ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9.
Available from http://www.i6doc.com/en/.

Multispectral Texture Classification in Agriculture

Mariya Shumska[1] and Kerstin Bunte[1]

1- University of Groningen - Johann Bernoulli Institute
for Mathematics and Computer Science, Groningen - the Netherlands

Abstract. Texture classification plays an important role in different domains including agricultural applications, where unmanned vehicles such as drones equipped with multispectral sensors are gaining more attention. Hence, a solution which does not require substantial computational resources is desired for real-time monitoring. In this contribution, we propose an efficient and interpretable Generalized Matrix Learning Vector Quantization based framework to classify multispectral images. We demonstrate the performance of different model designs and compare them to other benchmarks for the classification of a soil data set. Our framework yields comparable accuracy while providing interpretable results.

1 Introduction

Texture analysis is a branch of imaging science that aims to identify and quantify spatial patterns of pixels. Its methods are well suited for classification and segmentation tasks, as they provide unique information about the texture within the image region [1]. Particularly, texture classification is a topic of interest in agricultural applications, where remote sensing technologies collect data for crop monitoring [2, 3, 4]. A variety of methods have been developed for texture analysis including Gabor filtering [5] and co-occurrence matrices [6]. The state-of-art approaches such as Convolutional Neural Networks (CNN) have remarkable accuracy [7], however, they typically lack interpretability, demand large amounts of data for training, and require substantial computational resources. As unmanned aerial vehicles (UAVs) such as drones equipped with multispectral sensors are gaining more attention in the agricultural sector due to their flexibility [8], a more lightweight solution is desired for real-time monitoring.

The majority non-NN based texture classification methods are designed to operate on single-channel images, e.g. in the case of colour images, the input is first preprocessed with one of the standard RGB-to-greyscale transforms fixed for all classes. However, having local transformations can result in higher accuracy, as shown in [9] with the Colour Image Analysis Learning Vector Quantization (CIA-LVQ) framework. It bases on prototype learning with adaptive dissimilarities in the form of Generalized Matrix LVQ (GMLVQ) and a Gabor filter bank as a feature extractor. In [10, 11] CIA-LVQ was extended by an adaptive filter bank, which improved classification results even further.

In this work, we generalize CIA-LVQ to multispectral data, adjust the original adaptive dissimilarity measure and extend CIA-LVQ with a Parametrized Angle-based (PA) dissimilarity. We also demonstrate a special case of the transformation matrices, specifically designed for multichannel data which reduces the complexity while improving the generalization ability and explainability of

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

the models. Unlike [10, 9] we consider images in the spatial domain to allow more intuitive interpretation and discard the filtering operation as in contrast to [11], the focus of this work is the impact of dissimilarity measure and matrix format. Our results are comparable to benchmarks while being interpretable.

2 Methodology

The CIA-LVQ operates on patches from the original data. Without loss of generality, we consider all h channels and hence a patch of size $p \times p$ has the dimensionality of $n = p^2h$. The patches are vectorized and channel-wise concatenated to form feature vectors. Unlike [9, 10] we consider real input and no filtering, hence, analyze the use of Angle LVQ [12] for multispectral texture classification and propose a block-diagonal parameterization for channels.

2.1 LVQ for Multi-Channel Intensities

Learning Vector Quantization (LVQ) [13] is a supervised algorithm using the winner-takes-all scheme, in which a data point is classified according to the label of its closest prototype. Throughout the following, we assume a labelled training data set $\{(\mathbf{x}_i, y_i) \mid \mathbf{x}_i \in \mathbb{R}^n, \text{ and } y_i \in \{1, \ldots, C\}\}_{i=1}^N$ and a set of prototype vectors $\mathbf{w}^j \in \mathbb{R}^n$ with labels $c(\mathbf{w}^j) \in \{1, \ldots, C\}$. In contrast to the original heuristic prototype update the Generalized LVQ [14] introduced a training scheme as minimization of cost function:

$$E = \sum_{i=1}^{N} \Phi(\mu_i), \ \mu_i = \frac{d_i^J - d_i^K}{d_i^J + d_i^K} \ , \tag{1}$$

with distance $d_i^J = d(\mathbf{x}_i, \mathbf{w}^J)$ to the closest prototype $\mathbf{w}^J$with the same class label $y_i = c(\mathbf{w}^J)$ and the distance $d_i^K = d(\mathbf{x}_i, \mathbf{w}^K)$ to the closest prototype with non-matching label $y_i \neq c(\mathbf{w}^K)$. Φ is a monotonic function and we use the identity function in this contribution. The definition of d plays a central role in LVQ-based classifiers, as it determines the closest prototypes. In this paper, the quadratic form and angle-based dissimilarities are considered. These pseudo-metrics are not guaranteed to satisfy the triangle inequality, but we address them as "distances"/"dissimilarities" throughout this paper for readability.

GMLVQ [15] makes the distance adaptive by employing a positive semi-definite $n \times n$ matrix Λ, accounting for the pair-wise correlation between the features. To ensure positive semi-definiteness Λ can be decomposed as $\Lambda = \Omega^T\Omega$ with $\Omega \in \mathbb{R}^{m \times n}$, $m \leq n$. The corresponding quadratic form (QF) is defined as:

$$d_{QF}^{\Omega}(\mathbf{x}, \mathbf{w}) = (\mathbf{x} - \mathbf{w})^T \Omega^T \Omega (\mathbf{x} - \mathbf{w}) \ , \tag{2}$$

where Ω is learned along with the prototypes. Rectangular matrices Ω with $m < n$ imply dimensionality reduction by a linear transformation. In this paper, we adopt rectangular Ω's with $m = {}^n\!/_h$ to obtain a "quasi-greyscale" representation of the original multispectral patch and have a possibility to interpret Λ as the correlation matrix of spatio-spectral features. In addition to a full matrix we introduce the new option developed for channel intensity images (both shown in Figure 1). We term it a block-diagonal matrix:

$$\widehat{\Omega}_{ij} = \begin{cases} \widehat{\Omega}_{ij}, & \text{if } j = i + (l-1)m \text{ with } l = 1, \ldots, h \\ 0, & \text{otherwise .} \end{cases} \tag{3}$$

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

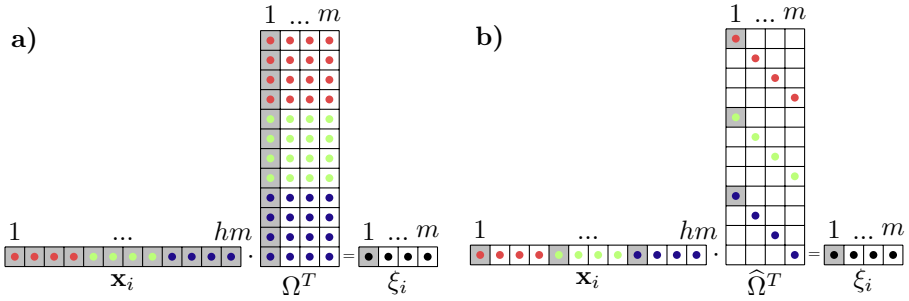

Fig. 1: $\mathbf{x}_i$ is projected with a full Ω^T (**a**) or block-diagonal $\widehat{\Omega}^T$ (**b**) transformation matrix resulting in a lower-dimensional representation ξ_i. Shading indicates contributions to the shaded element of ξ_i and empty cells indicate zero weights.

Hence, every row $\widehat{\Omega}_i$ contains only contributions of the same pixel from each of h channels reducing the number of free matrix parameters from nm to hm, which can prevent the risk of overfitting. Instead of a global transformation local matrices Ω^L or $\widehat{\Omega}^L$ attached to each prototype or class can be trained, changing the piece-wise linear decision boundaries into nonlinear ones.

2.2 Parametrized Angle Dissimilarity

A very recent extension of GMLVQ named Angle LVQ (ALVQ) [12] introduced a parameterized angle (PA) distance, that demonstrated very robust behaviour for heterogeneous data and imbalanced classes. The dissimilarity is defined as:

$$d_{PA}(\mathbf{x},\mathbf{w}) = g(b^{\Omega}(\mathbf{x},\mathbf{w}),\beta), \qquad \text{where } g(b^{\Omega},\beta) = \frac{e^{-\beta(b-1)}-1}{e^{2\beta}-1}\ , \tag{4}$$

$$\text{and } b^{\Omega}(\mathbf{x},\mathbf{w}) = \frac{\mathbf{x}^T\Omega^T\Omega\mathbf{w}}{\|\mathbf{x}\|_{\Omega}\|\mathbf{w}\|_{\Omega}}, \qquad \text{with } \|\mathbf{v}\|_{\Omega} = \sqrt{\mathbf{v}^T\Omega^T\Omega\mathbf{v}}\ . \tag{5}$$

The g function in (4) transforms the parameterized cosine similarity $b^{\Omega} = \cos\theta \in [-1,1]$ into a dissimilarity $\in [0,1]$. The hyperparameter β controls the slope weighting the contribution of samples within the receptive field based on their distance to the prototype. The angle-based distance classifies on the hypersphere instead of Euclidean space and hence does not consider the magnitude of vectors. Optimization of the prototypes and Ωs occurs through minimization of the non-convex cost function in eq. (1) by gradient methods, such as stochastic gradient descent or conjugate gradient. The corresponding partial derivatives for QF- and PA-based GMLVQ can be found in [15] and [12], respectively. Normalization of Ωs [15] and regularization [16] of E has been also proposed.

3 Experiments

To demonstrate the application of our framework to multispectral data, we use Statlog data set [17] of agricultural land. Each entry $\mathbf{x}_i$ is a 3×3 pixel neighbourhood containing information from $h = 4$ spectral bands: two of these are in the visible region (correspond approximately to green and red spectrum) and two are in the (near) infrared, yielding the dimensionality $n = 3^2 \cdot 4 = 36$. The

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

Fig. 2: Visualisation of first two (visible) bands for patches from different classes.

provided labels y_i are associated with the central pixel and belong to one of the 6 unbalanced soil classes, namely red (24.2%), cotton (10.8%), grey (21.7%), damp grey (9.4%), vegetation stubble (10.6%), and very damp grey (23.4%), which are shown in Figure 2. To correct for class imbalance, we apply weighting to the samples' contribution to the cost function as explained in [18]. 6435 data points were divided into training (4435) and test (2000) set by the author.

We experiment with 4 main types of models which differ in the distance measure d, and the transformation matrix Ω configuration. In all cases, we use local Ω^js with $m = {}^n/_h = 9$ so that the matrices project the original data into an 'intensity' image. The hyperparameters such as the number of prototypes per class k, β, regularization strength and presence of normalization are estimated with the grid search with 5-fold cross-validation on the training data. The best-performing model of each type is then evaluated on a test set.

4 Results and Discussion

The results of the experiments can be found in Table 1. Higher accuracy can be achieved with a larger number of prototypes and prototype-wise Ω^js, but we restrict them to be class-wise, i.e. Ω^j $(j = c(\mathbf{w}^L))$, and limit maximum prototypes' number to 3, reducing model complexity. According to our results, the best-performing model was trained with QF dissimilarity, full matrix, and 3 prototypes per class. Its counterpart with block-diagonal $\widehat{\Omega}^j$ shows rather similar results despite having lower complexity in terms of the number of trainable parameters. The PA-based models provide lower accuracy, struggling to differentiate between grey, damp grey and very damp grey soil types. These classes are the most problematic for all models, as can be seen, in Figure 3. It is likely due to their similarity as the names suggest and damp soil being the minority class. Therefore, further experimentation with weighing costs associated with data belonging to these classes is needed. We also compare our results to other publicly available benchmarks on Statlog. In [19], the authors report a test accuracy of 82.57% obtained with a NN. The results of the multilayer perceptron backpropagation neural network (MLP) (89.3%), support vector machine (SVM) (85.1%), and k-nearest neighbour (k-NN) (89.0%) are shown in [20]. Our best model achieves either higher or comparable accuracy while offering insights into the data. Specifically, correlation matrices after QF-based training which are shown in Figure 4 provide information on spatio-spectral feature relations specific to each class. For instance, a strong positive correlation can be observed within the visible red band and a negative correlation between the red and fourth infra-red bands of the grey soil class. While for very damp grey soil, the strongest positive correlation is found within the visible green channel. Moreover, since

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

Table 1: Accuracy and (std) in %. Last column contains hyperparameters and whether regularization and matrix normalization was used. Best result marked **bold**.

d	Ω	Val.	Train	Test	
QF	Full	88.86 (0.96)	89.81	**88.20**	$k = 3$, reg
QF	Block	87.84 (0.79)	88.28	87.70	$k = 2$, norm
PA	Full	81.40 (0.66)	86.52	81.40	$k = 2$, $\beta = 2$
PA	Block	80.25 (1.37)	83.47	79.90	$k = 3$, $\beta = 1$

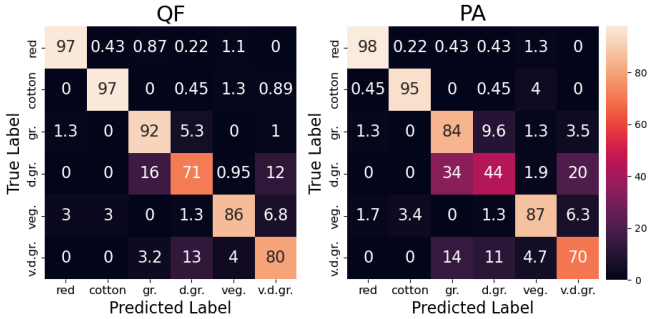

Fig. 3: Confusion matrices showing class-wise accuracy for models with full Ω^Ls.

transforming learned prototypes with corresponding Ω^j yields a single-channel output which can be visualized as an intensity image, the original multispectral input can have a more insightful depiction than if a fixed transform was used. The number of free parameters is $C(nm+kn)$ for full Ω^j or $C(n+kn)$ for block-diagonal $\widehat{\Omega}^j$ with C number of classes. For our best models, it totals to 2592 and 648, respectively. We estimate that the NN of [19] with two 65-neuron hidden layers has 6955 trainable parameters and 36-45-6 MLP of [20] has 1890. Hence, our framework, especially if block-diagonal matrices are used, is efficient to train and less likely to overfit. Moreover, the classification speed is linear with the total number of prototypes, which is much smaller than that of the reference set of k-NN or the number of support vectors in SVM. Finally, the LVQ model is transparent and allows further interpretation than presented in this work (see e.g. [18] for an alternative to classification activation maps).

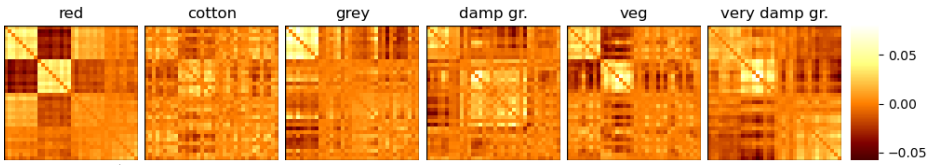

Fig. 4: Λ^j after training with diagonal elements set to 0. The quadrants show within- (on diagonal) and between-band (off-diagonal) correlation.

5 Conclusion and Future Work

In this work, we proposed an efficient and interpretable block-diagonal dissimilarity extension for the Generalized Matrix LVQ framework and demonstrated its application to agricultural multispectral images. We experimented with two adaptive distance measures that operate in Euclidean space or on a hyper-sphere,

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

and two matrix configurations. We conclude that for the Statlog data set classification in Euclidean space using the quadratic form provides the highest accuracy, while full and block-diagonal configurations of GMLVQ matrices give very similar results. Our models perform comparably with previously reported accuracy from NNs, SVM, and k-NN, while being interpretable, less complex and providing a way to visualize multispectral data. In future work, we will incorporate the learned texture similarities in morphological filter segmentation for images with intensity values in multiple channels, such as colour and hyperspectral images.

References

[1] W. Nailon. *Texture Analysis Methods for Medical Image Characterisation*, pages 75–100. IntechOpen, 2010.

[2] P. Kupidura. The comparison of different methods of texture analysis for their efficacy for land use classification in satellite imagery. *Remote Sens.*, 11(10), 2019.

[3] K. R. Gavhale, U. Gawande, et al. An overview of the research on plant leaves disease detection using image processing techniques. *IOSR J. Comput. Eng.*, 16(1):10–16, 2014.

[4] G.-H. Kwak and N.-W. Park. Impact of texture information on crop classification with machine learning and uav images. *Appl. Sci.*, 9:643, 2019.

[5] I. Y. Fogel and D. Sagi. Gabor filters as texture discriminator. *Biological Cybern.*, 61:103–113, 1989.

[6] R. M. Haralick, K. Shanmugam, and I. H. Dinstein. Textural features for image classification. *IEEE Trans. on syst., man, and cybern.*, SMC-3(6):610–621, 1973.

[7] S. Kumar and A. Gupta. Comparative review of machine learning and deep learning techniques for texture classification. In *Proc. of the Int. Conf. on Artif. Intell. Techniques for Elect. Eng. Syst.*, pages 95–112. Atlantis Press, 2022.

[8] L. Deng, Z. Mao, X. Li, Z. Hu, F. Duan, and Y. Yan. Uav-based multispectral remote sensing for precision agriculture: A comparison between different cameras. *Journal of Photogrammetry and Remote Sens,*, 146:124–136, 2018.

[9] K. Bunte, I. Giotis, N. Petkov, and M. Biehl. Adaptive matrices for color texture classification. In *CAIP*, 2011.

[10] I. Giotis, K. Bunte, N. Petkov, and M. Biehl. Adaptive matrices and filters for color texture classification. *Journal of Math. Imag. and Vis.*, 48:202–202, 2014.

[11] G. Luimstra and K. Bunte. Adaptive gabor filters for interpretable color texture classification. In *Eur. Symp. on Artif. Neural Netw*, pages 61–66. ESANN, 2022.

[12] S. Ghosh, P. Tiňo, and K. Bunte. Visualization and knowledge discovery from interpretable models. In *Int. Joint Conf. on Neural Netw.*, pages 1–8, Glasgow, UK, 2020. IEEE.

[13] T. Kohonen. *Learning Vector Quantization*, pages 175–189. Springer Berlin Heidelberg, Berlin, Heidelberg, 1995.

[14] A. Sato and K. Yamada. Generalized learning vector quantization. In D. Touretzky, M. Mozer, and M. Hasselmo, editors, *NeuIPS*, volume 8. MIT Press, 1995.

[15] P. Schneider, M. Biehl, and B. Hammer. Adaptive Relevance Matrices in Learning Vector Quantization. *Neural Comput.*, 21(12):3532–3561, 2009.

[16] P. Schneider, K. Bunte, H. Stiekema, B. Hammer, T. Villmann, and M. Biehl. Regularization in matrix relevance learning. *IEEE trans. neural netw.*, 21(5):831–840, 2010.

[17] A. Srinivasan. Statlog (landsat satellite) data set, 1993.

[18] S. Ghosh, E. S. Baranowski, M. Biehl, W. Arlt, P. Tino, and K. Bunte. Interpretable models capable of handling systematic missingness in imbalanced classes and heterogeneous datasets, 2022.

[19] M. L. Tej and S. Holban. Determining neural network architecture using data mining techniques. In *2018 Int. Conf. on Development and Appl. Syst.*, pages 156–163, 2018.

[20] S. Ghosh, S. Biswas, D. C. Sarkar, and P. P. Sarkar. A tutorial on different classification techniques for remotely sensed imagery datasets. *Smart Comput. Rev.*, 4:34–43, 2014.