



Complex-Valued Representation for RGB-D Object Recognition

Rim Trabelsi^{1,2,3(✉)}, Issam Jabri⁴, Farid Melgani⁵, Fethi Smach⁶,
Nicola Conci⁵, and Ammar Bouallegue²

¹ Advanced Digital Sciences Center, Singapore, Singapore
rim.t@adsc.com.sg

² SysCom Laboratory, National Engineering School of Tunis,
University of Tunis El Manar, Tunis, Tunisia

³ Hatem Bettaher IResCoMath Research Unit,

National Engineering School of Gabes, University of Gabes, Gabès, Tunisia

⁴ College of Computer and Information Systems, Al Yamamah University,
Riyadh, Kingdom of Saudi Arabia

⁵ Department of Information Engineering and Computer Science,
University of Trento, Trento, Italy

⁶ Profil Technology, 92120 Montrouge, France

Abstract. Object recognition methods usually tend to focus on single cues coming from traditional vision based systems but ignore to incorporate multi-modal data. With the advent of depth RGB-D sensors which provide synchronized multi-modal data with good quality, new opportunities have been emerged. In this paper, we make use of RGB and depth images to propose a new object recognition approach. Using a pixel-wise scheme, we propose a novel method to describe RGB-D images with a complex-valued representation. By means of neural network, we introduce a new CVNN (Complex-Valued Neural Network) with RBF neurons. Different from many RGB-D features, the proposed approach is able to jointly use RGB and depth data within a unified end-to-end learning framework. Category and instance object recognition tasks are evaluated through experiments carried out on a large scale RGB-D object dataset. Results show that our method can efficiently recognize objects in RGB-D images and outperforms state-of-the-art approaches.

Keywords: RGB-D representation · Object recognition
Complex-valued neural networks · Multi-modal data

1 Introduction

Object recognition is technologically challenging and practically useful problem in computer vision area owing to its wide spectrum of potential applications. This task deals with classifying an object into one of several predefined categories in an image. Strong solutions have been proposed in controlled environment [1, 2]. However, many issues still challenging until now with the presence

of color camouflage, cluttered backgrounds, objects occlusion and uncontrolled illumination. Most of the proposed attempts rely on traditional vision systems specifically on appearance data and their features. Recently, the development of 3D cameras and depth sensors have created new opportunities to advance the state-of-the-art of this field. In fact, depth information is less affected by those challenging matters. Yet, the extraction of depth may be itself affected by other issues including illumination changes. For this reason, the joint use of these multi-modal information, *i.e.*, appearance and depth, is very required to get robust features. With the advent of RGB-D sensors, depth maps can be extracted in real-time scenarios with good quality at low cost synchronized with RGB frames. Since the public release of RGB-D object dataset [3], a number of attempts have been made to recognize objects in RGB-D images [4–6]. Most of the proposed methods relied on region-based or holistic features that are combined in a trivial way from RGB and depth frames without joint fusion of the two modalities and ignores the particularity of depth maps and treat them the same way as appearance images.

In this paper, we propose to address the problem of object recognition in RGB-D images in a pixel-wise way. To this end, we introduced a new end-to-end strategy to classify images with a complex-valued representation. Inspired by the fact that point cloud, which corresponds to the mapping between RGB and depth images, could be easily seen as a complex-valued signal, we investigate complex-valued neural networks (CVNNs) to make use of both modalities in a joint way. Precisely, the main contributions of this paper are as follows. *(i)* A new RGB-D representation is proposed by projecting the real-valued data into the complex coordinate space where the depth is assumed as the imaginary part. *(ii)* Inspired by CVNNs [7, 8], a new end-to-end approach is introduced to solve the object recognition task in a pixel-wise fashion using RBF networks. *(iii)* Since RBF networks have a single hidden layer, their prototype vectors are here constructed using a K-means clustering algorithm with an adaptive method in order to fit complex-valued data. *(iv)* Evaluation of the proposed approach is finally evaluated over a large scale RGB-D dataset and compared with state-of-the-art methods.

The remaining of the paper is organized as follows. After reviewing related work in Sect. 2, we present the proposed method for object recognition with complex-valued representation in Sect. 3. Evaluation of two object recognition tasks over a large scale RGB-D dataset and comparisons with other state-of-the-art methods are reported in Sect. 4. Finally, in Sect. 5, the main contributions of the proposed approach in this paper are summarized.

2 Related Work

In this section, we will briefly highlight connections and differences between our approach and existing works mainly RGB-D representations designed for object recognition and since CVNNs are not employed so far to solve the target task, we present here a summary about their fundamental advances.

RGB-D based Representations. Using the RGB-D object recognition dataset published in 2011 [3], Bo et al. [9] succeeded to propose a new descriptor, named kernel descriptors, which enabled the use of multi-modal data by generalizing a set of features based on kernels. Lai et al. [10] proposed an efficient hierarchical classification approach where all hierarchy levels of the objects were used to enhance classification as well as pose estimation with stochastic gradient descent. In [11], the proposed method extracted hierarchical features from RGB-D images without supervision using hierarchical matching pursuit extended from [12].

Along with these hand-crafted methods, a quite interesting endeavors tried to adapt the revolutionary deep Convolution Neural Networks (CNNs) to fit RGB-D data. For example, using ImageNet pre-trained models, [13] proposed an architecture composed of two separate CNNs, one for the RGB and the other for the “D”. These two networks were combined with a late fusion network. An effective encoding to color space of depth images is proposed as well to fit model devoted to RGB images. Addressing objects detection problem, [5] come up with a new idea regarding the adaptation of depth information to the pre-trained color CNN model: the so-called HHA encoding. They extracted from depth image three channels at each pixel: horizontal disparity, height above ground, and the angle the pixels local surface normal makes with the inferred gravity direction. This representation has been intensively reused for further RGB-D tasks based CNN features.

Complex-Valued Neural Networks (CVNNs). In our daily lives, the large variety of information is dramatically increasing. It is hence expected to develop systems that process a wider range of information in more adaptive and effective ways just like human brain executes or better. So, this requires more suitable information’s representations. In order to make use of data with different modalities, we can model a couple of related real-valued signal as a complex-valued signal. With application to our context, we will later make use of such representation with visual 2D and 3D data. To this end, CVNNs were extended from the classic neural networks that we call here Real-Valued Neural Networks (RVNNs). CVNNs deal with information belonging to the complex coordinate space with complex-valued parameters and variables. “In relation to physicality, neural functions including learning and self-organization are influenced by sensorimotor interfaces that connect the neural network with the environment” [14]; this characteristic is of great importance also in CVNNs. Thus, there exist certain situations where CVNNs are inevitably required or greatly effective. Fundamental contributions to CVNNs were done by the pioneer Akira Hirose: the author of the first-ever concept of fully complex neural networks [15] and continuous complex-valued backpropagation [16] as well as a detailed survey of the critical concepts of CVNNs [14,17]. Regarding the learning algorithms for CVNNs, we should mention here the contributions of Fiori [18] which consists of generalizing the Hebbian learning for complex-valued neurons with an original optimization method which fits well CVNNs [19].

3 Multi-modal Representation by Means of Complex-Valued Neural Network

3.1 Overview

In this section, we are going to present the new multi-modal data representation and learning approach CVNN based. Our main goal is to build a robust representation of the image content that combines the advantages of the two modalities, *i.e.*, RGB and depth, to achieve high classification accuracy.

To formalize this learning problem, we consider the following notation. Let $\mathbb{X} \subset \mathbb{C}^m$ (\mathbb{C}^m is an m -dimensional complex coordinate space) be an input space and $\mathbb{L} = \{l_1, l_2, \dots, l_n\}$ be a finite real set of class labels. An instance $z \in \mathbb{X}$, represented in terms of features vector of dimension m as $z = [z_i]_{\{1 \leq i \leq m\}}$, is associated with a label $l \in \mathbb{L}$.

Let us also assume $\mathbb{T} = \{(z_i, l_i)\}_{i \in \{1, 2, \dots, n\}}$ a training set of n instances where $z_i \in \mathbb{X}$ and $l_i \in \mathbb{L}$. The purpose of this scheme of learning is to build a multi-class classifier: $\mathbb{M} : \mathbb{X} \rightarrow \mathbb{L}$ that optimizes some evaluation functions. To this end an RBF (Radial Basis Function) CVNN classifier is introduced.

3.2 Complex-Valued RBF-Networks

Motivation. By definition, an RBF is a function which has built into it a distance criterion with respect to a center. In the context of neural networks, the RBF function succeeded to replace the sigmoid activation function in multi-layer perceptron networks. The RBF neurons constitute the hidden layer units characterized by a center. In case where this function corresponds to a Gaussian, the network is trained by deciding on the number of hidden units/prototype vectors there should be as well as their centers and their sharpness (standard deviation), and then training up the output layer. Owing to the shallow yet wide architecture, RBF networks are able to extract a sparse model representation from a given training set. Motivated by these advantages of RBF-RVNNs, we propose to make use of RBF network to deal with complex-valued configuration for object recognition in RGB-D images. In the literature, this intuition has been exploited in other signal processing approximation and different classification tasks as in [20].

Architecture. Now, let us assume an RBF neural network that we call here RBF-CVNN which possesses complex-valued configuration (inputs, weights, activation functions, etc.). The network is composed of three layers: input, hidden and output layer as shown in Fig. 1. The input layer is composed of m nodes, each has a couple of inputs (a_i, b_i) . The functionality of this layer is to transform them into a complex values as $z_i(a_i, b_i) = a_i + \underline{j}b_i \forall i \in \{1, 2, \dots, m\}$ where \underline{j} is the imaginary unit, *i.e.*, the input prepared for the next layer is a complex-valued vector $z_i = (z_1, z_2, \dots, z_m)^T$. In what follows, $z_i(a_i, b_i)$ will be denoted by z_i .

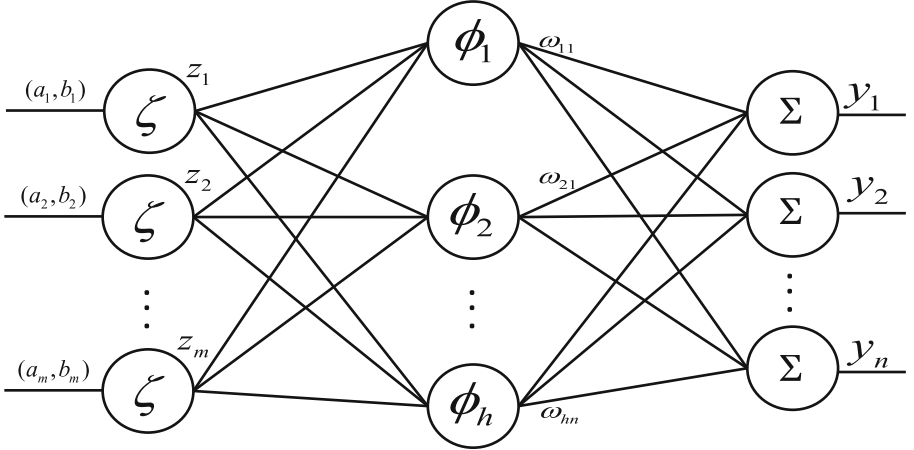


Fig. 1. Architecture of RBF-CVNN Network: composed of input, hidden and output layer (from left to right). The inputs of the network correspond to couples of real-valued numbers (a_i, b_i) , the ζ function converts each couple to a complex-valued number z_i which will be fed into the hidden neurons where a transformation will take place with a complex-valued activation function ϕ . Those prototype vectors have been choose using a specific clustering method as described in Sect. 3.2. A fully complex-valued gradient descent learning algorithm is exploited in order to learn weights between the hidden layer and the output layer which correspond a label vector of n category scores.

As for the hidden layer, it corresponds to the complex activation function RBF-based, specifically a gaussian-like one, defined as follows:

$$\phi_j(z_i) = \exp\left(\frac{-j\|z_i - c_j\|^2}{2\sigma_j^2}\right) \quad (1)$$

where $\|\cdot\|$ is the Euclidean distance, c_j is the center of the j^{th} hidden node and σ_j its corresponding variance. This function is suitable for CVNNs since it satisfies the property considered in [20, 21] which states that the fully complex non-linear activation function have to be analytic and bounded almost everywhere.

Using an unsupervised fashion, for a given number of hidden node h which corresponds to the number of prototype vectors, K-means [22] is used to determine their corresponding centers and means (c_j and σ_j). The clustering is proceeded on the training set \mathbb{T} using a specific setting. Instead of training K-means using couples of instance z_i and their corresponding label. To this end, a random instance is firstly chosen and assigned to the h neuron center. Then, for each element of the training set \mathbb{T} , the Euclidean distance is computed from each of the randomly chosen centers. Later, the instances of \mathbb{T} are clustered into h clusters depending on the minimum of the computed distance, *i.e.*, with objective to find:

$$\arg \min \sum_{j=1}^h \sum_{z_i \in \mathbb{T}} d(z_i, c_j) \quad (2)$$

with

$$d(z_i, c_j) = \sqrt{(z_i - c_j)(\overline{z_i - c_j})} \quad (3)$$

Next, the centers c_j are calculated as the mean of the instances belonging to each cluster. Once clustering is done, the distance between centers is checked and if it is less than the width of the cluster, those clusters will be joined together. This process is repeated until convergence, *i.e.*, until there is no changes of the values of c_j .

Regarding the output layer, it is constituted of n node, each one refereeing to the class label corresponding to the score of each category where the highest value is selected as the category. For a given instance/input vector z_i , the output vector Y is defined as $Y_i(z_i) = [y_k(z_i)]_{1 \leq k \leq n}$ where y_k are the score of z_i on the k^{th} category/class is given by Eq. (4). Each neuron output y_k is connected to all the h prototype vectors.

$$\begin{aligned} y_k(z_i) &= \sum_{j=1}^h \omega_{kj} \phi_j(z_i) \\ &= \sum_{j=1}^h (\text{Re}(\omega_{kj}) \text{Re}(\phi_j(z_i)) - \text{Im}(\omega_{kj}) \text{Im}(\phi_j(z_i))) \\ &\quad + \underline{j} (\text{Re}(\omega_{kj}) \text{Im}(\phi_j(z_i)) + \text{Im}(\omega_{kj}) \text{Re}(\phi_j(z_i))) \end{aligned} \quad (4)$$

In Eq. (4), ω_{kj} is complex-valued weight which is learned by minimizing the sum-squared errors (E) defined as:

$$E = \frac{1}{2} \sum_{i=1}^p \|T_i - Y_i\|^2 = \frac{1}{2} \sum_{i=1}^p \sum_{k=1}^n \|t_k - y_k\|^2 \quad (5)$$

where $T_i = [t_k]_{1 \leq k \leq n}$ and t_k are the target corresponding to z_i on the k^{th} class.

Using the fully complex-valued gradient descent learning algorithm proposed in [20], and according to Eq. (4), the update of output weights requires the differentiation of the E function with respect to ω_{kj} which allows us to obtain the following equation:

$$\frac{\partial E}{\partial \omega_{kj}} = -\overline{\phi_j} \frac{\partial E}{\partial y_k} \Leftrightarrow \Delta \omega_{kj} = \alpha \overline{\phi_j} \frac{\partial E}{\partial y_k} \quad (6)$$

where Δ is the delta rule, *i.e.*, a gradient descent learning rule for updating the weights here, α a complex-valued learning rate and $\overline{\phi_j}$ denotes the complex-conjugate of ϕ_j . Then, the update of the variance and the centers requires the differentiation of the E function with respect to the real and imaginary components of σ_j and c_j respectively, which allows us to write:

$$\Delta \sigma_j = \beta \overline{\phi_j} \left[\sum_{i=1}^p (\omega_{kj}^R \frac{\partial E}{\partial y_k^R} + \omega_{kj}^I \frac{\partial E}{\partial y_k^I}) \right] \frac{\|z_i - c_j\|^2}{\sigma_j^3} \quad (7)$$

$$\Delta c_j = \gamma \overline{\phi_j} \left[\frac{1}{\sigma_j^2} \sum_{i=1}^p (\omega_{kj}^R \frac{\partial E}{\partial y_k^R} \text{Re}(z_i - c_j) + \underline{j} \omega_{kj}^I \frac{\partial E}{\partial y_k^I} \text{Im}(z_i - c_j)) \right] \quad (8)$$

where β and γ are the learning rate parameters corresponding to σ_j and c_j respectively, ω_{kj}^R and ω_{kj}^I are the real and imaginary part of ω_{kj} respectively, Re and Im mean real and imaginary part respectively.

Thus, the fully complex-valued gradient descent learning algorithm allows us to update the parameters of our network ω , σ and c and correspond to each of them a learning rate parameter α , β and γ respectively.

3.3 Application to RGB-D Object Recognition

As reviewed earlier in the paper, significant advances have been made in quest of object recognition in RGB-D images, but much remains to be done, especially to improve the effective joint use of both modalities to take profit of their complementarities in a smarter way. To this end, we make use here of the proposed CVNN techniques explained above to define new solution using RGB-D images. RGB-D features have gained many computer vision tasks due to the complementarities between appearance and depth information. Here, we choose to investigate such type of data to enhance objects recognition using a joint pixel-wise classification strategy. Fusion between two different modes of data is done through complex-valued representation inspired by the fact that 3D point cloud, which corresponds to the mapping between RGB and depth images, could be easily seen as a complex-valued signal. Given a training set of n couples of RGB and depth images, we assume that each RGB-D image can be represented as a feature vector $z_i \in C_m$ in m dimensional space and assigned to a label l which corresponds to the instance category. Our objective is to obtain a robust description of z_i such that we can make use jointly of RGB and depth in an end-to-end classifier with higher accuracy. CVNN method is exploited with application to RGB-D object recognition using the same setting defined in Sect. 3.2.

4 Experimental Results

We evaluate our proposed RGB-D based CVNN approach using the large scale RGB-D object recognition dataset named ‘‘RGB-D Object Dataset’’ [3] with two evaluation settings: instance and category object recognition tasks. In fact, this dataset contains 41,877 images of common 300 household objects classified into 51 categories such as ‘‘Bowl’’, ‘‘Camera’’, ‘‘Hand towel’’, etc. Along with category labels, objects in this dataset are organized into instances: for example, the category ‘‘Food can’’ can be divided into physically unique instances like ‘‘Pepsi Can’’ and ‘‘Mountain Dew Can’’. RGB-D images were recorded in a multi-view scheme using Microsoft Kinect sensor (v.1) which provides RGB and depth images at a resolution of 640×480 . To be aligned the practices used in the literature, we follow the same evaluation process used in [3]. For category recognition, it consists of leaving one object instance out from each category for testing, and train models on the remaining objects, *i.e.*, 249 objects for training and 51 for testing at each trial. Reported results are obtained over a 10-fold cross validation procedure. As for instance recognition, we train models

on images captured from 30° and 60° elevation angles, and test them on the images of the 45° angle. Samples from RGB-D object dataset are provided in Fig. 2.

For better comparison with state-of-the-art approaches, we consider several baseline methods. Firstly, in order to prove the efficiency of using both RGB and depth data in a unified framework, we compare the results of the RGB-D based methods with their single-mode based variants, *i.e.*, RGB and depth separately. Then, to show the robustness of our complex-valued representation through neural networks, we compare it to a real-valued representation by means of RVNNs, specifically with an RBF-RVNN. Also, we compared our proposal to the state-of-the-art approaches coming from handcrafted features detailed earlier in the related work section: kernel descriptors [9], hierarchical matching pursuit (HMP) [12] and its unsupervised variant (U-HMP) [11].

Results for the category and instance recognition tasks are reported in Tables 1 and 2, respectively. It is clear that the use of multi-modal data is outperforming all single-based methods except for the RVNN baseline method where combining RGB and depth data in a trivial way is performing worse than its variant of single-based cues. Recognition methods proposed in [9, 11] outperform the proposed approach using single-based methods this is owing to their rich handcrafted feature and most important because our proposal is exclusively proposed to deal with RGB-D at once since its devoted to encapsulated RGB and depth in a unified way by means of complex-valued representation and using just a single data type will decrease its performance.

Regarding instance recognition, the best results is achieved by our proposal using RGB-D images and similarly to the above results our proposal is not able

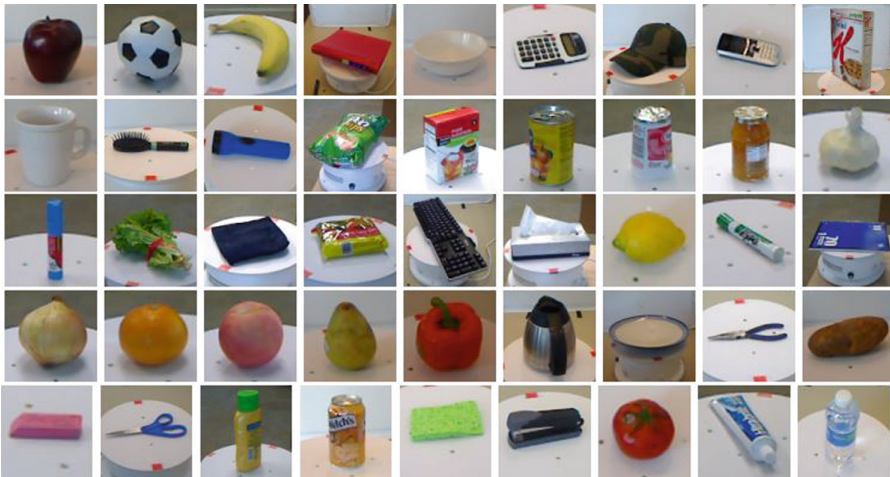


Fig. 2. Examples of objects with different categories presented in the large scale RGB-D objects dataset.

Table 1. Results for the category recognition task and evaluation against state-of-the-art method with different modalities: RGB, depth and RGB-D.

Approach	Modality		
	RGB	Depth	RGB-D
Kernel descriptors [9]	80.7 \pm 2.1	80.3 \pm 2.9	86.5 \pm 2.1
HMP [12]	74.7 \pm 2.5	70.3 \pm 2.2	82.1 \pm 3.3
U-HMP [11]	82.4 \pm 3.1	81.2 \pm 2.3	87.5 \pm 2.9
RVNN-baseline	76.8 \pm 2.6	75.0 \pm 3.0	71.8 \pm 3.2
CVNN (ours)	78.1 \pm 2.3	77.4 \pm 2.6	88.2 \pm 2.4

to cope with single-based modalities since it is designed for RGB-D data from the fine-grained information. It is notable here that depth data provides the worst results for all the approaches. This can be explained by the fact that objects belonging to the same category and different instances share in almost all the cases the same shape, however appearance in such cases will perform better.

Table 2. Results for the instance recognition task and evaluation against state-of-the-art method with different modalities: RGB, depth and RGB-D.

Approach	Modality		
	RGB	Depth	RGB-D
Kernel descriptors [9]	90.8	54.7	91.2
HMP [12]	75.8	39.8	78.9
U-HMP [11]	92.1	51.7	92.8
RVNN-baseline	83.7	50.9	69.2
CVNN (ours)	82.2	48.2	95.5

Finally, we can conclude that our RGB-D based representation is more robust for both instance and category tasks over the challenging large scale RGB-D dataset thanks to the complementarities between depth and appearance information and our fine-grained way of fusion and shows that it is able to deal with challenging images with texture-less items (like “bowls” or “apple”) or shape-less items (like “cereal boxes” or “hand towel”) captured under variation of viewpoint and variation of lighting conditions.

5 Conclusion

In this paper we addressed the problem of object recognition using multi-modal data. In contrast to majority of proposed recognition systems, we proposed a pixel-wise approach that fuse in early stage of learning process RGB and depth

data using a novel complex-valued representation within an end-to-end learning framework. An RBF layer is exploited in an adaptive way to construct a new CVNN network. Evaluation over the challenging large scale RGBD dataset is performed using two object recognition tasks shows that our proposal outperforms state-of-the-art methods. Increasing the number of layers and going deeper with our learning technique is very challenging but might be interesting and left for future work.

Acknowledgements. This work was supported by the European Union funding through ALYSSA program (ERASMUS-MUNDUS action 2 lot 6) and by the research grant from Singapore Agency for Science, Technology and Research (A*STAR) through the ARAP program.

References

1. Andreopoulos, A., Tsotsos, J.K.: 50 years of object recognition: directions forward. *Comput. Vis. Image Underst.* **117**(8), 827–891 (2013)
2. Bucak, S.S., Jin, R., Jain, A.K.: Multiple kernel learning for visual object recognition: a review. *IEEE Trans. Pattern Anal. Mach. Intell.* **36**(7), 1354–1369 (2014)
3. Lai, K., Bo, L., Ren, X., Fox, D.: A large-scale hierarchical multi-view RGB-D object dataset. In: 2011 IEEE International Conference on Robotics and Automation (ICRA), pp. 1817–1824. IEEE (2011)
4. Held, D., Thrun, S., Savarese, S.: Robust single-view instance recognition. In: 2016 IEEE International Conference on Robotics and Automation (ICRA), pp. 2152–2159. IEEE (2016)
5. Gupta, S., Girshick, R., Arbeláez, P., Malik, J.: Learning rich features from RGB-D images for object detection and segmentation. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8695, pp. 345–360. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10584-0_23
6. Li, X., Fang, M., Zhang, J.-J., Wu, J.: Learning coupled classifiers with RGB images for RGB-D object recognition. *Pattern Recogn.* **61**, 433–446 (2017)
7. Amin, M.F., Murase, K.: Single-layered complex-valued neural network for real-valued classification problems. *Neurocomputing* **72**(4), 945–955 (2009)
8. Savitha, R., Suresh, S., Sundararajan, N., Kim, H.J.: A fully complex-valued radial basis function classifier for real-valued classification problems. *Neurocomputing* **78**(1), 104–110 (2012)
9. Bo, L., Ren, X., Fox, D.: Depth kernel descriptors for object recognition. In: 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 821–826. IEEE (2011)
10. Lai, K., Bo, L., Ren, X., Fox, D.: A scalable tree-based approach for joint object and pose recognition. In: AAAI, vol. 1, p. 2 (2011)
11. Bo, L., Ren, X., Fox, D.: Unsupervised feature learning for RGB-D based object recognition. In: Desai, J., Dudek, G., Khatib, O., Kumar, V. (eds.) *Experimental Robotics*. Springer Tracts in Advanced Robotics, vol. 88, pp. 387–402. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-319-00065-7_27
12. Bo, L., Ren, X., Fox, D.: Hierarchical matching pursuit for image classification: architecture and fast algorithms. In: *Advances in Neural Information Processing Systems*, pp. 2115–2123 (2011)

13. Eitel, A., Springenberg, J.T., Spinello, L., Riedmiller, M., Burgard, W.: Multi-modal deep learning for robust RGB-D object recognition. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 681–687. IEEE (2015)
14. Hirose, A.: *Complex-Valued Neural Networks*. Springer Science & Business Media, Heidelberg (2006). <https://doi.org/10.1007/978-3-642-27632-3>
15. Hirose, A.: Dynamics of fully complex-valued neural networks. *Electron. Lett.* **28**(16), 1492–1494 (1992)
16. Hirose, A.: Continuous complex-valued back-propagation learning. *Electron. Lett.* **28**(20), 1854–1855 (1992)
17. Hirose, A.: *Complex-Valued Neural Networks: Theories and Applications*, vol. 5. World Scientific, Singapore (2003)
18. Fiori, S.: Nonlinear complex-valued extensions of Hebbian learning: an essay. *Neural Comput.* **17**(4), 779–838 (2005)
19. Fiori, S.: Learning by criterion optimization on a unitary unimodular matrix group. *Int. J. Neural Syst.* **18**(02), 87–103 (2008)
20. Savitha, R., Suresh, S., Sundararajan, N.: A fully complex-valued radial basis function network and its learning algorithm. *Int. J. Neural Syst.* **19**(04), 253–267 (2009)
21. Kim, T., Adali, T.: Fully complex multi-layer perceptron network for nonlinear signal processing. *J. VLSI Sig. Process. Syst. Sig. Image Video Technol.* **32**(1–2), 29–43 (2002)
22. Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., Wu, A.Y.: An efficient k-means clustering algorithm: analysis and implementation. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 881–892 (2002)