-segmentation accuracy," *Medical Physics*, vol. 47, no. 7, pp. 3250–3255, 2020.

[10] L. G. Roberts, *Machine Perception of Three-Dimensional Solids*, Massachusetts Institute of Technology, Cambridge, England, 1965.

[11] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.

[12] I. Sobel, "Camera models and machine perception," dissertation, stanford university, 1970.

[13] C. J. C. Burges, "A tutorial on Support Vector Machines for pattern recognition," *Data Min. Knowl. Discov.,* vol. 2, pp. 121–167, 1998.

[14] K. Adi, C. E. Widodo, A. P. Widodo, R. Gernowo, A. Pamungkas, and R. A. Syifa, "Naïve Bayes algorithm for lung cancer diagnosis using image processing techniques," *Advanced Science Letters*, vol. 23, no. 3, pp. 2296–2298, 2017.

[15] S. L. Salzberg, *Book Review: C4.5: Programs for Machine Learning*, J. Ross Quinlan, Ed., Springer, New York, NY, USA, 1993.

[16] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2015.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 9351, pp. 234–241, Springer International Publishing, Munich, Germany, October 2015.

[18] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[19] M. Gönen and E. Alpaydın, "Multiple kernel learning algorithms," *Journal of Machine Learning Research*, vol. 12, pp. 2211–2268, 2011.

[20] J. Mercer, "XVI. Functions of positive and negative type, and their connection the theory of integral equations," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 209, no. 441-458, pp. 415–446, 1909.

[21] N. R. Pal and J. C. Bezdek, "On cluster validity for the fuzzy c-means model," *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 3, pp. 370–379, 2002.

[22] A. Rakotomamonjy, F. R. Bach, S. Canu, and Y. Grandvalet, "Simple MKL," *Journal of Machine Learning Research*, vol. 9, no. 3, pp. 2491–2521, 2008.

[23] F. Milletari, N. Navab, and S. Ahmadi, "V-Net: fully convolutional neural networks for volumetric medical image segmentation," in *Proceedings of the Fourth International Conference on 3D Vision*, pp. 565–571, 3DV), Stanford, CA, USA, October 2016.

[24] O. Chapelle, V. Sindhwani, and S. S. Keerthi, "Optimization techniques for semi-supervised support vector machines," *Journal of Machine Learning Research*, vol. 9, no. 6, pp. 203–223, 2008.

[25] Z. Xu, Y. Tian, Z. Li, and Y. Li, "SEMG multi-class classification based on S4VM algorithm," *Lecture Notes in Electrical Engineering*, vol. 21, no. 4, pp. 495–503, 2013.

[26] N. García-Pedrajas and D. Ortiz-Boyer, "Boosting random subspace method," *Neural Networks*, vol. 21, no. 9, pp. 1344–1362, 2008.

[27] J. Park and I. Sandberg, "Universal approximation using radial-basis-function networks," *Neural Computation*, vol. 3, no. 2, pp. 246–257, 2014.

[28] J. Demiar and D. Schuurmans, "Statistical comparisons of classifiers over multiple data sets," *Journal of Machine Learning Research*, vol. 7, no. 1, pp. 1–30, 2006.

[29] S. Gar´ci and F. Herrera, "An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons," *Journal of Machine Learning Research*, vol. 9, pp. 2677–2694, 2008.

[30] K. H. Zou, S. K. Warfield, A. Bharatha et al., "Statistical validation of image segmentation quality based on a spatial overlap index1," *Academic Radiology*, vol. 11, no. 2, pp. 178–189, 2004.

[31] L. Fan, P. Qian, K.-H. Su et al., "Abdominal, multi-organ, auto-contouring method for online adaptive magnetic resonance guided radiotherapy: an intelligent, multi-level fusion approach," *Artificial Intelligence in Medicine*, vol. 90, pp. 34–41, 2018.