Proceedings of the IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI 2012)
Hong Kong and Shenzhen, China, 2-7 Jan 2012

Malignant Nodule Detection on Lung CT Scan Images with Kernel RX-algorithm

Aminmohammad Roozgard, Samuel Cheng, and Hong Liu

Abstract— **In this paper, we present a nonlinear anomaly detector called kernel RX-algorithm and apply it to CT images for malignant nodule detection. Malignant nodule detection is very similar to anomaly detection in military imaging applications where the RX-algorithm has been successfully applied. We modified the original RX-algorithm so that it can be applied to anomaly detection in CT images. Moreover, using kernel trick, we mapped the data to a high dimensional space to obtain a kernelized RX-algorithm that outperforms the original RX-algorithm. The preliminary results of applying the kernel RX-algorithm on annotated public access databases suggests that the proposed method may provide a means for early detection of the malignant nodules.**

I. Introduction

Lung cancer is the second highest cause of cancer death in both men and women in the United States and has one of the lowest 5-year survival rates of all cancers [1]. Only 15% of total lung cancer populations can potentially be cured and the average survival length is between 11 and 13 months after diagnosis.

In the human lung, there are generally two kinds of nodules: *malignant* and *benign*. These nodules are actually extremely common even though most of them are benign [2]. The early detection of a malignant nodule and cancerous tissue will increase the chance of survival. The conventional approaches of detection mainly rely on the judgment of physicians based on presence of fat [2], shape [3], size [4] and growth rate [5] of the nodule.

Lung cancer diagnosis using CT focuses on detection of pulmonary nodules, which has been an active research area in the last decade [6]–[10]. Current nodule detection systems usually consist of three steps such as, preprocessing, candidate detection and post processing. The preprocessing step aims to restrict the search space and reduce noise and the post processing step aims to reduce false positive alarm.

Numerous techniques have been proposed for candidate detection. These include multiscale Laplacian of Gaussian filtering [11], K-nearest neighbors [12], and various forms of clustering [13]. Many of these methods are able to detect most nodules. However, many of these nodule candidates are obviously false positive. To reduce classification complexity, numerous techniques have been proposed to reduce false positive. For example, these include rule-based approach [14] and soft margin SVM classifier [11].

The authors are with the Department of Electrical and Computer Engineering, University of Oklahoma, Tulsa and Norman, OK, USA.
This work was supported in part by Oklahoma Tobacco Research Center under a seed grant.

Malignant nodule detection is very similar to anomaly detection in military imaging applications, where the RX-algorithm has been successfully applied. In the past, RX-algorithms were usually applied in low dimensional space even though projecting features to a high dimensional space can potentially improve the performance. This is because a nonlinear classification boundary in a original low dimensional space can potentially be mapped to a linear boundary in a high dimensional space. The direct application of the aforementioned projection method is impractical due to intractable computational complexity. However, the problem can nonetheless be made manageable by using the so-called kernel trick [15]. Recently, the popular RX-algorithm for anomaly detection has been kernerlized and has shown promising results in military imaging applications [16], [17].

Encouraged by the success of the RX-algorithm and its variations for anomaly detection in military imaging applications, we developed and evaluated the same techniques for malignant nodule detection in this study. Moreover, we used images from public databases [18], [19] to evaluate our algorithms.

The goal of this research work is to increase lung cancer diagnosis accuracy by developing *novel* image/signal processing techniques for malignant nodule detection. To the best of our knowledge, there is no prior attempt of using the proposed RX-algorithm or kernel RX-algorithm for lung cancer detection. Preliminary result on annotated public access databases suggests that the proposed method may provide a means for early detection of the malignant nodules.

This paper is organized as follows. Section II provides a brief review of the RX-algorithm. Section III describes the feature space and the kernel-based anomaly detector. This section contains essential equations in understanding the kernel RX-algorithm. In Section IV, we propose a kernel RX-based anomaly detector that could be used for early stage cancer detection. Experimental results of the proposed method on CT image are given in Section V. Finally, we conclude the paper in Section VI.

II. RX-Algorithm

The RX-algorithm, a constant false alarm rate (CFAR) anomaly detection technique, has obtained its name from its inventors Reed and Xiaoli Yu [20]. It considers a two hypothesis problem with one hypothesis as foreground and the other as background. More precisely, the two hypotheses

978-1-4577-2177-9/12/$25 (C) 2012 IEEE

are given by

$$\begin{array}{ll} (Target\ Absent) & \mathcal{H}_0: \ \mathbf{r} = \mathbf{n}, \\ (Target\ Present) & \mathcal{H}_1: \ \mathbf{r} = \mathbf{s} + \mathbf{n}, \end{array} \tag{1}$$

where $\mathbf{s} = [s_1, s_2, \cdots, s_{p \cdot c}]^T$ is a vector generated by a target image patch composed of p pixels each with c components, $\mathbf{n}$ is a background noise vector characterized by a covariance matrix C_b, and $\mathbf{r}$ is a vector generated by a candidate image patch where a malignant nodule may be present or absent.

Since malignant nodules are rare, we can accurately estimate the mean and covariance matrix of the background noise even if we do not exclude the contribution of pixels containing a potential nodule. Given an input image and a region of interest (ROI), denote $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M$ as column vector representations of image patches each centered at a pixel insider the ROI. For the convenience of exposition, we also combine the patches and define a data matrix $\mathbf{X}_b = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M]$. We can then estimate from $\mathbf{X}_b$ the mean and covariance of the background noise, denoted as $\hat{\mu}_b$ and $\hat{C}_b$, respectively.

Given an input patch $\mathbf{r}$, the RX-algorithm estimates the degree of anomaly (from the background) of the patch by computing

$$\begin{aligned} RX(\mathbf{r}) \ = \ & (\mathbf{r} - \hat{\mu}_b)^T \\ & \cdot \Big(\frac{M}{M+1} \hat{C}_b + \frac{1}{M+1} (\mathbf{r} - \hat{\mu}_b)(\mathbf{r} - \hat{\mu}_b)^T \Big)^{-1} \\ & \cdot (\mathbf{r} - \hat{\mu}_b). \end{aligned} \tag{2}$$

With a predetermined threshold η, the RX-algorithm decides that a target is present (hypothesis $\mathcal{H}_1$) if $RX(\mathbf{r}) > \eta$ and a target is absent (hypothesis $\mathcal{H}_0$) otherwise. As $M \to \infty$, Equation (2) converges to

$$RX(\mathbf{r}) = (\mathbf{r} - \hat{\mu}_b)^T \hat{C}_b^{-1} (\mathbf{r} - \hat{\mu}_b). \tag{3}$$

III. Kernel RX-Algorithm

A. Kernel trick

As the dimension of data increases, it becomes more and more likely that two groups of data can be separable by a linear hyperplane [15]. Therefore, mapping data into a very high (close to infinite) dimensional space will typically make a classification problem easier since we can limit ourselves to linear decision boundaries. More precisely, consider two-class case with a linear discriminant model (characterized by a vector $\mathbf{a}$ and a threshold b) and a mapping function $\Psi(\cdot)$. An input data $\mathbf{r}$ will be treated as in one class when $\langle \Psi(\mathbf{r}), \Psi(\mathbf{a}) \rangle > b$ and in the alternate class when $\langle \Psi(\mathbf{r}), \Psi(\mathbf{a}) \rangle < b$. While the linear model is very simple, the complexity to compute the inner product $\langle \Psi(\mathbf{r}), \Psi(\mathbf{a}) \rangle$ becomes prohibitively high when the dimension of the range of $\Psi(\cdot)$ is very high.

Instead of computing the inner product directly, the kernel trick suggests to compute it through a kernel function $k(\cdot, \cdot)$ such that [15]

$$\langle \Psi(\mathbf{r}), \Psi(\mathbf{a}) \rangle \ = \ k(\mathbf{r}, \mathbf{a}). \tag{4}$$

Note that we can reduce the computational complexity significantly since $\mathbf{r}$ and $\mathbf{a}$ lie in a space of much lower dimension than that of the space where $\Psi(\mathbf{r})$ and $\Psi(\mathbf{a})$ lie.

B. kernel-based detector

Using the same assumption of RX-algorithm which was introduced in Section II, we can map the input date space to a high-dimensional feature space and express two hypothesis as

$$\begin{array}{ll} (Target\ Absent) & \mathcal{H}_0^\Psi: \ \mathbf{r} = \mathbf{n}^\Psi, \\ (Target\ Present) & \mathcal{H}_1^\Psi: \ \mathbf{r} = \Psi(\mathbf{s}) + \mathbf{n}^\Psi, \end{array} \tag{5}$$

where $\mathbf{n}^\Psi$ and $\Psi(\mathbf{s})$ are the noise vector and mapped input data in the feature space, respectively. In the feature space, the RX-function in (3)) can be rewritten as

$$RX(\mathbf{r}) = (\Psi(\mathbf{r}) - \hat{\mu}_b^\Psi)^T \big(\hat{C}_b^\Psi\big)^{-1} (\Psi(\mathbf{r}) - \hat{\mu}_b^\Psi), \tag{6}$$

where

$$\hat{\mu}_b^\Psi = \frac{1}{M} \sum_{i=1}^{M} \Psi\big(\mathbf{x}_i\big) \tag{7}$$

and

$$\hat{C}_b^\Psi = \frac{1}{M} \sum_{i=1}^{M} \Big(\Psi\big(\mathbf{x}_i\big) - \hat{\mu}_b^\Psi \Big) \Big(\Psi\big(\mathbf{x}_i\big) - \hat{\mu}_b^\Psi \Big)^T \tag{8}$$

are the estimated mean and covariance in the feature space, respectively.

Denote $\mathbf{k}(\mathbf{X}_b, \mathbf{r})$ and $\mathbf{K}(\mathbf{x}_i, \mathbf{X}_b)$ as column vectors with kernels $k(\mathbf{x}_i, \mathbf{r})$, $i = 1, \cdots, M$, and kernels $k(\mathbf{x}_i, \mathbf{x}_j)$, $j = 1, \cdots, M$ as their components. Further denote $\mathbf{K}_b$ as the uncentered Gram matrix such that $\mathbf{K}_b = [\mathbf{K}(\mathbf{x}_1, \mathbf{X}_b), \mathbf{K}(\mathbf{x}_2, \mathbf{X}_b), \cdots, \mathbf{K}(\mathbf{x}_M, \mathbf{X}_b)]$. By applying the kernel trick on equation (6), we can rewrite the kernelized RX-function, $RX_K(\mathbf{r})$, as

$$RX_K(\mathbf{r}) = (\mathbf{K}_\mathbf{r}^T - \mathbf{K}_{\hat{\mu}_b}^T)^T \hat{\boldsymbol{K}}_b^{-1} (\mathbf{K}_\mathbf{r}^T - \mathbf{K}_{\hat{\mu}_b}^T), \tag{9}$$

where

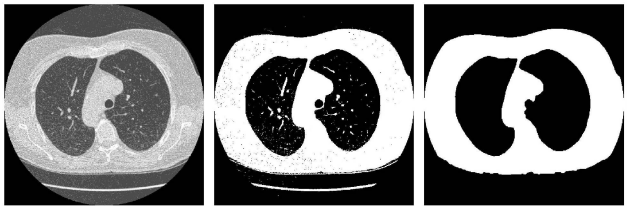
$$\begin{aligned} \mathbf{K}_\mathbf{r}^T \ &= \ \mathbf{k}(\mathbf{X}_b, \mathbf{r})^T - \frac{1}{M} \sum_{i=1}^{M} k(\mathbf{x}_i, \mathbf{r}) \\ &= \ \Psi(\mathbf{r})^T \mathbf{X}_b^\Psi, \end{aligned} \tag{10}$$

$$\begin{aligned} \mathbf{K}_{\hat{\mu}_b}^T \ &= \ \frac{1}{M} \sum_{i=1}^{M} \mathbf{K}(\mathbf{x}_i, \mathbf{X}_b) - \frac{1}{M^2} \sum_{i=1}^{M} \sum_{j=1}^{M} k(\mathbf{x}_i, \mathbf{x}_j) \\ &= \ (\hat{\mu}_b^\Psi)^T \mathbf{X}_b^\Psi, \end{aligned} \tag{11}$$

and $\hat{\boldsymbol{K}}_b$ is the centered Gram matrix which can be estimated from the uncentered Gram matrix as [21]

$$\hat{\boldsymbol{K}}_b = \mathbf{K}_b - \mathbf{1}_M \mathbf{K}_b - \mathbf{K}_b \mathbf{1}_M - \mathbf{1}_M \mathbf{K}_b \mathbf{1}_M, \tag{12}$$

where $\mathbf{1}_M$ is an $M \times M$ matrix with all of its coefficient equal to $\frac{1}{M}$.



(a) (b) (c)

Fig. 1: Pre-processing of a CT image. The images in (a), (b) and (c) are the original image, the segmented tissue after watershed segmentation, and the refined segmentation result after applying morphological and median filters, respectively.

IV. Kernel RX-Algorithm for lung anomaly detection

The kernel RX-algorithm described in Section III has been applied to the detection of military targets over hyperspectral images [16]. Since each pixel contains many components in that case, it is sufficient to use a single pixel as a patch for anomaly detection. However, since we only have few components in medical images, a single-pixel patch will not work in general. Moreover, the shape and size of patches will also affect the detector performance. In this paper, we have considered four patch models as follows:

1) Linear patch $N_L(x, y, z)$: The set of $(2 \times t + 1)$ pixels contains pixels from $(x, y, z - t)$ to $(x, y, z + t)$.
2) Pyramid patch $N_P(x, y, z)$: The set of seven pixels contains $(x, y, z \pm 1)$, $(x, y \pm 1, z)$, $(x \pm 1, y, z)$ and (x, y, z).
3) Cubic patch $N_C(x, y, z)$: The set of nine pixels contains $(x \pm 1, y \pm 1, z \pm 1)$ and (x, y, z).
4) Mixture patch $N_M(x, y, z)$: The set of fifteen pixels contains $(x, y, z \pm 1)$, $(x, y \pm 1, z)$, $(x \pm 1, y, z)$, $(x \pm 1, y \pm 1, z \pm 1)$ and (x, y, z).

As described in Section II, patches are vectorized and combined into a data matrix for the estimation of means and covariance matrices of the background noise (in the original and the feature spaces). To maintain good local estimates, we gather the background noise statistics of a pixel only from patches with centers near the pixel. However, the computational complexity is rather high if we estimate noise statistics independently for each pixel. Fortunately, the background noise statistics are very similar for adjacent pixels and thus we may approximate them to be the same.

In practice, we partition a region of interest into $q \times q$ windows. Within each window, we approximate the statistics to be the same, which are gathered from a larger concentric window of size $k \times k$. By using this speed up, we can approximately reduce the complexity of algorithm by $O(q^2)$.

V. Experimental results

To both reduce noise and increase the speed of algorithm, we apply a simple preprocessing step that segments the lung tissue out from a CT image. Then we classify the healthy and non-healthy areas by Kernel RX-Algorithm only on segmented lung tissue of CT images.

We apply a simple watershed segmentation in the $CIELAB$ color space of an input CT image. Fig. 1(b) shows the segmentation result. To clean out some segmentation noises, we further apply median filter, morphological dilation and erosion filters to the bit map of the segmented tissue, where the final preprocessing result is shown in Fig. 1(c).

In our experiment, we used CT images from [18], [19] of size 512×512. Also, we chose $k = 15$ and $q = 7$ for the outer window and the inner window sizes, respectively ($M = k \times k = 225$).

In addition, for kernel function, we used the Gaussian radial basis function (RBF) given by

$$k(\mathbf{x}_i, \mathbf{x}_j) = \frac{\exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2)}{C}, \tag{13}$$

where C is a non-zero positive constant and is set to 5 in our experiment. The main reason for choosing Gaussian RBF kernel is the *translation-invariant* property of this kernel function [22]. As Equation (13) shows, the kernel function only depends on the difference between the inputs $\mathbf{x}_i$ and $\mathbf{x}_j$ and the absolute values of the vectors are not important.

We applied the kernel and non-kernel RX-algorithms and stored the maximum (max) and minimum (min) values of the RX function outputs. Then we set the threshold η as $0.1 \times min + 0.9 \times max$ for each slice of CT image. The pixels which had values more than η were marked as anomaly. We applied the proposed method on twelve sets of CT images with slice thickness from 2 to $5mm$. The kernel-RX algorithm is able to detect all nodules but there are four false positives. It is interesting to see that all false positives lie on the boundaries and so boundary effect could be an issue in the current implementation. This suggests a possible direction for future improvement.

In Fig. 2, we show an example result of one slice of CT image (Fig. (a)) after applying RX-algorithm (Fig. (b)) and kernel RX-algorithm with Gaussian RBF (Equation 13). Figs. (c)–(f) shows the result of the kernel RX-algorithm with different patches as described in Section IV. Fig. 3 shows the repeated experiement on a CT scan image of the same patient a year later. A cancer nodule was found and circled in the marked area based on LIDC-IDRI [19] annotated files. All kernel RX-algorithms detected the nodule but cubic N_C and mixture N_M patches result in outputs farther from the thresholds than the outputs of the linear N_L and pyramid N_P patches.

VI. Conclusion

In this paper, we presented an anomaly detector for cancer nodule based on the kernel RX-algorithm, which has been successfully adopted in military target detection in hyperspectral images. To apply to medical images, we modified the algorithm and in particular considered the effect of different sizes and shapes of data patches. Simple segmentation and additional speed up tricks have also been adopted to decrease the computational complexity of the algorithm. Preliminary results have suggested that the proposed approach could be an efficient technique for early lung cancer detection.

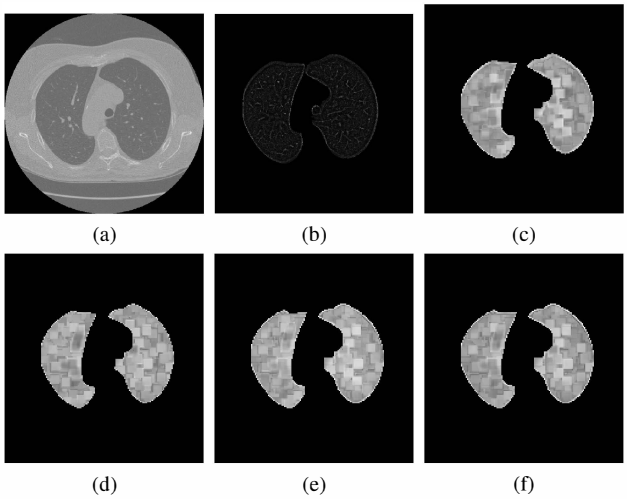


Fig. 2: The examples of processed CT image. a) original CT image; b) RX-algorithm output; c) kernel RX-algorithm output with N_L neighborhood; d) kernel RX-algorithm output with N_P neighborhood; e) kernel RX-algorithm output with N_C neighborhood; f) kernel RX-algorithm output with N_M neighborhood;

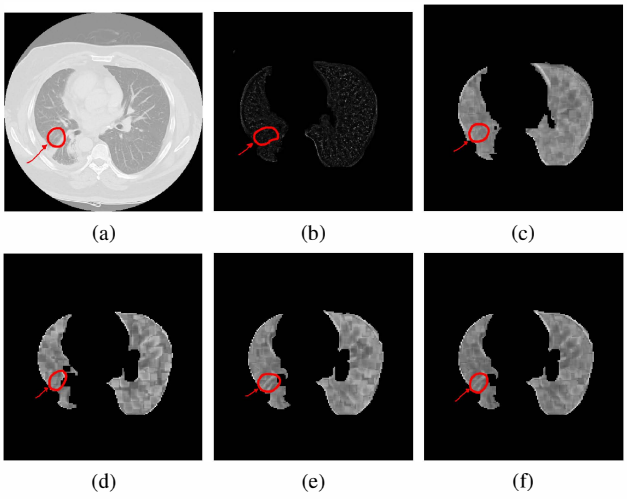


Fig. 3: The CT image of patient which has cancer in marked area one year later. a) original CT image; b) RX-algorithm output; c) kernel RX-algorithm output with N_L neighborhood; d) kernel RX-algorithm output with N_P neighborhood; e) kernel RX-algorithm output with N_C neighborhood; f) kernel RX-algorithm output with N_M neighborhood;

ACKNOWLEDGEMENT

The authors would like to thank Mrs. Nafise Barzigar for her valuable comments and her helps on algorithm implementation and debugging.

REFERENCES

[1] A. Jemal, R. Siegel, E. Ward, Y. Hao, J. Xu, T. Murray, and M.J. Thun, "Cancer statistics, 2008," *CA: a cancer journal for clinicians*, vol. 58, no. 2, pp. 71, 2008.

[2] D. Wormanns and S. Diederich, "Characterization of small pulmonary nodules by ct," *European radiology*, vol. 14, no. 8, pp. 1380–1391, 2004.

[3] S. Takashima, S. Sone, F. Li, Y. Maruyama, M. Hasegawa, T. Matsushita, F. Takayama, and M. Kadoya, "Small solitary pulmonary nodules (¡= 1 cm) detected at population-based ct screening for lung cancer: Reliable high-resolution ct features of benign lesions," *American Journal of Roentgenology*, vol. 180, no. 4, pp. 955, 2003.

[4] X. Ye, M. Siddique, A. Douiri, G. Beddoe, and G. Slabaugh, "Graph-cut based automatic segmentation of lung nodules using shape, intensity and spatial features," in *Proc. MICCAI09*. Citeseer, 2009.

[5] W. Mullally, M. Betke, J. Wang, and J.P. Ko, "Segmentation of nodules on chest computed tomography for growth assessment," *Medical physics*, vol. 31, pp. 839, 2004.

[6] S.G. Armato, M.L. Giger, C.J. Moran, J.T. Blackburn, K. Doi, and H. MacMahon, "Computerized detection of pulmonary nodules on ct scans1," *Radiographics*, vol. 19, no. 5, pp. 1303, 1999.

[7] M.N. Gurcan, B. Sahiner, N. Petrick, H.P. Chan, E.A. Kazerooni, P.N. Cascade, and L. Hadjiiski, "Lung nodule detection on thoracic computed tomography images: preliminary evaluation of a computer-aided diagnosis system," *Medical Physics*, vol. 29, pp. 2552, 2002.

[8] K. Doi, "Current status and future potential of computer-aided diagnosis in medical imaging," *British Journal of Radiology*, vol. 78, no. 1, pp. 3–19, 2005.

[9] R.J. van Klaveren, M. Oudkerk, M. Prokop, E.T. Scholten, K. Nackaerts, R. Vernhout, C.A. van Iersel, K.A.M. van den Bergh, S. van't Westeinde, C. van der Aalst, et al., "Management of lung nodules detected by volume ct scanning," *New England Journal of Medicine*, vol. 361, no. 23, pp. 2221–2229, 2009.

[10] B. van Ginneken, S.G. Armato III, B. de Hoop, S. van Amelsvoort-van de Vorst, T. Duindam, M. Niemeijer, K. Murphy, A. Schilham, A. Retico, M.E. Fantacci, et al., "Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: The anode09 study," *Medical Image Analysis*, vol. 14, no. 6, pp. 707–722, 2010.

[11] S.V. Fotin, A.P. Reeves, A.M. Biancardi, D.F. Yankelevitz, and C.I. Henschke, "A multiscale laplacian of gaussian filtering approach to automated pulmonary nodule detection from whole-lung low-dose ct scans," in *Proc. of SPIE Vol*, 2009, vol. 7260, pp. 72601Q–1.

[12] K. Murphy, B. van Ginneken, AMR Schilham, BJ de Hoop, HA Gietema, and M. Prokop, "A large-scale evaluation of automatic pulmonary nodule detection in chest ct using local image features and k-nearest-neighbour classification," *Medical image analysis*, vol. 13, no. 5, pp. 757–770, 2009.

[13] N. Homma, K. Takei, and T. Ishibashi, "Combinatorial effect of various features extraction on computer aided detection of pulmonary nodules in x-ray ct images," *WSEAS Transactions on Information Science and Applications*, vol. 5, no. 7, pp. 1127–1136, 2008.

[14] H. Arimura, S. Katsuragawa, K. Suzuki, F. Li, J. Shiraishi, S. Sone, and K. Doi, "Computerized scheme for automated detection of lung nodules in low-dose computed tomography images for lung cancer screening1," *Academic radiology*, vol. 11, no. 6, pp. 617–629, 2004.

[15] B. Schölkopf and A.J. Smola, *Learning with kernels*, The MIT Press, 2002.

[16] H. Kwon and N.M. Nasrabadi, "Kernel rx-algorithm: a nonlinear anomaly detector for hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 2, pp. 388–397, 2005.

[17] A.M. Thomas, "Extending the rx anomaly detection algorithm to continuous spectral and spatial domains," in *Southeastcon, 2008. IEEE*. IEEE, 2008, pp. 557–562.

[18] National Cancer Institute, "National biomedical image archive," https://imaging.nci.nih.gov/ncia/, May 2011.

[19] National Cancer Institute, "The cancer imaging archive," https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI, September 2011.

[20] I.S. Reed and X. Yu, "Adaptive multiple-band cfar detection of an optical pattern with unknown spectral distribution," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 38, no. 10, pp. 1760–1770, 1990.

[21] B. Schölkopf, A. Smola, and K.R. Müller, "Kernel principal component analysis," *Artificial Neural Networks*, pp. 583–588, 1999.

[22] C. Franke and R. Schaback, "Convergence order estimates of meshless collocation methods using radial basis functions," *Advances in Computational Mathematics*, vol. 8, no. 4, pp. 381–399, 1998.