A Bag-of-Meaningful-Words Integrating Scattering and Structure Information for PolSAR Image Classification

Radu Tănase, European Union Satellite Centre, radu.tanase@satcen.europa.eu, Spain
Gottfried Schwartz, German Aerospace Centre, gottfried.schwartz@dlr.de, Germany
Mihai Datcu, German Aerospace Centre, mihai.datcu@dlr.de, Germany

Abstract

We propose a feature descriptor that integrates both the physical scattering properties of local targets and the contextual structure of HR PolSAR image patches. Firstly, the physical scattering properties are integrated by learning a meaningful scattering vocabulary, composed of the physically optimized entropy\anisotropy\alpha (H\A\$\boldsymbol{\alpha}$) classification labels, computed on small windows (e.g., of $\mathbf{5 \times 5}$ pixels). Then, in a Bag-of-Meaningful-Words (BoMW) fashion, we compute the histograms of these labels over the learned vocabulary on larger windows (e.g., of $\mathbf{64 \times 64}$ pixels), thus capturing the structure of image patches. We validated the quality of the obtained feature descriptors by computing *K*-Nearest-Neighbor (*K*NN) classifications of two HR, L-band, airborne polarimetric SAR images of the F-SAR and UAVSAR sensors, showing significant classification improvements compared with other state-of-the-art methods.

1 Introduction

The scattering matrix representation of PolSAR images embeds information about the physical interactions between the emitted waves and the ground targets. In order to exploit this information, several coherent and incoherent target decomposition theorems were proposed in the state-of-the-art literature. A relevant example is the H\A\$\boldsymbol{\alpha}$ decomposition and classification algorithm [1] which, when applied to individual pixels or to small neighbourhoods, reveals the different backscattering mechanisms within each resolution cell.

However, when dealing with image patches, the problem is more complex. In principle, the H\A\$\boldsymbol{\alpha}$ method can also be applied to larger windows, but only if these meet the stationarity condition. Otherwise, the averaging operator applied to the computation of the H, A, and $\boldsymbol{\alpha}$ parameters would cause a significant loss of information. The stationarity hypothesis can be taken for granted in the case of low-resolution data, but it no longer holds for the tremendous volume of currently available highly diverse HR data. Therefore, in this paper we present a solution for adapting the H\A\$\boldsymbol{\alpha}$ method to non-stationary, HR PolSAR image patches. Instead of computing the H\A\$\boldsymbol{\alpha}$ classification on large windows, we propose computing it on small neighbourhoods (e.g., of $\mathbf{5 \times 5}$ pixels), thus capturing the local scattering properties, and then retrieving the histograms of the resulting labels over larger windows (e.g., of $\mathbf{64 \times 64}$ pixels), thus capturing the contextual structure of the image patches (which could - in case of relatively homogeneous areas - be regarded as texture). In this way, we obtain a robust feature descriptor which integrates scattering with structure information, and which is not biased by the stationarity hypothesis of PolSAR image analysis. This BoW approach has been used before for texture representation and classification of remote sensing images [2] [3], but, as an innovation, here we apply it on a meaningful vocabulary (learned from the resulting H\A\$\boldsymbol{\alpha}$ classification labels), in which each word is related to physical backscattering properties (thus, the BoMW acronym).

In order to validate the quality of the proposed feature descriptor, we used it in supervised *K*NN classification tests on two HR images of the F-SAR (Traunstein AOI) and UAVSAR (Los Angeles AOI) instruments. Then, we compared our results with the ones of other state-of-the-art methods: the polarimetric supervised Wishart [4], the polarimetric unsupervised H\A\$\boldsymbol{\alpha}$-Wishart [5] and the traditional BoW technique, in which the vocabulary consists of the local statistics of the detected image.

2 H\A\$\boldsymbol{\alpha}$ classification

Our test images, two polarimetric, L-band, F-SAR and UAVSAR datasets, respectively, were delivered in the single-look complex scattering matrix format. When dealing with this kind of images, the first operation that has to be performed is to reduce the inherent speckle noise. In order to do this, we vectorized the scattering matrix into the Pauli basis, resulting in a target vector $\vec{\boldsymbol{k}}$, and then we computed the multi-looked coherence matrix $\langle \boldsymbol{T} \rangle$, resulting a $\mathbf{4480 \times 3712}$ pixel F-SAR image, and a $\mathbf{9600 \times 11520}$ pixel UAVSAR image, both having a ground resolution of approximately $\mathbf{1.5}$ m (Fig. 1).

For homogeneous, stationary areas, the three real-valued elements on the main diagonal of $\langle \boldsymbol{T} \rangle$ show the objects characterized by odd-bounce, even-bounce, and volume scattering. However, for HR, heterogeneous, non-stationary scenes, the physical properties of the recorded ground objects can no longer be extracted directly from the coherence matrix, thus incoherent decompositions of this matrix have to be computed. Such a decomposition is the well-known H\A\$\boldsymbol{\alpha}$ algorithm, introduced in [1],

ISBN 978-3-8007-4636-1 / ISSN 2197-4403

which relies on the eiegendecomposition of $\langle \boldsymbol{T} \rangle$. Here, we computed the rule-based H\A\$\boldsymbol{\alpha}$ classification of the two datasets on $\mathbf{5} \times \mathbf{5}$ pixel windows, resulting in two maps composed of 16 labels.

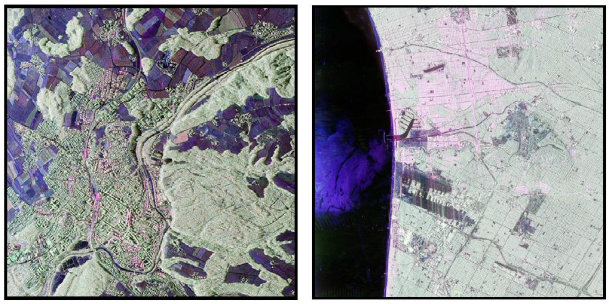

Figure 1: Pauli RGB representations of our multi-looked test images: F-SAR (left) and UAVSAR (right).

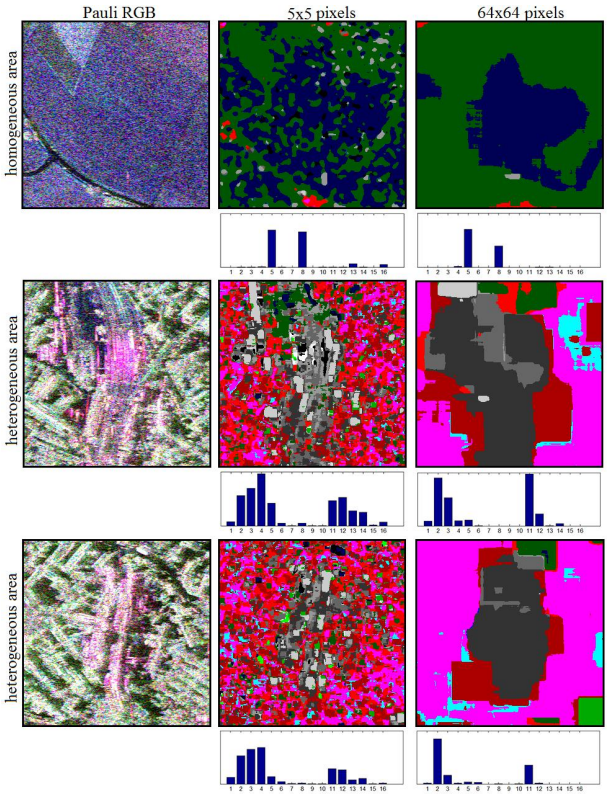

Figure 2: H\A\$\boldsymbol{\alpha}$ classification over both homogeneous (stationary) and heterogeneous (non-stationary) areas, using different window sizes. Over homogeneous areas (upper row), more or less similar results are obtained independently of the window size. However, over heterogeneous areas (middle and bottom rows) the use of a large window excludes the possibility of detecting local, strong backscattering targets like railway tracks or buildings. This is obvious from the label histograms shown beneath each classification map: a small window allows the detection of a large variety of scattering mechanisms, whereas a large window limits the number of scattering mechanisms that can be detected.

However, considering the huge amount of remote sensing image data, an urgent current challenge is to extend fast classification methods from single pixels to spatially extended patches. This challenge arises from the necessity of automatically classifying heterogeneous (non-stationary) ground areas that comprise various semantically meaningful categories, composed of multiple objects. H\A\$\boldsymbol{\alpha}$ classification can be successfully applied to small windows, but it is not suitable for large, non-stationary windows, because the averaging operator employed for computing $\langle \boldsymbol{T} \rangle$, H, A, and $\boldsymbol{\alpha}$ causes a significant loss of small-scale information. This fact is demonstrated in Fig. 2, on F-SAR image patches.

3 BoMW-H\A\$\boldsymbol{\alpha}$ classification

In computer vision, and particularly in image classification, the BoW model treats image features as words. Thus, a BoW representation of an image consists of a vector which counts the number of occurrences of each primitive feature in a vocabulary, i.e., its histogram over the vocabulary. Usually, the primitive features can be the simple pixel values [6], or they can be learned from common feature description methods like Gabor filters or local image statistics [2].

As an innovation, in the present paper we learn physically meaningful primitive features from the H\A\$\boldsymbol{\alpha}$ classification on small windows (of $\mathbf{5} \times \mathbf{5}$ pixels), resulting in a dictionary composed of $\mathbf{16}$ words. These primitive features reveal the local physical scattering properties of the recorded target areas. Then, in a BoMW approach, we compute the histogram of these meaningful primitive features over the vocabulary in order to integrate information on the structure of HR PolSAR image patches, which can be regarded as local texture within homogeneous areas. It was shown in [7] that the BoW model is appropriate for texture representation, given that for stochastic textures only the types of basic texture elements are important, and not their spatial distribution [8]. Therefore, by computing local H\A\$\boldsymbol{\alpha}$ classifications using a small sliding window, and then the BoMW on larger adjacent windows (of $\mathbf{64} \times \mathbf{64}$ pixels), we manage to overcome the non-stationarity issue by capturing both the local scattering properties and the structure of HR PolSAR image patches. For a HR PolSAR image, given in terms of its single-look complex scattering matrix $\boldsymbol{S}$, the proposed BoMW-H\A\$\boldsymbol{\alpha}$ feature extraction and classification algorithm follows the steps shown in Alg. 1.

Algorithm 1: The BoMW-H\A\$\boldsymbol{\alpha}$ classification alg.
Input : a HR SLC PolSAR image S
Output : a patch-based classification map
1: **function** BoMW-H\A\$\boldsymbol{\alpha}$ ($\boldsymbol{S}$)
2: compute $\vec{\boldsymbol{k}}$
3: compute $\langle \boldsymbol{T} \rangle$
4: compute eigendecomposition of $\langle \boldsymbol{T} \rangle$
5: compute H, A, $\boldsymbol{\alpha}$ parameters
6: compute px-based H\A\$\boldsymbol{\alpha}$ classif. map
7: split map into patches

ISBN 978-3-8007-4636-1 / ISSN 2197-4403

```
    for each patch in map do
        compute BoMW-H\A\α feature descriptor
    end for
    compute supervised KNN classification
    return new patch-based classification map
end function
```

First, we compute the target vector $\vec{k}$ and the multi-looked coherence matrix $\langle \boldsymbol{T} \rangle$ on windows resulting from averaging over several azimuth samples (three in the case of the F-SAR image and six in the case of the UAVSAR image, respectively). Then, following the algorithm presented in [1], we compute the global H\A\$\alpha$ classification map using $\mathbf{5} \times \mathbf{5}$ pixel windows. Next, we split the resulting map into non-overlapping adjacent patches, and for each patch we compute the BoMW-H\A\$\alpha$ feature descriptor, which is the histogram of the H\A\$\alpha$ classification labels over the vocabulary. The patch size was set to $\mathbf{64} \times \mathbf{64}$ pixels, as we consider that this is the minimum size that allows the detection of semantically meaningful categories in HR PolSAR images. In this way, we obtain robust feature descriptors that integrate both information about the local physical scattering properties, and about the spatial structure of HR PolSAR image patches. Finally, we perform a supervised *K*NN classification based on the resulting feature descriptors (*K* was set to one, and for training, 20 typical feature descriptors were selected for each category), obtaining a new, patch-based classification map. The results of the proposed algorithm, compared with the results of other state-of-the-art methods, are presented in the following Section.

4 Results and discussions

4.1 F-SAR dataset

In a first experiment, we considered the F-SAR image, which covers the Traunstein test area in southern Germany, and we selected the three most relevant categories, as follows: *urban* (urban areas, covered by buildings), *agriculture* (agricultural fields) and *forest* (forested areas). These categories differ significantly, both in terms of local scattering mechanisms and structure, as shown in Fig. 3. Then, in order to assess the performance of the proposed method, we classified the test image with the following five alternative algorithms:

1. The unsupervised H\A\$\alpha$-Wishart [5] ($\mathbf{64} \times \mathbf{64}$ pixel image patches);
2. A supervised Wishart [4], ($\mathbf{64} \times \mathbf{64}$ px);
3. A supervised BoW-MI, ($\mathbf{64} \times \mathbf{64}$ px): here the primitive features in the BoW model are the local mean pixel intensities within $\mathbf{5} \times \mathbf{5}$ pixel windows;
4. A supervised BoW-Var, ($\mathbf{64} \times \mathbf{64}$ px): here the primitive features in the BoW model are the local variances within $\mathbf{5} \times \mathbf{5}$ pixel windows;
5. The proposed BoMW-H\A\$\alpha$ ($\mathbf{64} \times \mathbf{64}$ px);

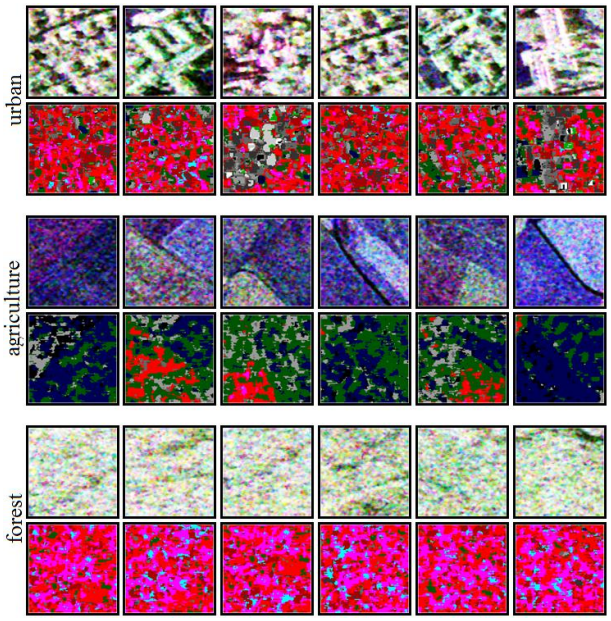

Figure 3: Examples of image patches belonging to three different categories: *urban*, *agriculture* and *forest*. For each category, the upper row consists of Pauli RGB representation of image patches, while the lower row consists of the corresponding H\A\$\alpha$ classification maps. It is obvious from these maps that the three categories are different both in terms of local scattering mechanisms (the presence of different labels) as well as structure (the distribution of labels). All this information is integrated into the histogram of labels over the vocabulary, which defines our proposed final BoMW-H\A\$\alpha$ feature descriptor.

For all the supervised algorithms presented above, we used **60** feature descriptors for training (**20** for each category), which were selected from the same locations in the image. Also, all feature descriptors were computed using non-overlapping adjacent image patches. The classification results are shown as confusion matrices in Table 1, and graphically in Fig. 4.

Table 1: Confusion matrices of 3 categories and 5 algorithms (F-SAR)

77.2 %	**Unsupervised H\A\α-Wishart** (t = 101″)						No. of
	urban		agriculture		forest		patches
urban	**534**	**49.2 %**	82	7.5 %	470	43.3 %	1086
agriculture	40	2.9 %	**1286**	**93.3 %**	52	3.8 %	1378
forest	157	9.8 %	126	7.9 %	**1313**	**82.3 %**	1596
79.3 %	**Supervised Wishart** (t = 10″)						
	urban		agriculture		forest		
urban	**600**	**55.3 %**	50	4.6 %	436	40.1 %	1086
agriculture	53	3.8 %	**1256**	**91.2 %**	69	5.0 %	1378
forest	146	9.1 %	87	5.5 %	**1363**	**85.4 %**	1596
78.6 %	**Supervised BoW-MI** (t = 59″)						
	urban		agriculture		forest		
urban	**934**	**86.0 %**	22	2.0 %	130	12.0 %	1086
agriculture	221	16.1 %	**1097**	**79.6 %**	60	4.3 %	1378
forest	400	25.0 %	38	2.4 %	**1158**	**72.6 %**	1596
77.0 %	**Supervised BoW-Var** (t = 127″)						
	urban		agriculture		forest		
urban	**873**	**80.4 %**	88	8.1 %	125	11.5 %	1086
agriculture	120	8.7 %	**1138**	**82.6 %**	120	8.7 %	1378
forest	196	12.3 %	287	18.0 %	**1113**	**69.7 %**	1596
88.0 %	**Supervised BoMW-H\A\α** (t = 138″)						
	urban		agriculture		forest		
urban	**958**	**88.2 %**	100	9.2 %	28	2.6 %	1086
agriculture	76	5.5 %	**1270**	**92.2 %**	32	2.3 %	1378
forest	153	9.6 %	99	6.2 %	**1344**	**84.2 %**	1596

ISBN 978-3-8007-4636-1 / ISSN 2197-4403

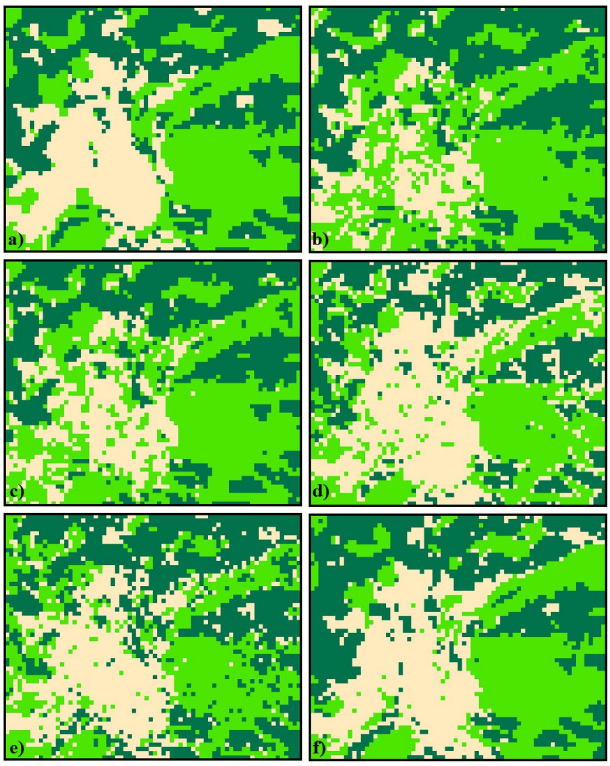

Figure 4: Classification results derived from non-overlapping image patches of $\mathbf{64} \times \mathbf{64}$ pixels. a) Manually annotated ground-truth; b) Unsupervised H\A\$\alpha$-Wishart; c) Supervised Wishart; d) Supervised BoW-MI; e) Supervised BoW-Var; f) Supervised BoMW-H\A\$\alpha$.

As can be seen in Table 1 and Fig. 4, the unsupervised H\A\$\alpha$-Wishart algorithm yields a very unsatisfactory true-positive classification rate for urban areas. This demonstrates the unsuitability of H\A\$\alpha$ classification on large image patches, especially when dealing with heterogeneous data. Slightly better results are achieved by the supervised Wishart algorithm, but still, too many *urban* patches are misclassified as *forest*, since this algorithm only relies on the mean backscatter within the considered image patches, disregarding their structure. The supervised BoW-MI and BoW-Var produce true-positive classification rates of at least 70% for all categories, which is a positive aspect. However, the overall classification accuracies of less than 80% indicate that these feature descriptors are limited by the relatively low performances of their primitive features. Finally, the supervised BoMW-H\A\$\alpha$ method yields an overall accuracy of 88%, exceeding the other methods by about 10%. In addition, all categories are correctly detected with a percentage of at least 84%. This result demonstrates the robustness of the proposed feature as it successfully integrates both information about the local scattering mechanisms and information about the structure of PolSAR image patches.
As a remark, the good classification results of our proposed method come at the cost of longer processing time. As expected, the fastest algorithm is the supervised Wishart, since it does not assume computing pixel-level primitive features. All other methods are significantly more time consuming, especially the BoMW-H\A\α, since computing the eigenvalues and the eigenvectors of the coherence matrix is much more complex than computing sample means or sample variances. However, we consider that the 10% gain in classification accuracy is worth the effort.
Further, a second experiment was performed, for more complicated scenarios, to attempt the detection of the following very heterogeneous, complex categories:

- Forest edges (*EDG*): boundary areas between forests and agricultural fields;
- Agricultural fields (*AGR*): areas covered by various crop types;
- Residential areas (*RES*): areas covered by houses and trees, characterized by regular geometries;
- Metal structure buildings (*MET*): areas covered by buildings that generate strong radar responses, mainly due to steel structures;
- Forest (*FOR*): forested areas.

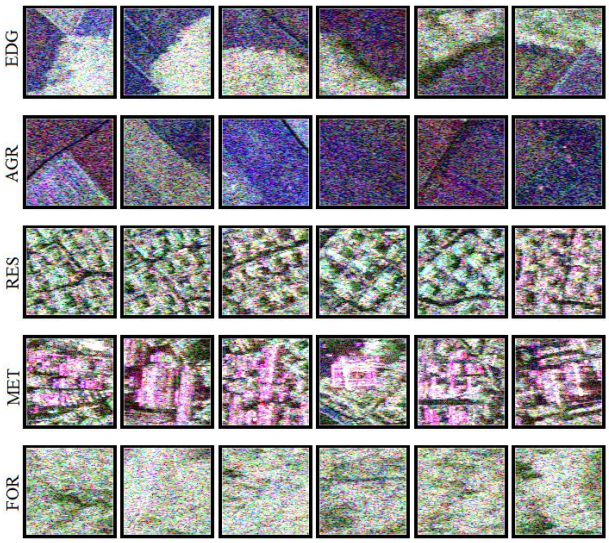

Figure 5: Examples of F-SAR image patches belonging to 5 different categories.

The goal of this experiment was to show that the proposed feature descriptors are able to characterize very complex categories that are composed of multiple objects. As a rule, the image patches belonging to these categories are highly non-stationary, as shown in Fig. 5.
In order to conduct the experiment, we selected $\mathbf{1000}$ representative image patches of $\mathbf{64} \times \mathbf{64}$ pixels ($\mathbf{200}$ for each category) and we tested all the supervised algorithms listed above. The unsupervised H\A\$\alpha$-Wishart algorithm was disregarded because it would have been very difficult to map its $\mathbf{16}$ classes to the five considered categories, and because it returned very unsatisfactory results in the first experiment. For training, we selected ten image patches for each category, while the *K*NN order was set to one. The classification results are shown in Table 2.

ISBN 978-3-8007-4636-1 / ISSN 2197-4403

Table 2: Confusion matrices of 5 categories and 4 algorithms (F-SAR)

75.0 %	a) Supervised Wishart										No. of patches
	EDG		AGR		RES		MET		FOR		
EDG	**124**	**62 %**	34	17 %	18	9 %	0	0 %	24	12 %	200
AGR	8	4 %	**190**	**95 %**	2	1 %	0	0 %	0	0 %	200
RES	70	35 %	0	0 %	**82**	**41 %**	6	3 %	42	21 %	200
MET	6	3 %	0	0 %	14	7 %	**168**	**84 %**	12	6 %	200
FOR	12	6 %	0	0 %	2	1 %	0	0 %	**186**	**93 %**	200
82.2 %	**b) Supervised BoW-MI**										
	EDG		AGR		RES		MET		FOR		
EDG	**136**	**68 %**	30	15 %	12	6 %	22	11 %	0	0 %	200
AGR	16	8 %	**172**	**86 %**	4	2 %	2	1 %	6	3 %	200
RES	4	2 %	0	0 %	**164**	**82 %**	32	16 %	0	0 %	200
MET	0	0 %	0	0 %	34	17 %	**166**	**83 %**	0	0 %	200
FOR	0	0 %	0	0 %	16	8 %	0	0 %	**184**	**92 %**	200
79.8 %	**c) Supervised BoW-Var**										
	EDG		AGR		RES		MET		FOR		
EDG	**156**	**78 %**	28	14 %	12	6 %	2	1 %	2	1 %	200
AGR	38	19 %	**152**	**76 %**	2	1 %	2	1 %	6	3 %	200
RES	34	17 %	0	0 %	**138**	**69 %**	28	14 %	0	0 %	200
MET	8	4 %	0	0 %	28	14 %	**164**	**82 %**	0	0 %	200
FOR	0	0 %	2	1 %	10	5 %	0	0 %	**188**	**94 %**	200
89.4 %	**d) Supervised BoMW-H\A\α**										
	EDG		AGR		RES		MET		FOR		
EDG	**156**	**78 %**	4	2 %	40	20 %	0	0 %	0	0 %	200
AGR	12	6 %	**188**	**94 %**	0	0 %	0	0 %	0	0 %	200
RES	8	4 %	0	0 %	**190**	**95 %**	0	0 %	2	1 %	200
MET	2	1 %	0	0 %	34	17 %	**160**	**80 %**	4	2 %	200
FOR	0	0 %	0	0 %	0	0 %	0	0 %	**200**	**100 %**	200

Again, the supervised BoMW-H\A\α method yields the best classification results, exceeding all other tested algorithms. The supervised Wishart achieves good results only for the categories *AGR* and *FOR*, which are rather homogeneous, and which are mainly characterized by single scattering mechanisms (surface and volume scattering, respectively). The other categories, which exhibit a more complex structure, are mostly misclassified. The next two algorithms (BoW-MI and BoW-Var) produce better results, since they also consider the structure of complex image patches. However, their performance will be reduced if we use inappropriate primitive features with limited representation capability. The overall 89.4% true-positive rate achieved by our proposed algorithm demonstrates that it efficiently handles both homogeneous as well as heterogeneous, non-stationary image patches as it considers local scattering information together with their spatial structure.

4.2 UAVSAR dataset

In order to further assess the performances of the proposed BoMW-H\A\α feature descriptor, we also tested it on a HR UAVSAR image covering the metropolitan area of Los Angeles, California, United States of America. This area stands out through its vast urban diversity, providing a completely different scenario from the first one. Therefore, to experience a classification, we identified the following heterogeneous, complex categories:

- Tall building residential areas (*TAL*): residential areas composed of high buildings, with poor vegetation, mainly characterized by multiple scatterings;
- Low building residential areas (*LOW*): residential areas composed of single-level houses, surrounded by trees, characterized by both multiple and volume scatterings;
- Motorways (*MOT*): motorways (or very wide streets) and their adjacent areas, which are characterized by surface scatterings and a mixture of some other scattering mechanisms;
- Water bodies (*WAT*): the ocean area west of the city;
- Recreational areas (*REC*): parks, composed of grasslands and trees, characterized by surface and volume scatterings.

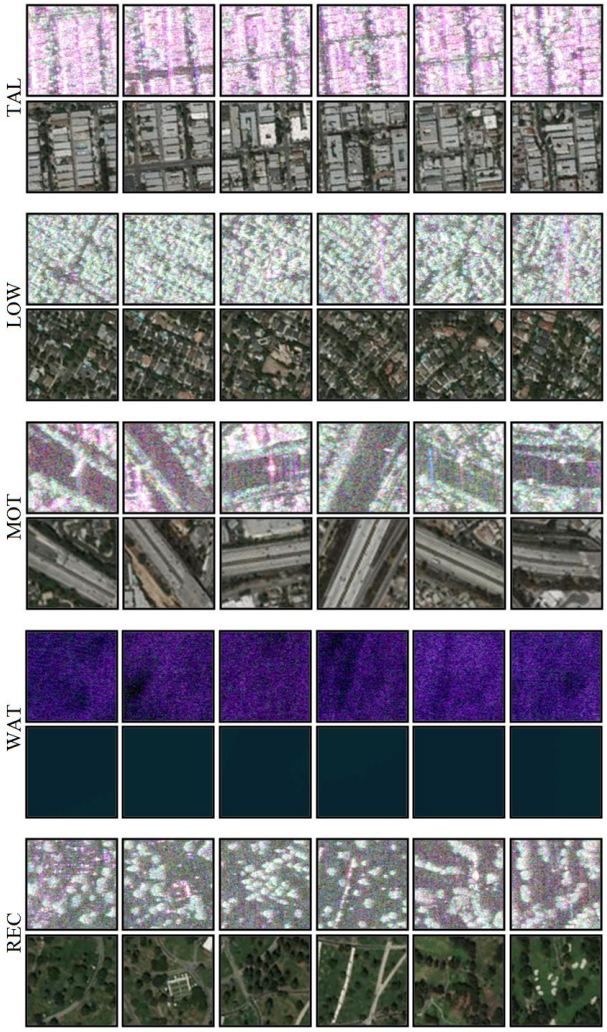

Figure 6: Examples of UAVSAR image patches belonging to 5 different categories, in Pauli RGB (upper rows) and optical (bottom rows) representations.

Some image patches belonging to the five above mentioned categories are shown in Fig. 6, in both Pauli RGB and optical representations. As we can notice in this figure, the five selected categories are very complex, composed of several different objects, thus a pixel-based classification would return meaningless results, in terms of semantics. Therefore, we experienced the same patch-based, supervised classification methods that were detailed in the previous Subsection. The confusion matrices are shown in Table 3.

ISBN 978-3-8007-4636-1 / ISSN 2197-4403

Table 3: Confusion matrices of 5 categories and 4 algorithms(UAVSAR)

85.7 %	a) Supervised Wishart										No. of patches
	TAL		LOW		MOT		WAT		REC		
TAL	**500**	**100 %**	0	0.0 %	0	0.0 %	0	0.0 %	0	0.0 %	500
LOW	13	2.6 %	**454**	**90.8 %**	33	6.6 %	0	0.0 %	0	0.0 %	500
MOT	16	5.3 %	42	14.0 %	**174**	**58.0 %**	0	0.0 %	68	22.7 %	300
WAT	0	0.0 %	0	0.0 %	0	0.0 %	**491**	**98.2 %**	9	1.8 %	500
REC	11	3.7 %	3	1.0 %	106	35.3 %	0	0.0 %	**180**	**60.0 %**	300
89.6 %	**b) Supervised BoW-MI**										
	TAL		LOW		MOT		WAT		REC		
TAL	**499**	**99.8 %**	1	0.2 %	0	0.0 %	0	0.0 %	0	0.0 %	500
LOW	1	0.2 %	**498**	**99.6 %**	1	0.2 %	0	0.0 %	0	0.0 %	500
MOT	0	0.0 %	2	0.7 %	**212**	**70.7 %**	0	0.0 %	86	28.6 %	300
WAT	0	0.0 %	0	0.0 %	1	0.2 %	**475**	**95.0 %**	24	4.8 %	500
REC	0	0.0 %	0	0.0 %	95	31.7 %	7	2.3 %	**198**	**66.0 %**	300
88.0 %	**c) Supervised BoW-Var**										
	TAL		LOW		MOT		WAT		REC		
TAL	**500**	**100 %**	0	0.0 %	0	0.0 %	0	0.0 %	0	0.0 %	500
LOW	0	0.0 %	**493**	**98.6 %**	7	1.4 %	0	0.0 %	0	0.0 %	500
MOT	0	0.0 %	2	0.7 %	**176**	**58.7 %**	0	0.0 %	122	40.6 %	300
WAT	0	0.0 %	0	0.0 %	0	0.0 %	**496**	**99.2 %**	4	0.8 %	500
REC	0	0.0 %	0	0.0 %	112	37.3 %	5	1.7 %	**183**	**61.0 %**	300
92.9 %	**d) Supervised BoMW-H\A\α**										
	TAL		LOW		MOT		WAT		REC		
TAL	**500**	**100 %**	0	0.0 %	0	0.0 %	0	0.0 %	0	0.0 %	500
LOW	0	0.0 %	**499**	**99.8 %**	0	0.0 %	0	0.0 %	1	0.2 %	500
MOT	0	0.0 %	2	0.7 %	**228**	**76.0 %**	0	0.0 %	70	23.3 %	300
WAT	0	0.0 %	0	0.0 %	0	0.0 %	**497**	**99.4 %**	3	0.6 %	500
REC	0	0.0 %	0	0.0 %	69	23.0 %	4	1.3 %	**227**	**75.7 %**	300

As shown in the Table 3, the proposed supervised BoMW-H\A\$\alpha$ algorithm returned a 92.9% true-positive rate of classification, surpassing all other experienced methods. Similar to the F-SAR image, the supervised Wishart method returned a poor true-positive rate for the complex, heterogeneous *MOT* and *REC* categories, because this method relies only on the mean backscattering mechanisms within image patches, disregarding their contextual structure. Further, the two traditional BoW techniques that we experienced (BoW-MI and BoW-Var) returned slightly better overall classification accuracies, since they are able to exploit the structure of PolSAR image patches, but still, the fact that they do not consider in any way the local backscattering mechanisms represents a major drawback. Finally, the best results returned by the BoMW-H\A\$\alpha$ algorithm proves the fact that integrating both information regarding the physical backscattering properties of local targets and the structure of image patches represents a major advantage when dealing with HR, non-stationary PolSAR image processing or analysis tasks.

5 Conclusions

A simple, yet very practical new method for PolSAR image content feature extraction and classification was proposed in this paper. This comes in the context of today's multitude of non-stationary HR images, when the accent falls on patch-based classification, and the traditional pixel-based algorithms are no longer suitable. The method integrates both the local scattering mechanisms present in PolSAR image patches as well as the structure of the patches; it can handle both homogeneous and heterogeneous (non-stationary) images. The information about local scattering is embedded in the H\A\α classification labels, while the information about the spatial structure of the image patches is embedded in the histogram of labels over the semantically meaningful vocabulary. Therefore, by combining these two types of information, we obtained a powerful, compact feature descriptor which, when tested with a *K*NN classifier on HR PolSAR datasets, returned very good results compared to other state-of-the-art algorithms.

Acknowledgements

We would like to thank Dr.-Eng. Rolf Scheiber from the Microwaves and Radar Institute of the German Aerospace Center for interesting discussions and for making available the F-SAR data used in this paper.
UAVSAR data courtesy NASA/JPL-Caltech.

References

[1] S. R. Cloude and E. Pottier. An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 35(1):68–78, Jan 1997.

[2] R. Bahmanyar, S. Cui, and M. Datcu. A Comparative Study of Bag-of-Words and Bag-of-Topics Models of EO Image Patches. *IEEE Geoscience and Remote Sensing Letters*, 12(6):1357–1361, June 2015.

[3] R. Tanase, R. Bahmanyar, G. Schwarz, and M. Datcu. Discovery of Semantic Relationships in PolSAR Images Using Latent Dirichlet Allocation. *IEEE Geoscience and Remote Sensing Letters*, 14(2):237–241, Feb 2017.

[4] J. S. Lee, M. R. Grunes, and R. Kwok. Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution. *International Journal of Remote Sensing*, 15(11):2299–2311, 1994.

[5] J. S. Lee, M. R. Grunes, T. L. Ainsworth, L. J. Du, D. L. Schuler, and S. R. Cloude. Unsupervised classification using polarimetric decomposition and the complex Wishart classifier. *IEEE Transactions on Geoscience and Remote Sensing*, 37(5):2249–2258, Sep 1999.

[6] S. Cui, G. Schwarz, and M. Datcu. Remote Sensing Image Classification: No Features, No Clustering. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(11):5158–5170, Nov 2015.

[7] O. G. Cula and K. J. Dana. Compact representation of bidirectional texture functions. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages I–1041–I–1047 vol.1, 2001.

[8] B. Julesz. Textons, the elements of texture perception, and their interactions. *Nature*, 290(5802):91–97, 1981.