



# Leveraging colour-based pseudo-labels to supervise saliency detection in hyperspectral image datasets

Annalisa Appice<sup>1,2</sup> · Angelo Cannarile<sup>1</sup> · Antonella Falini<sup>1</sup> · Donato Malerba<sup>1,2</sup> · Francesca Mazza<sup>1</sup> · Cristiano Tamborrino<sup>3</sup>

Received: 11 December 2020 / Revised: 11 June 2021 / Accepted: 12 July 2021 /

Published online: 11 August 2021

© The Author(s) 2021

## Abstract

Saliency detection mimics the natural visual attention mechanism that identifies an imagery region to be salient when it attracts visual attention more than the background. This image analysis task covers many important applications in several fields such as military science, ocean research, resources exploration, disaster and land-use monitoring tasks. Despite hundreds of models have been proposed for saliency detection in colour images, there is still a large room for improving saliency detection performances in hyperspectral imaging analysis. In the present study, an ensemble learning methodology for saliency detection in hyperspectral imagery datasets is presented. It enhances saliency assignments yielded through a robust colour-based technique with new saliency information extracted by taking advantage of the abundance of spectral information on multiple hyperspectral images. The experiments performed with the proposed methodology provide encouraging results, also compared to several competitors.

**Keywords** Saliency detection · Remote sensing · Imagery data analysis · Earth's observation

## 1 Introduction

Hyperspectral (HS) remote sensing is a major breakthrough in remote sensing technology for Digital Earth (Fu et al., 2020). HS sensors, mounted on aircraft or satellites, produce digital images (HyperSpectral Images - HSI) of an observed scene, by recording reflected light in hundreds of narrow frequencies covering the visible, near-infrared and shortwave infrared bands (pixel spectrum) (Appice & Malerba, 2019; Bioucas-Dias et al., 2013; Hoye & Fridman, 2013; Stuart et al., 2019). Such an abundance of spectral data represents an invaluable source of knowledge regarding the physical nature of the different materials possibly observed. In particular, the high-dimensionality of the measured pixel spectrum may be useful for saliency analysis in HS imagery data, in order to possibly characterise a landscape having higher contrast with its surroundings (Liang et al., 2013).

---

✉ Annalisa Appice  
annalisa.appice@uniba.it

In general, the concept of saliency refers to identifying parts, regions, objects or features that first draw visual attention and, hence, can be considered notable and important. In remote sensing, saliency analysis mimics human intelligence and detects the most attracting objects or pixels in a sensed scene. This analysis has great potential in many fields such as military science, ocean research, resources exploration, disaster and land-use monitoring. In fact, boosted by a variety of applications (e.g. scene understanding, geological and environmental monitoring and precision-guided system), various studies in imagery saliency detection have been progressed better and better in recent years (Borji et al., 2019). These studies commonly describe machine learning-based techniques that address the task of saliency detection as a binary pixel labelling problem. In general, these techniques analyse the pixel data of an input image to learn a binary saliency matrix that reveals foreground and background imagery pixels, respectively.

Despite HS imaging techniques have been progressively gaining popularity in the recent Earth observation, current saliency detection technologies have achieved real maturity in the analysis of traditional colour images. The current state-of-the-art of literature mainly describes techniques for estimating saliency in colour data. With the recent boom of deep learning, we can already count on various powerful deep learning models (Hou et al., 2019; Wang et al., 2020; Gao et al., 2019; Favorskaya & Jain, 2019; Liu et al., 2020; Luo et al., 2020) that have been accurately trained on a big amount of colour images and can be applied to estimate accurate saliency in new colour data. Nevertheless, a few studies have occurred on the stage a few years ago to start the investigation of saliency detection in HS images (e.g. Liang et al., 2013; Cao et al., 2015; Le Moan et al., 2013; Imamoglu et al., 2018; Du & Zhang, 2014; Du et al., 2016; Yan et al., 2016; Zhang et al., 2018; Falini et al., 2020; Appice et al., 2020).

The spectral variability and the curse of dimensionality are the major challenges to deal, in order to properly analyse HS imagery data. The spectral variability is commonly caused by many conditions such as incident illumination, atmospheric effects, unwanted shade and shadow, natural spectrum variation and instrument noises. It has been indicated as one of the main issue of HS classification learning (Appice & Malerba, 2019) and recently of HS saliency detection (Zhang et al., 2018). On the other hand, the curse of dimensionality, that is due to the abundance of information in HS spectrum (Hughes, 1968; Appice & Malerba, 2019), prevents us from processing HS imagery data by simply applying one of the powerful techniques designed for saliency detection in colour images. A naive approach to fit colour imagery techniques to HS imagery data is based on the idea of recovering a colour composition schema for HS imagery display (Du et al., 2008) and analysing colour data in place of HS pixels (Liang et al., 2013). Of course the consequence of this approach is that the abundant HS information is given away for the saliency analysis.

A recent research trend has started estimating the saliency in HS images without resorting to their colour rendering. Following this research direction, several studies concentrate on mitigating the curse of dimensionality by applying data reduction techniques to HS data, in order to extract non-redundant informative features that aid in highlighting discriminative properties of salient regions (Le Moan et al., 2013; Falini et al., 2020; Appice et al., 2020). A few studies perform the saliency analysis of HS imagery data according to statistical properties of HS pixels processed through an anomaly detector (Du & Zhang, 2014; Du et al., 2016). In alternative, contrast techniques coupled to clustering are applied to gradient spectral data (Yan et al., 2016; Zhang et al., 2018). All these studies estimate saliency in the HS imagery data by elaborating the abundance of information enclosed in HS pixels. Spatial information is often coupled to spectral data to deal with spectral variability. In any

case, these studies turn down the great progress made in saliency detection techniques for colour images. In addition, they analyse the imagery pixels of a single HS image.

In this paper, we revamp the idea of leveraging powerful methodologies fine-tuned in the literature to estimate saliency in colour data. However, we complement initial colour-based saliency assignments with spectral-based saliency refinements that allow us to better separate a salient landscape from its un-salient surroundings. In addition, we handle the HS imagery saliency detection task in the common scenario of surveillance missions, where multiple HS images of various scenes are collected using the same HS sensor technology. In this scenario, the task of saliency detection is performed on a dataset of HS images. As reported in Imamoglu et al. (2018), any saliency detection methodology designed for a single HS image can, in principle, be used for saliency detection in HS image datasets. In fact, independent HS saliency detection processes can be run in parallel on separate HS images to detect saliency in each HS image independently of other HS images. Differently, in this paper, we intend to take advantage of the collaboration among multiple saliency detection patterns learned from multiple HS images so that every pattern can possibly refine saliency assignments of other patterns by gaining accuracy.

To this aim, we propose an HS saliency detection methodology, named AGNES (pseudo-Label Generation for uNsupEvised Saliency detection in Hyperspectral image datasets), that takes a dataset of HS images acquired with the same HS sensor as input and yields the binary saliency matrices of the input images as output. In particular, the proposed methodology takes advantage of saliency classification patterns that can be learned by jointly elaborating the multiple HS images of the input dataset. The learning process is conducted on both the colour mode and the HS mode of each input image. First it applies a pre-trained, dataset-independent, saliency detection pattern to elaborate colour data and produce saliency pseudo-labels for all pixels of each input HS image. Then it leverages the produced colour-based pseudo-labels to supervise learning of a distinct HS saliency classification pattern from each input HS image. The PCA is used to deal with the curse of dimensionality during HS classification learning. Each classification pattern learned in this step is able to predict the binary saliency based on the HS pixel spectrum. The use of supervision allows us to learn an ensemble of multiple HS saliency classification patterns from multiple images. This ensemble is used to construct the final saliency matrices of the input images (or any other new testing images acquired with the same HS sensor).

To the best of our knowledge, the novelty of this study consists in the specific use adopted for colour-based saliency detection patterns within an HS ensemble learning methodology and the effectiveness of the combination that actually outperforms several state-of-the-art competitors on a benchmark HS image dataset. In particular, colour-based saliency detection patterns are used to fuel supervision in HS imagery analysis without requiring the acquisition of ground truth labels. So, this study contributes to proving that the proposed formulation is an effective means to delineate salient pixels in an unsupervised manner by taking actually advantage of the information enclosed in abundant spectrum of HS pixels. Another contribution is the use of the ensemble in the final labelling step. In this paper, we show that the ensemble can limit the effects of the lack of generality due to the spectral variability that may occur during the sensing operations of each single image. In fact, this lack of generality commonly affects saliency detection patterns learned processing HS data acquired with a single acquisition. Definitively, the empirical validation proves that the methodology gains in accuracy with the ensemble compared to other HS saliency detection algorithms.

The remainder of the paper is organised as follows. The next Section reports a brief overview of the recent literature. Section 3 introduces basic concepts, while Section 4 illustrates the proposed methodology. Section 5 provides the details of the experiments, the results and some discussions about them. In particular, the experiments described show the effectiveness of each component of the proposed methodology and compare the performance to that of various recent competitors. Finally, Section 6 summarises the conclusions.

## 2 Related work

Extensive literature studies have dealt with the saliency detection in traditional colour images. These methodologies can be roughly divided into two categories – heuristic saliency detection methodologies and deep learning-based saliency detection methodologies. The heuristic methodologies are mainly inspired by the human attention model in the neuroscience (Koch & Ullman, 1987) and have rooted in the Itti's model defined in Itti et al. (1998). The Itti's model mimics the human attention model through a centre-surround contrast technique that calculates the Euclidean distance of the considered pixel to its surrounding ones in the colour space. While simple, the Itti's model has inspired the most part of the subsequent literature (see Borji et al., 2019; Ullah et al., 2020 for recent surveys). On the other hand, deep neural networks have recently gained great success in a wide range of computer vision applications. So, a significant research effort is devoted to train deep neural network-based saliency detection patterns from big amounts of annotated colour images (Hou et al., 2019; Wang et al., 2020; Gao et al., 2019; Favorskaya & Jain, 2019; Liu et al., 2020; Luo et al., 2020). In addition to these salient detection methodologies that are defined for traditional imagery data, a few studies investigate the saliency topic in more challenging domains, such as video (Jun et al., 2015; Wang et al., 2020) or audio (Zlatintsi et al., 2015). In this paper we focus the overview mainly on the literature that explores the saliency detection topic in HS imagery data.

The first saliency detection methodology for HS imagery data is presented in Liang et al. (2013). It first converts an HS image into a colour image by projecting the HS space into the CIELAB colour space. Then it applies the Itti's model to the colour image.

The authors of Cao et al. (2015) seminaly formulate a notion of saliency as the extent to which a group of pixels stands out in an HS image in terms of reluctance instead of colourimetric contrast. They extract various HS conspicuity features to construct the unique saliency matrix and they apply the winner-take-all strategy to identify salient targets: pixels with the highest salience are regarded as salient targets. Features of various dimensionalities (reflectance curves, principal components and Gabor-filtered pixels) are also extracted in Le Moan et al. (2013). They are in turn compared using either the Euclidean distance or the angle distance.

A few studies propose anomaly detection methodologies for saliency detection in HS images. For example, the authors of Du and Zhang (2014) describe an anomaly detection methodology that resorts to the manifold feature to divide HS pixels into a potential anomaly (salient) part and a potential background part. In Du et al. (2016), statistical characteristics of HS pixels are processed in combination with a sparse representation. In particular, this study illustrates an anomaly detection approach that sparsely and linearly represents a pixel with different atoms under different hypotheses and assumes that the noise has the same

covariance structure, but different variances under the two competing hypotheses. It uses the generalised likelihood ratio test to construct the anomaly detector.

In Yan et al. (2016), the authors introduce the idea of resorting to a region-based spectral gradient contrast techniques to analyse HS imagery data. They first compute the gradient along each HS pixel. Then they apply segmentation and clustering techniques to gradient data to get a group of imagery regions. Finally, they resort to centre prior and local contrast to compute the saliency score of each region, with which the salient area can be obtained. A spectral gradient technique is also described in Zhang et al. (2018), where the authors propose to yield various saliency matrices through constructing an imagery region-based hierarchical structure. Each region is evaluated in multiple scales with a saliency pattern that depicts the region contrast with the spectral gradient.

Finally, very recent studies have applied sophisticated data dimensionality reduction techniques to saliency detection in HS images. In Falini et al. (2020), the authors propose an HS saliency detection methodology that cascades non-negative matrix factorisation and clustering on spectral distances computed between the input HS image and the reconstructed image. In Falini et al. (2020), the authors extend the approach described in Falini et al. (2020) by introducing new distance measures. In Appice et al. (2020), HS imagery data are reconstructed through a deep autoencoder neural network. Similarly to Falini et al. (2020), various distance measures are used to quantify the saliency degree in the data encoded and decoded through the autoencoder, while a clustering stage is performed in order to separate the salient information from the background. However, experiments reported in Appice et al. (2020) show that this methodology commonly achieves the best performance when autoencoders and clustering are coupled to the computation of the spectral spatial distance introduced in Yang and Mueller (2007).

In general, the HS saliency detection methods that have been defined in the recent literature have focused the attention on processing spectral data without taking advantage of supervision mechanisms. To the best of our knowledge, supervised algorithms have been widely investigated for saliency detection in colour images, where a big amount of annotated colour-based data is actually available. Instead, the lack of a significant amount of ground-truth labels for HS data has prevented recent saliency detection algorithms from taking advantage of supervision mechanisms with HS data. Therefore, introducing the use of supervision is a progress in the state-of-the-art of HS saliency detection literature. In any case, we differ from traditional supervised algorithms as we replace ground-truth labels, that remain unavailable during the learning stage, with saliency pseudo-labels. These pseudo-labels are predicted by applying accurate, pre-trained, colour-based saliency detection patterns to colour-based transformations of HS images. Another novel contribution is that existing HS saliency detection algorithms, in general, apply machine learning and numerical algorithms that learn the saliency matrix of a specific HS image without yielding a general pattern that can be also applied to a different HS image. In particular, they account for the spectral information acquired with the HS image under consideration only. So, if applied to a dataset of HS images, these algorithms process every HS image independently of each other. We note that this leads to overlook knowledge enclosed in saliency patterns potentially discovered from multiple HS images. A different approach is investigated in this paper. In particular, we resort to the ensemble learning theory (Brown, 2010) and explore how multiple saliency detection patterns learned from a dataset of multiple HS images may be handled as a “committee” of decision makers to achieve better overall accuracy on the input multiple HS images than the individual committee members.

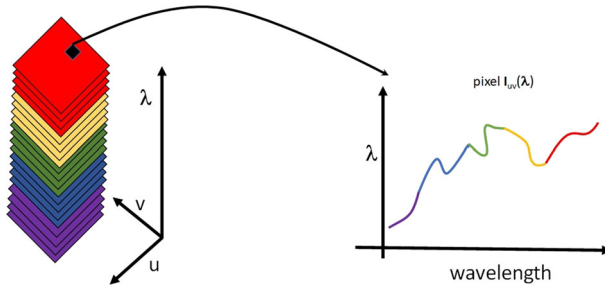


Fig. 1 HS imagery data

### 3 Preliminary concepts

Let  $\mathcal{I}$  be a dataset of HS images – digital images of observed scenes which are acquired using an HS sensor. The HS sensor records reflected light in hundreds of narrow frequencies covering the visible, near-infrared and shortwave infrared bands of a wavelength  $\lambda$  (also called spectrum). The spectrum is an  $m$ -dimensional feature vector (spectral vector), so that  $\lambda$  is spanned on the numeric spectral features  $\lambda_1, \lambda_2, \dots$ , and  $\lambda_m$ .

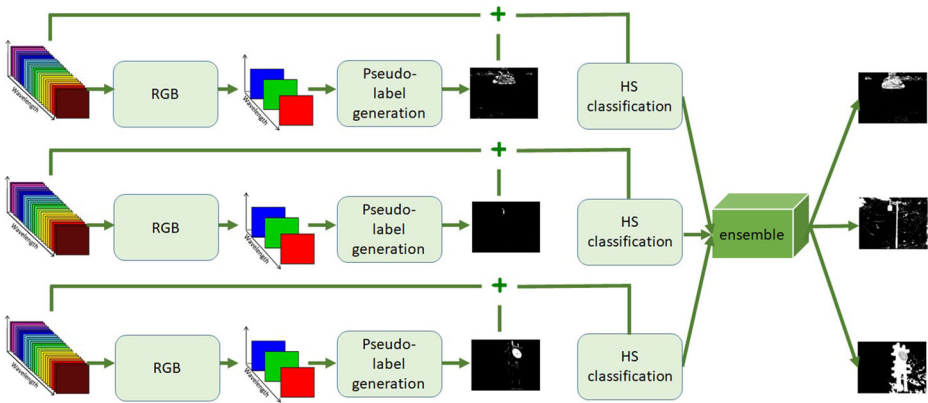
Every HS image  $\mathbf{I}_\lambda \in \mathcal{I}$  (see Fig. 1) is a three-dimensional set of pixels (called hypercube) with values representing spectral reflectance indexed by spatial coordinates  $u, v$  and spectrum  $\lambda$ . A pixel  $\mathbf{I}_\lambda(u, v)$  is a region of around a few square meters of the Earth’s surface that is a function of the sensor spatial resolution. Specifically, it is a one-dimensional spectrum section of hyper-cube  $\mathbf{I}_\lambda$  indexed by spatial coordinates  $(u, v)$  within the sensor resolution of the camera. Every pixel spectral value  $\mathbf{I}_\lambda(u, v, \lambda_i)$  is numeric and expresses how much the radiation is reflected, on average, at the  $i$ -th band of  $\lambda$  from the resolution cell of the considered pixel  $\mathbf{I}_\lambda(u, v)$ .

In a task of saliency detection, every HS imagery pixel can, in principle, be labelled according to an unknown binary target function, whose range is a finite set of two distinct labels, i.e. “salient” and “no-salient”. According to this function, a saliency matrix  $\mathbf{S}$  can be associated to an HS image  $\mathbf{I}_\lambda$ . In particular,  $\mathbf{S}$  is a two-dimensional set of saliency values with every value  $\mathbf{S}(u, v)$  representing the saliency label of the HS pixel  $\mathbf{I}_\lambda(u, v)$  indexed by the spatial coordinates  $u, v$ .

### 4 Learning algorithm

We propose a machine learning methodology named AGNES (pseudo-IABel Generation for uNsupEvised Saliency detection in Hyperspectral image datasets) for saliency detection in an HS image dataset. It takes as input a dataset of HS images (all images are sensed on the same spectrum) and learns an ensemble of HS saliency classification patterns by iterating colour-based, saliency, pseudo-label generation and HS classification analysis on every HS image of the dataset. The ensemble is used to build the final HS-driven saliency matrices of the HS image dataset taken as input.

We point out that, in the proposed methodology, both colour data and HS data analyses are coupled, since the learning process is performed jointly on these two different representations of the same imagery data. In particular, the saliency pseudo-labels, which are



**Fig. 2** The block diagram of AGNES methodology. It takes as input a dataset of HS images. It iterates colour-based, saliency, pseudo-label generation and HS classification analysis to populate an ensemble. It uses the ensemble to yield the saliency matrices of the processed HS images

produced from the colour display of HS imagery data, boost the supervision in the HS classification learning stage. In this way, we are actually able to take advantage of the progress achieved in colour-based saliency detection to yield good saliency pseudo-labels. These pseudo-labels allow us to perform classification analyses of HS imagery data although ground truth saliency labels are unavailable for supervision – HS imagery data are collected in an unsupervised scenario. Finally, the ensemble strategy (Opitz & Maclin, 1999) allows us to strengthen the accuracy of single HS saliency classification patterns that are learned from a dataset of multiple HS images as members of an ensemble. In particular, the ensemble strategy allows us to actually leverage a finite set of different HS saliency classification patterns for saliency detection. The better flexibility of the ensemble structure turns out to be more robust to spectral variability.

The block diagram of AGNES is illustrated in Fig. 2, while the pseudo-code is reported in Algorithm 1. The main used symbols are introduced in Table 1. The learning stages of the methodology (colour-based, saliency pseudo-label generation, HS classification learning and ensemble classification) are described in Sections 4.1–4.3. The analysis of the time complexity of the algorithm is reported in Section 4.4. Finally, the implementation details of the algorithm are reported in sub-Section 4.5.

**Table 1** Main used symbols

Symbol	Meaning
$\mathcal{I}$	HS image dataset (input)
$\mathcal{S}$	Saliency matrix dataset (output)
$\lambda$	Spectrum (spectral space)
$\mathbf{I}_\lambda$	HS image
$\mathbf{I}_{RGB}$	Colour rendering of $\mathbf{I}_\lambda$
$\mathbf{PC}_\lambda$	Principal component space of $\lambda$
$\phi_\lambda$	HS classification function
$\Phi$	Ensemble

**Algorithm 1** AGNES: pseudo-Ilabel Generation for uNsupEvised Saliency detection in Hyperspectral image datasets.

---

**Input:**  $\mathcal{I}$ : a dataset of HS images of scenes sensed with an HS sensor on a wavelength  $\lambda$

**Output:**  $\mathcal{S}$ : a data set of binary saliency matrices produced from HS images of  $\mathcal{I}$

```

1  $\Phi = \emptyset$ 
2  $\mathcal{S} = \emptyset$ 
3 for  $\mathbf{I}_\lambda \in \mathcal{I}$  do
    // colour-based, saliency pseudo-label generation
4    $\mathbf{I}_{\text{RGB}} = \text{rgb}(\mathbf{I}_\lambda)$ 
5    $\mathbf{S}_{\text{RGB}} \leftarrow \text{ASnet}(\mathbf{I}_{\text{RGB}})$ 
    // HS classification learning
6    $\mathbf{PC}_\lambda \leftarrow \text{pca}(\mathbf{I}_\lambda)$ 
7    $\mathbf{IPC}_\lambda \leftarrow \text{fit}(\mathbf{PC}_\lambda, \mathbf{I}_\lambda)$ 
8    $\phi_\lambda \leftarrow \text{classificationLearning}(\mathbf{IPC}_\lambda \oplus \mathbf{S}_{\text{RGB}})$ 
9    $\Phi = \Phi \cup \{(\mathbf{PC}_\lambda, \phi_\lambda)\}$ 
10 for  $\mathbf{I}_\lambda \in \mathcal{I}$  do
    // ensemble-based saliency detection
11    $\mathbf{I}_{\text{RGB}} = \text{rgb}(\mathbf{I}_\lambda)$ 
12    $\Phi \leftarrow \text{sort}(\Phi, \mathbf{I}_{\text{RGB}})$ 
13    $\langle \mathbf{PC}_\lambda^\uparrow, \phi_\lambda^\uparrow \rangle \leftarrow \text{topranked}(\Phi, \mathbf{S}_{\text{RGB}})$ 
14    $\mathbf{IPC}_\lambda^\uparrow \leftarrow \text{fit}(\mathbf{PC}_\lambda^\uparrow, \mathbf{I}_\lambda)$ 
15    $\mathbf{S}_{\phi_\lambda^\uparrow} \leftarrow \text{apply}(\phi_\lambda^\uparrow, \mathbf{IPC}_\lambda^\uparrow)$ 
16    $\mathcal{S} = \mathcal{S} \cup \{\mathbf{S}_{\phi_\lambda^\uparrow}\}$ 

```

---

#### 4.1 Colour-based saliency pseudo-label generation

In the first stage, we leverage the information enclosed in a colorimetric rendering of HS imagery data, in order to generate colour-based, saliency pseudo-labels (lines 4-5, Algorithm 1). These pseudo-labels will be used to supervise the subsequent HS classification learning stage. The pseudo-label generation procedure is repeated on every HS image of  $\mathcal{I}$  and proceeds as follows.

Let us consider an HS image  $\mathbf{I}_\lambda \in \mathcal{I}$ . First we determine  $\mathbf{I}_{\text{RGB}}$  that is the colorimetric rendering of  $\mathbf{I}_\lambda$ . For this operation, we apply the technique described in Foster and Amano (2019). First the pixel spectrum is scaled to range from 0 to 1 and multiplied by the global illuminant value,<sup>1</sup> converted to CIE XYZ tristimulus values according to equations:

$$X(u, v) = \kappa \int \mathbf{I}_\lambda(u, v) \bar{x}(\lambda) d\lambda, \quad (1)$$

$$Y(u, v) = \kappa \int \mathbf{I}_\lambda(u, v) \bar{y}(\lambda) d\lambda, \quad (2)$$

$$Z(u, v) = \kappa \int \mathbf{I}_\lambda(u, v) \bar{z}(\lambda) d\lambda, \quad (3)$$

<sup>1</sup>The illuminant is calculated from the imaging day light according to the calculator available at <https://www.waveformlighting.com/tech/calculate-cie-1931-xy-coordinates-from-cct/>. If the imaging light-day is unknown, then the default illuminant at 10000K can be adopted.



where  $\kappa$  is chosen so that  $Y=100$  for a perfectly white surface under full illumination, while  $\bar{x}(\lambda)$ ,  $\bar{y}(\lambda)$  and  $\bar{z}(\lambda)$  are the CIE XYZ colour-matching functions for the second standard observer (Foster & Amano, 2019). Then, CIE XYZ is scaled to representation to range from 0 to 1 and transformed to the default RGB color space sRGB according to the linear transformation (IEC, 1998):

$$\begin{bmatrix} R(u, v) \\ G(u, v) \\ B(u, v) \end{bmatrix} = \begin{bmatrix} 3.2406 & -1.5372 & -0.4986 \\ -0.9689 & 1.8758 & 0.0415 \\ 0.0557 & -0.2040 & 1.0570 \end{bmatrix} \begin{bmatrix} X(u, v) \\ Y(u, v) \\ Z(u, v) \end{bmatrix}. \quad (4)$$

According to guidelines reported in Foster and Amano (2019), sRGB values less than 0 are set to 0 and sRGB values greater than 1 are set to 1, in order to satisfy range constraints. Finally, a nonlinear correction is applied to compensate approximately for the input–output function of the display device. The typical approximate correction (Foster & Amano, 2019) has the form:

$$\begin{bmatrix} R'(u, v) \\ G'(u, v) \\ B'(u, v) \end{bmatrix} = \begin{bmatrix} R(u, v)^{0.4} \\ G(u, v)^{0.4} \\ B(u, v)^{0.4} \end{bmatrix}. \quad (5)$$

Then we use the ASNet pattern – a deep learning neural network pattern described in Wang et al. (2020) – to yield accurate pixel-wise saliency estimation of  $\mathbf{I}_{\mathbf{RGB}}$ . Specifically, the ASNet pattern produces a grey-level image that displays the numeric intensity of saliency estimated at every colour pixel of  $\mathbf{I}_{\mathbf{RGB}}$ . In this study, we use the Otsu’s algorithm (Otsu, 1979) to determine the upper threshold for separating the grey-level saliency values estimated by ASNet into two classes: foreground (“salient”) and background (“no-salient”). In particular, we assign imagery pixels with grey-level saliency intensity higher than the Otsu’s threshold to the pseudo-label “salient”, while we assign the remaining pixels to the pseudo-label “no-salient”. In this way, we produce the colour-based saliency matrix  $\mathbf{S}_{\mathbf{RGB}}$  associated to  $\mathbf{I}_\lambda$ , where every value  $\mathbf{S}_{\mathbf{RGB}}(u, v)$  represents the saliency pseudo-label that is colour-based assigned to the HS pixel  $\mathbf{I}_\lambda(u, v)$ .

The ASNet pattern is the colorimetric saliency estimation pattern that is described in Wang et al. (2020). It was learned by training an Attentive Saliency Network – a deep neural network architecture that was trained with a fixation map, derived at the upper network layers, which mimics human visual attention mechanisms and captures a high-level understanding of the scene from a global view. In particular, the ASNet architecture views saliency as fine-grained object-level saliency segmentation that is progressively optimized with the guidance of the fixation map in a top-down manner. It is a hierarchy of convLSTMs that offers an efficient recurrent mechanism to sequentially refine the saliency features over multiple steps. The pattern described in Wang et al. (2020) was trained on 30160 colour images collected in problems of human fixation prediction and salient object detection.<sup>2</sup> In this paper, the choice of this specific ASNet pattern is due to the extensive empirical validation illustrated in Wang et al. (2020), which proves that the same ASNet pattern that we use here can yield accurate pixel-wise grey-level saliency estimation in a high variety of colour images by outperforming 15 recent deep learning-based alternatives and 4 classical non-deep learning models.

The Otsu’s algorithm is an adaptive threshold algorithm introduced in Otsu (1979) commonly used in image binarization problems. It determines a saliency threshold in a grey-level image by minimising the intra-class intensity variance defined as a weighted sum of variances of the two classes.

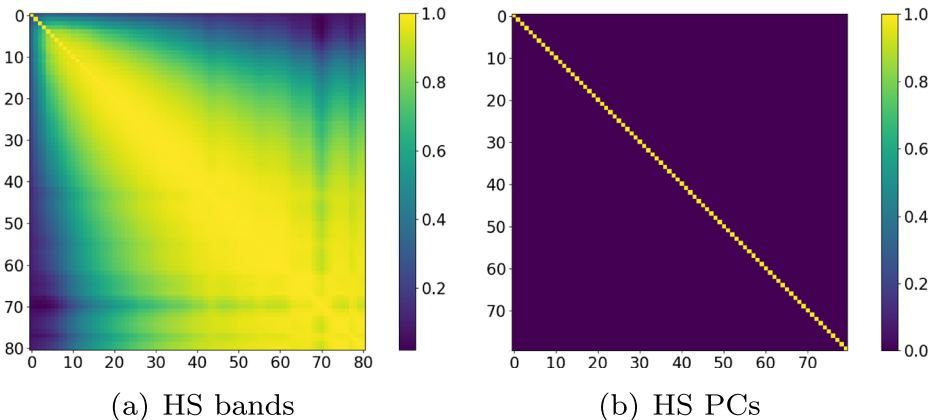
<sup>2</sup><https://github.com/wenguanwang/ASNet>

## 4.2 HS classification learning

In the second stage, we leverage the information enclosed in the spectral vectors of HS imagery data performing a process of HS classification learning. The purpose is to take advantage of the abundance of spectral information enclosed in HS pixels, in order to refine the colour-based saliency assignments that may have the highest difficulty in better separating the salient landscape from the surrounding landscape. The HS classification learning stage is repeated on every HS image of  $\mathcal{I}$  (lines 6–9, Algorithm 1) learning a new HS saliency classification pattern for feeding the ensemble  $\Phi$ . The classification analysis of every HS image proceeds as follows.

Let us consider an HS image  $\mathbf{I}_\lambda \in \mathcal{I}$ . First, we build  $\mathbf{I}_{PC_\lambda}$  – the hyper-cube produced by approximating the HS pixel of  $\mathbf{I}_\lambda \in \mathcal{I}$  on the independent feature space  $\mathbf{PC}_\lambda$  spanned by  $PC_1, PC_2, \dots$ , and  $PC_H$ . This feature space represents the top-ranked  $H$  principal components of  $\lambda$  in  $\mathbf{I}_\lambda$ . Then we couple the colour-based, saliency pseudo-labels of  $\mathbf{SRGB}$  to  $\mathbf{I}_{PC_\lambda}$ , in order to construct the training set  $\mathbf{I}_{PC_\lambda} \oplus \mathbf{SRGB}$ . Finally, we train an HS classification function  $\phi_\lambda: \mathbf{PC}_\lambda \mapsto \{\text{salient}, \text{un-salient}\}$  from  $\mathbf{I}_{PC_\lambda} \oplus \mathbf{SRGB}$  and add the learned classification function  $\phi_\lambda$  coupled to the principal component space  $\mathbf{PC}_\lambda$  to the ensemble  $\Phi$ . We point out that  $\langle \mathbf{PC}_\lambda, \phi_\lambda \rangle$  defines an HS saliency classification pattern that can be used to predict the saliency label of any HS pixel sensed on spectrum  $\lambda$ .

Principal Component Analysis (PCA) is one of the most widely used linear feature extraction techniques, which has been proved to be a powerful HS imagery data reduction strategy in tasks of HS classification (Xia et al., 2018; Appice & Malerba, 2019) or HS change detection (Lopez-Fandino et al., 2018; Appice et al., 2020). Specifically, PCA seeks to reduce the dimension of the data and drops the curse of dimensionality by finding a few orthogonal directions (the Principal Components – PCs) that express the original spectral bands whose projections on the PCs have the largest variance. In HS image analysis, the preference for the use of PCA for data reduction is also motivated by its ability to derive a collinearity-free characterisation of the spectrum. The spectral bands are strongly contemporaneously correlated with each other in the near spectrum, while the spectral principal components are contemporaneously uncorrelated with each other. An illustration of



**Fig. 3** IMG02 of HS-SOD dataset (see Section 5.1 for further details): heatmap of the Pearson correlation matrix of spectral bands (Fig. 3a) and heatmap of Pearson correlation matrix of the Principal Components, expressing an orthogonal reduced representation of the entire spectrum (Fig. 3b)

this phenomenon can be seen in Fig. 3a and b. We note that the collinearity phenomenon among near spectral bands may not be simply neglected, as it leads to a series of problems, such as unreliable coefficients and predictions, as well as aggravated data redundancy and computational complexity (Howley et al., 2006). In general, as discussed in Pravišević et al. (2017) and Pravišević et al. (2018), PCA is a mandatory step in improving the learning performance, by removing collinearity, speeding up the learning process and reducing the data storage requirements.

There are also valid alternatives to PCA. For example, autoencoders, that belong to the neural network family, are similar to PCA as they can be used for finding a low-dimensional representation of input data (Charte et al., 2018). They minimize the same objective function as PCA, but they are more flexible than PCA, due to the activation function that can introduce non-linearities in the encoding. Although autoencoders are really a big class of potentially extremely complex models, the advantage of PCA is that it is simple and efficient to train in comparison. Assuming that the linear transformation of PCA fits the spectral data accurately, it is much better to train PCA than try to select some complex deep model. In Appice and Malerba (2019), the viability of both PCA and autoencoders is compared in various benchmark HS imaging scenarios concluding that no significant improvement can be actually achieved in HS imagery analysis by considering auto-encoding instead of principal components.

For the classification analysis, we select XGBoost (Chen & Guestrin, 2016) as a valuable classification algorithm for this stage. Although, deep neural networks tend to outperform all other classification algorithms in various applications (included image analysis), traditional algorithms are still considered best-in-class when classification analysis comes to small-to-medium tabular data. In this paper, the choice of XGboost is due to the fact that it is currently one of the most popular machine learning algorithms (outside deep learning) in both academy and industry. It is a highly flexible and versatile algorithm that learns a decision-tree-based ensemble. It uses a gradient boosting framework to minimise the error of sequential models. In particular, XGBoost is efficient, as the process of sequential tree building is performed using parallelized implementation. It is designed to make efficient use of hardware resources and it is implemented with the depth-first approach that contributes to improve computational performance significantly. XGBoost is able to penalise more complex models through, backward tree pruning, LASSO (L1) and Ridge (L2) regularization to prevent overfitting. In addition, it employs the distributed weighted Quantile Sketch algorithm to effectively find the optimal split points among weighted datasets. Finally, there are several studies (Loggenberg et al., 2018; Samat et al., 2020; Zhou et al., 2020), where XGBoost has been applied to HS imaging with great success.

### 4.3 Ensemble classification

In the third stage, we use  $\Phi$  – the ensemble populated with the HS saliency classification patterns learned from  $\mathcal{I}$  – to yield  $\mathcal{S}$  – the set of the final saliency matrices produced for the HS images of  $\mathcal{I}$  (lines 10-16, Algorithm 1).

In particular, let us consider an HS image  $\mathbf{I}_\lambda \in \mathcal{I}$ , we use  $\Phi$  to build the saliency matrix  $\mathbf{S}_{\phi_\lambda}^\uparrow$  associated with  $\mathbf{I}_\lambda$  and add  $\mathbf{S}_{\phi_\lambda}^\uparrow$  to  $\mathcal{S}$ . The computation of  $\mathbf{S}_{\phi_\lambda}^\uparrow$  proceeds as follows. First we sort  $\Phi$  by an estimate of the accuracy of every pattern  $\langle \mathbf{PC}_\gamma, \phi_\gamma \rangle \in \Phi$  on  $\mathbf{I}_\lambda$ . To measure the accuracy of  $\langle \mathbf{PC}_\gamma, \phi_\gamma \rangle$  on  $\mathbf{I}_\lambda$ , we compare:

- $\mathbf{S}_{\phi_\gamma}^\lambda$  – the HS-based saliency matrix populated with saliency predictions yielded by  $\langle \mathbf{PC}_\gamma, \phi_\gamma \rangle$  on  $\mathbf{I}_\lambda$ , and

- $\mathbf{S}_{\text{RGB}}$  – the colour-based saliency matrix of  $\mathbf{I}_\lambda$  as it has been already computed in the first stage of the algorithm.

In particular, we measure:

$$\text{accuracy}(\langle \mathbf{PC}_\gamma, \phi_\gamma \rangle, \mathbf{I}_\lambda) = \text{AUCBorji}(\mathbf{S}_{\phi_\gamma}^\lambda, \mathbf{S}_{\text{RGB}}), \quad (6)$$

where  $\text{AUCBorji}()$  is the Borji variant of the Area Under the Roc Curve (AUC), which is commonly adopted in measuring accuracy in saliency detection tasks (Borji et al., 2013).

After sorting  $\Phi$  according to (6), we select the top-ranked HS saliency classification pattern  $\langle \mathbf{PC}_\lambda^\uparrow, \phi_\lambda^\uparrow \rangle \in \mathcal{I}$ , so that:

$$\langle \mathbf{PC}_\lambda^\uparrow, \phi_\lambda^\uparrow \rangle = \arg \max_{\langle \mathbf{PC}_\gamma, \phi_\gamma \rangle \in \Phi} \text{accuracy}(\langle \mathbf{PC}_\gamma, \phi_\gamma \rangle, \mathbf{I}_\lambda). \quad (7)$$

Finally, we use  $\langle \mathbf{PC}_\lambda^\uparrow, \phi_\lambda^\uparrow \rangle$  to build the final saliency matrix  $\mathbf{S}_{\phi_\lambda^\uparrow}$  associated with  $\mathbf{I}_\lambda$ . In  $\mathbf{S}_{\phi_\lambda^\uparrow}$ , the saliency label  $\mathbf{I}_\lambda(u, v)$  indexed by  $u, v$  is the prediction yielded by  $\phi_\lambda^\uparrow$  on the principal component values defined by fitting the principal component space  $\mathbf{PC}_\lambda^\uparrow$  to the HS imagery pixel  $\mathbf{I}_\lambda(u, v)$ .

#### 4.4 Time complexity

Let us consider that: 1) the dataset  $\mathcal{I}$  collects  $N$  images with  $U \times V$  pixel resolution acquired with an HS sensor covering a spectrum  $\lambda$  spanned on  $m$  spectral bands; 2) the used neural network ASNet was pre-trained estimating  $w$  weights; 3) the algorithm PCA is used to remove the collinearity in a system of  $m$  spectral bands; 4) the algorithm XGBoost is used to learn XGBoost patterns with  $K$  trees spanned over  $d$  layers; 5) the algorithm AUCBorji is implemented averaging the AUC values computed on  $s$  split trials with  $t$  equal-width steps processed in each split trial. Based upon these premises, the time complexity of AGNES is computed by summing-up the cost of: (i) generating the saliency pseudo-labels with ASNet, (ii) learning the HS saliency classification patterns with PCA and XGBoost and (iii) using the ensemble of the learned HS saliency classification patterns coupled with AUCBorji to produce the final saliency matrices of the input images. Note that the pipeline composed of the steps (i) and (ii) may be also run in parallel on each independent image of  $\mathcal{I}$ . Again, once the ensemble of the HS saliency classification patterns is fully learned by completing steps (i) and (ii) on all images of  $\mathcal{I}$ , the step (iii) may be run in parallel on each independent image of  $\mathcal{I}$ .

**Saliency pseudo-label generation** for each HS image  $\mathbf{I}_\lambda \in \mathcal{I}$ , the color image  $\mathbf{I}_\lambda$  is determined and the pseudo-labels  $\mathbf{S}_{\text{RGB}}$  are predicted. The time cost of building  $\mathbf{I}_{\text{RGB}}$  is proportional to  $mUV$ , while the time cost of generating  $\mathbf{S}_{\text{RGB}}$  is proportional to  $wUV$ . Considering that  $w \gg m$ , the time complexity of completing this step on all HS images of  $\mathcal{I}$  may be re-written as  $\mathbf{O}(N(wUV))$ .

**HS Classification learning** for each HS image  $\mathbf{I}_\lambda \in \mathcal{I}$ , its principal components are computed and then XGBoost pattern is trained. The time cost of determining a system of principal components for an HS image is  $\mathbf{O}(mUV \times \min(m, UV) + m^3)$ . The time cost of training a XGBoost pattern is  $Kdn_z + n_z \log r_b$  where  $K$  is the number of trees,  $d$  is the number of tree layers,  $n_z$  is the number of non-missing entries in the training dataset and  $r_b$  is the maximum number of rows in each block. Note that according to the theory reported

in Chen and Guestrin (2016), the block structure is adopted to speed-up the computation on large datasets. Therefore, assuming  $m \ll UV$ , the time complexity of completing this step on all HS images of  $\mathcal{I}$  is  $\mathbf{O}(N(m^2UV + Kdn_z + n_z \log r_b))$ .

**Ensemble classification** for each HS image  $\mathbf{I}_\lambda \in \mathcal{I}$ , the XGBoost pattern that achieves the highest *AUCBorji* predicting the pseudo-labels is selected and the final saliency matrix is built. The cost of measuring the *AUCBorji* value of an XGBoost pattern with respect to the saliency pseudo-labels of an HS image is proportional to  $stUV$ , where  $s$  denotes the number of split trials and  $t$  is the number of steps. This operation is repeated on each XGBoost pattern enclosed in the ensemble. The cost of selecting the top XGBoost pattern of the ensemble is  $\mathbf{O}(N)$ , while the cost of predicting the final labels with the selected XGBoost pattern is  $\mathbf{O}(Kd)$ . Therefore, the cost of performing this phase on all HS images of  $\mathcal{I}$  is  $\mathbf{O}(N(NstUV + N + Kd))$ . As  $N < NUUV$  and assuming that  $Kd \ll UV$ , the time complexity of this step can be rewritten as  $\mathbf{O}(N^2stUV)$ .

#### 4.5 Implementation details

AGNES is written in Python 3.7. It uses the ASNet pattern learned in Wang et al. (2020)<sup>3</sup> using Keras 2.3<sup>4</sup> – a high-level neural network API with TensorFlow<sup>5</sup> as the back-end. In addition, AGNES imports:

- The implementation of Otsu’s algorithm from `skimage.filters.threshold_otsu`.<sup>6</sup>
- The implementation of PCA from `sklearn.decomposition.PCA`.<sup>7</sup> In particular, we compute the PCA with the full SVD calling the standard LAPACK solver (`svd_solver='full'`) and with Minka’s MLE (Minka, 2001) for the automatic choice of dimension  $H$  (`component='mle'`). By adopting this choice, the algorithm called probabilistic principal component analysis (PPCA) described in Tipping and Bishop (2006) is used.
- The implementation of XGBoost from `xgboost.XGBoostClassifier`,<sup>8</sup> by adopting the default parameter set-up is reported in the documentation.<sup>9</sup> In particular, the number of trees  $K = 100$  and the number of tree layers  $d = 6$ .
- The implementation of the AUCBotji algorithms is done by following the guidelines reported in Borji et al. (2013). In particular, the AUC is measured on  $t = 10$  equal-width steps and repeated on  $s = 100$  split trials. The final AUC is computed as the average of AUC values computed on the split trials.

### 5 Experimental evaluation and discussion

To provide a compelling evaluation of the accuracy of our methodology, we have conducted a range of experiments on the benchmark HS saliency detection dataset called HS-SOD

<sup>3</sup><https://github.com/wenguanwang/ASNet>

<sup>4</sup><https://keras.io/>

<sup>5</sup><https://www.tensorflow.org/>

<sup>6</sup>[https://scikit-image.org/docs/dev/api/skimage.filters.html#skimage.filters.threshold\\_otsu](https://scikit-image.org/docs/dev/api/skimage.filters.html#skimage.filters.threshold_otsu)

<sup>7</sup><https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>

<sup>8</sup><https://xgboost.readthedocs.io/en/latest/index.html>

<sup>9</sup><https://xgboost.readthedocs.io/en/latest/parameter.html>

(Imamoglu et al., 2018). This dataset is available with the ground-truth saliency images. The main objective of these experiments is to evaluate the effectiveness of the proposed learning methodology along its various learning dimensions – learning on spectrum, classification with principal components, ensemble strategy (Section 5.2). In addition, we investigate how the proposed methodology compares to state-of-the-art HS saliency detection competitors (Section 5.3). In all these experiments, the accuracy performance is evaluated with the Borji variant of the Area Under the ROC Curve (AUCBorji) (Borji et al., 2013). This metric has been already used in Imamoglu et al. (2018), Appice et al. (2020), Falini et al. (2020), and Falini et al. (2020) for the analysis of the performance of various HS saliency detection methodologies defined in the literature and evaluated on the dataset HS-SOD.

## 5.1 HS-SOD

The dataset HS-SOD (Imamoglu et al., 2018)<sup>10</sup> is a benchmark HS image dataset that collects 60 HS images with  $1024 \times 768$  pixel resolution. These HS images were sensed in various scenes of the public parks of Tokyo Waterfront (City in Odaiba, Tokyo, Japan) in several days between August - September 2017 when the weather was sunny or partially cloudy. They exhibit different characteristics in terms of sensed landscapes, salient landscape size, foreground-background contrast and salient landscape position on the image. In particular, each HS image of HS-SOD was acquired using the NH-AIK hyper-spectral camera that covers a spectrum between  $350 \mu\text{m}$  and  $1100 \mu\text{m}$  spanned on 150 spectral bands with spectral resolution  $5 \mu\text{m}$ . However, the authors of HS-SOD prepared the images on the visible spectrum ( $380\text{--}780 \mu\text{m}$ ) spanned on 81 spectral bands. In addition, they produced a ground-truth binary image for every HS image, so that salient pixels were labelled within each ground-truth image. We point out that ground-truth labels have been ignored during the execution of the saliency detection methodologies, while they are considered to evaluate the accuracy of the saliency matrices constructed.

## 5.2 Learning components analysis

In this Section we analyse the effectiveness of the various learning components of AGNES, in order to answer the following questions:

1. How does the analysis of the spectrum information affects the accuracy of the saliency matrices computed?
2. How does the accuracy of the HS imaging analysis change by introducing PCA to deal with the curse of dimensionality?
3. Is the idea of using the ensemble strategy more powerful than considering single classification patterns?

To this purpose, we perform an ablation study where we consider four configurations as baselines. These are in turn defined by removing HS information analysis, PCA and/or ensemble learning from the whole methodology of AGNES. In particular, these baseline configurations are defined as follows:

1. ASNet that takes as input each HS image of the input dataset and applies the ASNet pattern to construct the imagery saliency matrix of the image from the colour rendering

<sup>10</sup><https://github.com/gistairc/HS-SOD>

**Table 2** Characteristics of the several configurations of AGNES evaluated

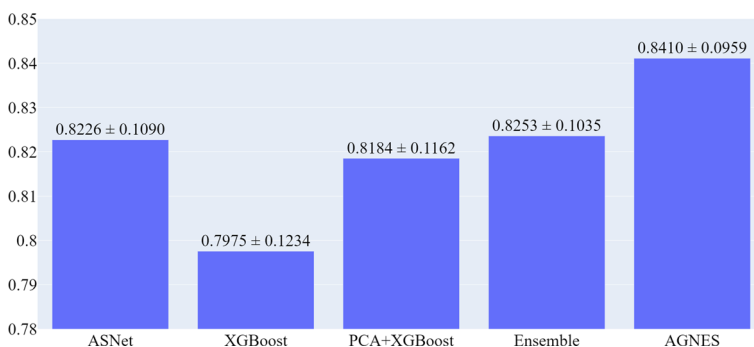
Configuration	Colour information	Spectral data analysis	PCA	Ensemble learning
ASNet	Yes	No	No	No
XGBoost	Yes	Yes	No	No
PCA+XGBoost	Yes	Yes	Yes	No
Ensemble	Yes	Yes	No	Yes
AGNES	Yes	Yes	Yes	Yes

of the HS imagery pixels. This baseline considers the colour pixel representation only by giving away both HS information analysis, PCA and ensemble learning.

- XGBoost that takes as input each HS image of the input dataset and applies the ASNet pattern to generate the saliency pseudo-labels from the colour rendering of the HS imagery pixels. It uses these colour-based, saliency pseudo-labels to supervise learning of HS classification patterns with XGBoost. It constructs the final imagery saliency matrix of each HS image with the saliency predictions produced by the XGBoost pattern learned on the HS imagery pixels of the considered image. This baseline couples the colour-based analysis to the spectral-based analysis giving away PCA and ensemble learning.
- PCA+XGBoost that is like the configuration of XGBoost, but it performs PCA of HS information before learning HS classification patterns with XGBoost. This baseline couples the colour-based analysis to PCA and spectral-based analysis giving away ensemble learning.
- Ensemble that populates an ensemble with multiple HS classification patterns that are learned with the configuration XGBoost from the multiple HS images of the input dataset. It uses this ensemble to build the final imagery saliency matrices of the HS images in the input dataset. This baseline couples the colour-based analysis to the spectral-based analysis and considers the ensemble learning giving away PCA during HS information analysis.

A summary of the characteristics of the compared configurations is reported in Table 2.

We evaluate the performance of ASNet, XGBoost, PCA+XGBoost, Ensemble and AGNES processing all the HS images of HS-SOD. In particular, Fig. 4 reports the AUCBorji

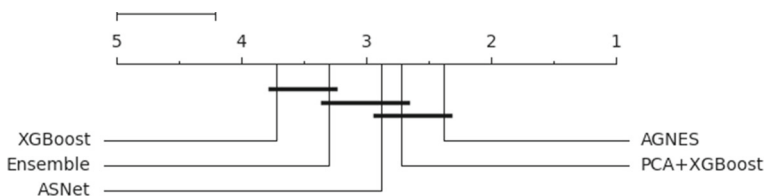


**Fig. 4** AUCBorji (mean ± stdev) of ASNet, XGBoost, PCA+XGBoost, Ensemble and AGNES computed on the saliency matrices yielded for all the HS images of HS-SOD

(mean and standard deviation) of the compared configurations confirming that AGNES outperforms, in average, all its baselines. These results point out that decoupling the spectral-based analysis from both PCA and ensemble learning (XGBoost) leads to a drop in accuracy with respect to the baseline decision of saliency assignments based on the colour-based information only (ASNet). On the other hand, the configuration that couples PCA to HS imagery classification learning without ensemble (PCA+XGBoost), as well as the configuration that uses an ensemble of HS classification patterns learned without PCA (Ensemble) have performance which stay close to that of ASNet. Our interpretation of these results is that PCA contributes to deal with the curse of dimensionality (as well as with HS information collinearity) in agreement with the conclusions already drawn in both (Appice & Malerba, 2019) and (Appice et al., 2020) for HS classification and HS change detection, respectively. Ensemble learning contributes to handle possible phenomena of spectral variability that may occur in single HS images by taking advantage of HS multiple classifiers in place of a single one (Ceamanos et al., 2009). So, the winning strategy that allows AGNES to properly take advantage of HS information dealing with both issues simultaneously is derived by combining achievements of both PCA and ensemble learning.

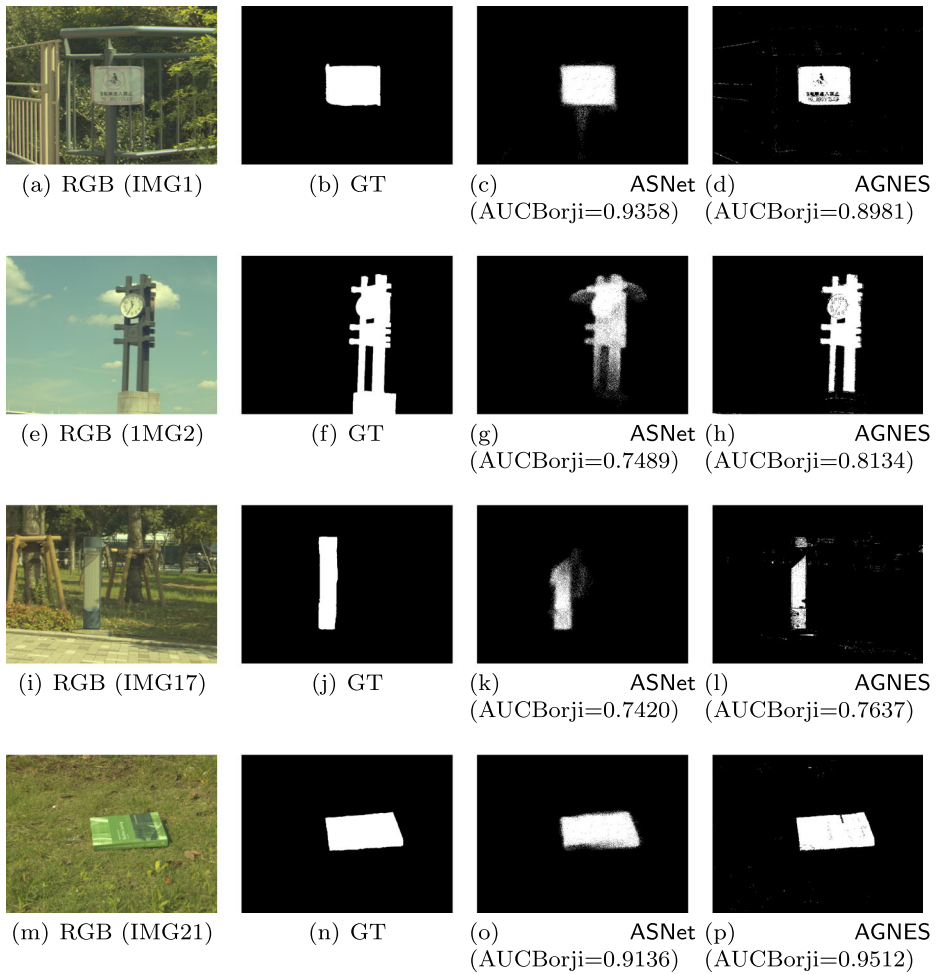
To statistically test whether the improvement in accuracy of AGNES is significant, we use the Friedman's test. This is a non-parametric test that is commonly used to compare multiple approaches over multiple datasets (Demšar, 2006). The Friedman test compares the average ranks of the approaches, so that the best performing approach gets the rank of 1, the second best gets rank 2. The null-hypothesis states that all the configurations are equivalent. Under this hypothesis, ranks of compared approaches should be equal. We perform this test on the AUCBorji scores of the compared configurations on each image of the dataset HS-SOD and reject the null hypothesis with  $p\text{-value} \leq 0.05$ . As the null-hypothesis is rejected, we use a post-hoc test—the Nemenyi test—for pairwise comparisons (Demšar, 2006). The display of the results of this test, reported in Fig. 5, confirms that AGNES is ranked higher than all its baseline configurations. In particular, the critical difference diagram, obtained using a 0.05 significance level, shows that AGNES is on average the best performing approach with the configuration PCA+XGBoost as runner-up.

At the completion of this discussion, we compare the visual display of a few saliency matrices produced by processing colour-based information, as well as joining colour-based and HS-based information. Figure 6 shows the visual displays of both the input and output of ASNet and AGNES for the HS images 1, 2, 17 and 21 of HS-SOD. These displays highlight that, in all the HS images, AGNES can better delineate salient regions along their boundaries. This is a confirmation of the fact that we propose a saliency detection methodology that can actually take advantage of the HS information to better separate landscapes. We also note that the AUCBorji score achieved by AGNES is lower than the AUCBorji score achieved by ASNet on image 1. This score is computed with the ground truth provided



**Fig. 5** Comparison between ASNet, XGBoost, PCA+XGBoost, Ensemble and AGNES with the Nemenyi test using the AUCBorji. Groups of configurations that are not significantly different (at  $p \leq 0.05$ ) are connected





**Fig. 6** IMG1, IMG2, IMG17, IMG21 of HS-SOD dataset: RGB rendering (Fig. 6a, e, i and m), saliency ground-truth (Fig. 6b, f, j and n), saliency images built by ASNet (Fig. 6c, g, k and o), saliency images built by AGNES (Fig. 6d, h, l and p)

by the authors of the dataset. In this image, AGNES recognises the sign as the salient part of the image, but it also separates the illustration within the sign from the white scene. This separation, that is non reported in ground truth, depends on the fact that the sign background and the illustration within the sign actually define different landscapes. This suggests that a possible direction to extend this methodology comprises a mechanism to derive a multi-level, global-to-local, representation of the saliency information within the same imagery scene (i.e., the local salient detail emerging within a global salient area).

### 5.3 Competitor analysis

Finally, we compare the AUC-Borji scores competitors considered in the recent literature for problems of saliency detection in HS images. For this comparative study, we consider:

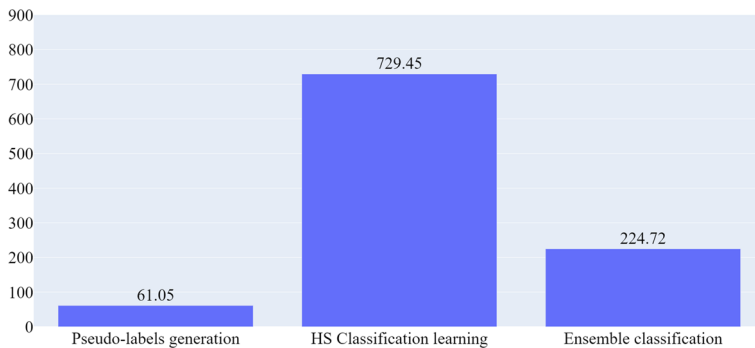
- AISA – the HS saliency detection methodology described in Appice et al. (2020). It uses an auto-encoder representation of HS imagery data, spectral-spatial difference between imagery data and reconstructed data, as well as clustering.
- SSMF and SSMF1 – the HS saliency detection methodologies presented in Falini et al. (2020) and (Falini et al., 2020) respectively. They use a sparse non-negative matrix factorization algorithm together with several error functions, based on spectral and spatial measures, and with Gaussian-based clustering techniques.
- Itti's model (Itti et al., 1998) computed on the colour rendering of the HS images.
- SED and SAM – the HS saliency detection methodologies described in Liang et al. (2013). It is used with spectral Euclidean distance and spectral angular distance, respectively.
- GS – the HS saliency detection methodology presented in Liang et al. (2013). It first divides spectral bands into four groups (G1,G2,G3,G4) and then measures the colour opponency by computing the Euclidean distance between these vectors (G1-G3 and G2-G4).
- SED-OCM-GS and SED-OCM-SAD – the HS saliency detection combinations proposed in Liang et al. (2013). They use orientation-based salient features (OCM).
- SGC – the HS saliency detection methodology illustrated in Yan et al. (2016). It determines super-pixels by computing spectral and spatial gradients and identifies local region contrasts from super-pixels.

Table 3 reports the AUCBorji scores of all the compared methodologies. AGNES achieves a better average accuracy compared to all the other competitors tested in this study. The best performance of AGNES is mainly due to the ensemble learning process, coupled with the specific cascade of colour-based analysis and HS-based analysis. We note that we take advantage of the ensemble learning as we count on a dataset of HS images. This condition, generally, happens in surveillance missions (e.g. environmental surveillance done with a drone), where multiple images of various scenes are, commonly, sensed using the same HS technology within the same mission.

To complete this study, we analyse the computation time spent completing the saliency detection task on HS-SOD dataset. The computation time is measured in minutes on an Intel(R) Core(TM) i7-4720U CPU@2.60 GHz and 16 GB RAM running Microsoft Windows 8.1 (64 bits). Figure 7 reports the total time spent completing the three steps of AGNES

**Table 3** Competitor analysis

	Saliency detection methodology	AUCBorji
	AGNES	0.8410
	AISA	0.7540
AUCBorji score of AGNES and its competitors computed on HS-SOD. The AUCBorji of AISA is reported in Appice et al. (2020). The AUCBorji of SSMF and SSMF1 are reported in Falini et al. (2020) and Falini et al. (2020). The AUCBorji of Itti's model, SED, SAM, GS, SED-OCM-GS, SED-OCM-SAD and SGC are reported in Imamoglu et al. (2018)	SSMF1	0.7592
	SSMF	0.7329
	Itti's model	0.7694
	SED	0.6415
	SAM	0.7521
	GS	0.7597
	SED-OCM-GS	0.7863
	SED-OCM-SAD	0.8008
	SGC	0.8295



**Fig. 7** Total computation time spent (in minutes) completing the three steps of AGNES on all the HS images of HS-SOD dataset

(i.e. colour-based saliency pseudo-label generation, HS classification learning and ensemble classification). Table 4 compares the computation times spent completing the saliency detection task with both AGNES and its competitors AISA, SSMF1 and SSMF.<sup>11</sup> These results confirm that the higher accuracy achieved by AGNES is at the cost of the more computation time spent both performing the supervision in the HS classification learning step and adopting the ensemble for producing the final saliency matrices. Note that, in this evaluation, we have not use a parallel computation infrastructure, in order to perform the learning steps of the compared algorithms on several HS images simultaneously. In principle, assuming the availability of a parallel computation infrastructure, AISA, SSMF1 and SSMF may process all the HS images of HS-SOD dataset in parallel spending, in average, 9.94 minutes, 6.75 minutes and 2.95 minutes, per image, respectively. On the other hand, AGNES spends, in average, 16.92 minutes per image. This estimate is derived by assuming that both the pseudo-label generation and the HS classification learning steps are completed in parallel on each HS image of HS-SOD dataset. Similarly, the ensemble classification of each HS image is also completed in parallel on each HS image of HS-SOD dataset once all the classification patterns of the ensemble have been learned.<sup>12</sup>

## 6 Conclusion

This paper illustrates a learning methodology for analyzing a dataset of HS images and detecting, in each HS image, the salient pixel region that can be considered more notable than the background.

The proposed methodology takes advantage of a learning process done on a dataset of HS images sensed through the same HS sensor. Learning is done by jointly processing the imagery pixel information represented on both the colour mode and the HS mode. In particular, the proposed methodology yields saliency pseudo-labels from a colourimetric

<sup>11</sup>The code of these algorithms is available to be run for this comparison.

<sup>12</sup>This approximate estimate of the computation time possibly spent completing the considered saliency detection task within a parallel computation infrastructure does not consider any estimate of the communication times. In any case, the investigation of the performance of a parallel strategy to implement the proposed algorithm is out of the scope of the current paper.

**Table 4** Competitor analysis

Saliency detection methodology	TOTAL TIME	AVG TIME
AGNES	1015.22	16.92
AISA	596.84	9.94
SSMF1	405.0	6.75
SSMF	177.0	2.95

Computation time spent in minutes with AGNES and its competitors AISA, SSMF1 and SSMF completing the saliency detection task on HS-SOD dataset. The TOTAL TIME (column 2) is the time spent processing all the HS images of HS-SOD dataset. The AVG TIME (column 3) is the time spent, in average, processing a single HS image of HS-SOD dataset

rendering of HS imagery data. It uses these pseudo-labels to supervise learning of HS classification patterns trained with XGboost. These HS classification pattern predicts saliency assignments based on the HS pixel spectrum. It uses PCA, in order to deal with the curse of dimensionality during HS classification learning, as well as ensemble learning, in order to strengthen the accuracy of multiple HS classification patterns trained for saliency detection on multiple HS images.

The experiments are performed on a benchmark HS image dataset that comprises both HS images and ground-truth saliency images. Ground-truth labels are considered to evaluate the accuracy of the saliency assignments learned. The performed experiments investigate the sensitivity of the performance to the steps of the learning methodology proving that every component of the methodology contributes to the gain in detection accuracy. The results also reveal that the proposed methodology is able to provide competitive accuracy compared to state-of-the-art HS models. In fact, with the encouraging performance of the proposed methodology, precise salient areas of various scenes may be identified.

Some directions for further work are still to be explored. Appropriate deep learning architectures can be considered, in order to improve the accuracy of the HS classification learning stage. Active learning mechanisms can be applied for label acquisition in a possible weakly supervised enhancement of the classification analysis. In addition, we plan to explore the possibility of introducing a mechanism for a global-to-local mechanism of saliency detection to recognize (and possibly classify) local components belonging to different landscapes within the same global salient region. In addition, we intend to investigate different HS feature engineering algorithms to feed the classification stage of the learning methodology proposed in this study. In particular, we plan to explore the performance of Gabor features (Jia et al., 2015), autocorrelation features (Appice & Malerba, 2019), morphological features (Appice et al., 2016; 2017) and frequency features (Guccione et al., 2015) in the investigated HS saliency detection scenario. Finally, we plan to extend the investigation of the feasibility of a parallel strategy for implementing the proposed algorithm.

**Acknowledgements** This work fulfills the research objectives of the PON Ricerca e Innovazione 2014-2020 project RPASInAir Integrazione dei Sistemi Aeromobili a Pilotaggio Remoto nello spazio aereo non segregato per servizi (ARS01 00820), funded by the Italian Ministry for Universities and Research (MIUR). The research of Antonella Falini is funded by PON Project AIM 1852414 CUP H95G18000120006 ATT1. The research of Cristiano Tamborrino is funded by PON Project Change Detection in Remote Sensing CUP H94F18000260006. The authors Antonella Falini, Cristiano Tamborrino and Francesca Mazzia are members of the INdAM Research group GNCS.

**Funding** Open access funding provided by Università degli Studi di Bari Aldo Moro within the CRUI-CARE Agreement.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.





## References

- Appice, A., Guccione, P., Acciaro, E., & Malerba, D. (2020). Detecting salient regions in a bi-temporal hyperspectral scene by iterating clustering and classification. *Applied Intelligence*, *50*(10), 3179–3200. <https://doi.org/10.1007/s10489-020-01701-8>.
- Appice, A., Guccione, P., & Malerba, D. (2016). Transductive hyperspectral image classification: toward integrating spectral and relational features via an iterative ensemble system. *Machine Learning*, *103*(3), 343–375. <https://doi.org/10.1007/s10994-016-5559-7>.
- Appice, A., Guccione, P., & Malerba, D. (2017). A novel spectral-spatial co-training algorithm for the transductive classification of hyperspectral imagery data. *Pattern Recognition*, *63*, 229–245. <https://doi.org/10.1016/j.patcog.2016.10.010>.
- Appice, A., Lomuscio, F., Falini, A., Tamborrino, C., Mazzia, F., & Malerba, D. (2020). Saliency detection in hyperspectral images using autoencoder-based data reconstruction. In D. Helic, G. Leitner, M. Stettinger, A. Felfernig, & Z. W. Ras (Eds.) *Foundations of intelligent systems - 25th international symposium, ISMIS 2020, Graz, Austria, September 23-25, 2020, Proceedings, Lecture Notes in Computer Science*, (Vol. 12117 pp. 161–170). Springer. [https://doi.org/10.1007/978-3-030-59491-6\\_15](https://doi.org/10.1007/978-3-030-59491-6_15).
- Appice, A., & Malerba, D. (2019). Segmentation-aided classification of hyperspectral data using spatial dependency of spectral bands. *ISPRS Journal of Photogrammetry and Remote Sensing*, *147*, 215–231. <https://doi.org/10.1016/j.isprsjprs.2018.11.023>.
- Bioucas-Dias, J. M., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N., & Chanussot, J. (2013). Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and Remote Sensing Magazine*, *1*(2), 6–36. <https://doi.org/10.1109/MGRS.2013.2244672>.
- Borji, A., Cheng, M. M., Hou, Q., Jiang, H., & Li, J. (2019). Salient object detection: a survey. *Computational Visual Media*, *5*(2), 117–150. <https://doi.org/10.1007/s41095-019-0149-9>.
- Borji, A., Tavakoli, H. R., Sihite, D. N., & Itti, L. (2013). Analysis of scores, datasets, and models in visual saliency prediction. In *2013 IEEE International conference on computer vision* (pp. 921–928). <https://doi.org/10.1109/ICCV.2013.118>.
- Brown, G. (2010). *Ensemble Learning*, (pp. 312–320). Boston: Springer. [https://doi.org/10.1007/978-0-387-30164-8\\_252](https://doi.org/10.1007/978-0-387-30164-8_252).
- Cao, Y., Zhang, J., Tian, Q., Zhuo, L., & Zhou, Q. (2015). Salient target detection in hyperspectral images using spectral saliency. In *2015 IEEE China Summit and international conference on signal and information processing (chinaSIP)* (pp. 1086–1090). <https://doi.org/10.1109/ChinaSIP.2015.7230572>.
- Ceamanos, X., Waske, B., Benediktsson, J. A., Chanussot, J., & Sveinsson, J.R. (2009). Ensemble strategies for classifying hyperspectral remote sensing data. In J. A. Benediktsson, J. Kittler, & F. Roli (Eds.) *Multiple classifier systems* (pp. 62–71). Berlin: Springer.
- Charte, D., Charte, F., García, S., del Jesus, M. J., & Herrera, F. (2018). A practical tutorial on autoencoders for nonlinear feature fusion: taxonomy, models, software and guidelines. *Information Fusion*, *44*, 78–96.
- Chen, T., & Guestrin, C. (2016). Xgboost: a scalable tree boosting system. In B. Krishnapuram, M. Shah, A. J. Smola, C. C. Aggarwal, D. Shen, & R. Rastogi (Eds.) *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, KDD 2016* (pp. 785–794). ACM. <https://doi.org/10.1145/2939672.2939785>.
- Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, *7*, 1–30.
- Du, B., & Zhang, L. (2014). A discriminative metric learning based anomaly detection method. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(11), 6844–6857. <https://doi.org/10.1109/TGRS.2014.2303895>.
- Du, B., Zhang, Y., Zhang, L., & Tao, D. (2016). Beyond the sparsity-based target detector: a hybrid sparsity and statistics-based detector for hyperspectral images. *IEEE Transactions on Image Processing*, *25*(11), 5345–5357. <https://doi.org/10.1109/TIP.2016.2601268>.

- Du, Q., Raksuntorn, N., Cai, S., & Moorhead, R. (2008). Color display for hyperspectral imagery. *Geoscience and Remote Sensing. IEEE Transactions on*, *46*, 1858–1866. <https://doi.org/10.1109/TGRS.2008.916203>.
- Falini, A., Castellano, G., Tamborrino, C., Mazzia, F., Mininni, R. M., Appice, A., & Malerba, D. (2020). Saliency detection for hyperspectral images via sparse-non negative-matrix-factorization and novel distance measures. In *2020 IEEE Conference on evolving and adaptive intelligent systems, EAIS 2020* (pp. 1–8). IEEE. <https://doi.org/10.1109/EAIS48028.2020.9122749>.
- Falini, A., Tamborrino, C., Castellano, G., Mazzia, F., Mininni, R. M., Appice, A., & Malerba, D. (2020). Novel reconstruction errors for saliency detection in hyperspectral images. In *6th International Conference on Machine Learning, Optimization, and Data Science, LOD 2020, Lecture Notes in Computer Science*, Vol. 12565-12566. Springer. [https://doi.org/10.1007/978-3-030-64583-0\\_12](https://doi.org/10.1007/978-3-030-64583-0_12).
- Favorskaya, M., & Jain, L. (2019). Saliency detection in deep learning era: trends of development. *Information and Control Systems* 10–36. <https://doi.org/10.31799/1684-8853-2019-3-10-36>.
- Foster, D., & Amano, K. (2019). Hyperspectral imaging in color vision research: tutorial. *Optical Society of America Journal A Optics Image Science, and Vision*, *36*(4), 606. <https://doi.org/10.1364/JOSAA.36.000606>.
- Fu, W., Ma, J., Chen, P., & Chen, F. (2020). Remote sensing satellites for digital earth. In H. Guo, M. F. Goodchild, & A. Annoni (Eds.) *Manual of Digital Earth* (pp. 55–123). Singapore: Springer. [https://doi.org/10.1007/978-981-32-9915-3\\_3](https://doi.org/10.1007/978-981-32-9915-3_3).
- Gao, S., Cheng, M., Zhao, K., Zhang, X., Yang, M., & Torr, P.H.S. (2019). Res2net: a new multi-scale backbone architecture. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1–1. <https://doi.org/10.1109/TPAMI.2019.2938758>.
- Guccione, P., Mascolo, L., & Appice, A. (2015). Iterative hyperspectral image classification using spectral-spatial relational features. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(7), 3615–3627. <https://doi.org/10.1109/TGRS.2014.2380475>.
- Hou, Q., Cheng, M., Hu, X., Borji, A., Tu, Z., & Torr, P.H.S. (2019). Deeply supervised salient object detection with short connections. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *41*(4), 815–828. <https://doi.org/10.1109/TPAMI.2018.2815688>.
- Howley, T., Madden, M. G., O'Connell, M. L., & Ryder, A.G. (2006). The effect of principal component analysis on machine learning accuracy with high-dimensional spectral data. *Knowledge-Based Systems*, *19*(5), 363–370. AI 2005 SI.
- Hoye, G., & Fridman, A. (2013). The mixel camera — keystone-free hyperspectral images. In *2013 5Th workshop on hyperspectral image and signal processing: Evolution in remote sensing (WHISPERS)* (pp. 1–4). <https://doi.org/10.1109/WHISPERS.2013.8080703>.
- Hughes, G. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, *14*(1), 55–63.
- IEC (1998). Colour management in multimedia systems—part 2: colour management, part 2.1: default rgb colour space—srgb. International Electrotechnical Commission, IEC/4WD 61966-2-1.
- Imamoglu, N., Oishi, Y., Zhang, X., Ding, G., Fang, Y., Kouyama, T., & Nakamura, R. (2018). Hyperspectral image dataset for benchmarking on salient object detection. In *2018 Tenth international conference on quality of multimedia experience (QoMEX)* (pp. 1–3). <https://doi.org/10.1109/QoMEX.2018.8463428>.
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259. <https://doi.org/10.1109/34.730558>.
- Jia, S., Shen, L., & Li, Q. (2015). Gabor feature-based collaborative representation for hyperspectral imagery classification. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(2), 1118–1129. <https://doi.org/10.1109/TGRS.2014.2334608>.
- Jun, X., Qin, T., Cuiwei, L., Ran, G., & Aidong, M (2015). Video saliency map detection based on global motion estimation. In *2015 IEEE International conference on multimedia expo workshops (ICMEW)* (pp. 1–6). <https://doi.org/10.1109/ICMEW.2015.7169845>.
- Koch, C., & Ullman, S. (1987). Shifts in selective visual attention: Towards the underlying neural circuitry. In L. M. Vaina (Ed.) *Matters of Intelligence: Conceptual Structures in Cognitive Neuroscience* (pp. 115–141). Netherlands: Springer. [https://doi.org/10.1007/978-94-009-3833-5\\_5](https://doi.org/10.1007/978-94-009-3833-5_5).
- Le Moan, S., Mansouri, A., Hardeberg, J., & Voisin, Y. (2013). Saliency for spectral image analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 1–8. <https://doi.org/10.1109/JSTARS.2013.2257989>.
- Liang, J., Zhou, J., Bai, X., & Qian, Y. (2013). Salient object detection in hyperspectral imagery. In *2013 IEEE International conference on image processing* (pp. 2393–2397). <https://doi.org/10.1109/ICIP.2013.6738493>.

- Liu, Z., Tang, J., Xiang, Q., & Zhao, P. (2020). Salient object detection for rgb-d images by generative adversarial network. *Multimedia Tools and Applications* 79. <https://doi.org/10.1007/s11042-020-09188-8>.
- Loggenberg, K., Strever, A., Greyling, B., & Poona, N. (2018). Modelling water stress in a shiraz vineyard using hyperspectral imaging and machine learning. *Remote Sensing* 10. <https://doi.org/10.3390/rs10020202>.
- Lopez-Fandino, J., Garea, A. S., Heras, D. B., & Argüello, F. (2018). Stacked autoencoders for multiclass change detection in hyperspectral images. In *2018 IEEE International geoscience and remote sensing symposium, IGARSS 2018, valencia, spain, july 22-27, 2018* (pp. 1906–1909). IEEE.
- Luo, R., Huang, H., & Wu, W. (2020). Salient object detection based on backbone enhanced network. *Image and Vision Computing*, 95, 103,876. <https://doi.org/10.1016/j.imavis.2020.103876>.
- Minka, T. (2001). Automatic choice of dimensionality for pca. Technical Report 514, MIT Media Lab Vision and Modeling Group.
- Opitz, D., & Maclin, R. (1999). Popular ensemble methods: an empirical study. *Journal of Artificial Intelligence Research*, 11, 169–198. <https://doi.org/10.1613/jair.614>.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems Man and Cybernetics*, 9, 62–66.
- Pravilovic, S., Appice, A., & Malerba, D. (2018). Leveraging correlation across space and time to interpolate geophysical data via cokriging. *International Journal of Geographical Information Science*, 32(1), 191–212.
- Pravilovic, S., Bilancia, M., Appice, A., & Malerba, D. (2017). Using multiple time series analysis for geosensor data forecasting. *Information Sciences*, 380, 31–52.
- Samat, A., Li, E., Wei, W., Liu, S., Lin, C., & Abuduwaili, J. (2020). Meta-xgboost for hyperspectral image classification using extended mser-guided morphological profiles. *Remote Sensing* 12. <https://doi.org/10.3390/rs12121973>.
- Stuart, M., McGonigle, A., & Willmott, J. (2019). Hyperspectral imaging in environmental monitoring: a review of recent developments and technological advances in compact field deployable systems. *Sensors (Basel)*, 19(14), 1–17. <https://doi.org/10.3390/s19143071>.
- Tipping, M. E., & Bishop, C. M. (2006). Mixtures of probabilistic principal component analysers. *Neural Computation*, 11(2), 443–482.
- Ullah, I., Jian, M., Hussain, S., Guo, J., Yu, H., Wang, X., & Yin, Y. (2020). A brief survey of visual saliency detection. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-020-08849-y>.
- Wang, Q., Zhang, L., Zou, W., & Kpalma, K. (2020). Salient video object detection using a virtual border and guided filter. *Pattern Recognition*, 97, 106,998. <https://doi.org/10.1016/j.patcog.2019.106998>.
- Wang, W., Shen, J., Dong, X., Borji, A., & Yang, R. (2020). Inferring salient objects from human fixations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(8), 1913–1927. <https://doi.org/10.1109/TPAMI.2019.2905607>.
- Xia, J., Ghamisi, P., Yokoya, N., & Iwasaki, A. (2018). Random forest ensembles and extended multiextinction profiles for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(1), 202–216.
- Yan, H., Zhang, Y., Wei, W., Zhang, L., & Li, Y. (2016). Salient object detection in hyperspectral imagery using spectral gradient contrast. In *2016 IEEE International geoscience and remote sensing symposium (IGARSS)* (pp. 1560–1563). <https://doi.org/10.1109/IGARSS.2016.7729398>.
- Yang, Z., & Mueller, R. (2007). Spatial-spectral cross-correlation for change detection : a case study for citrus coverage change detection. In *ASPRS 2007 Annual conference*, (Vol. 2 pp. 767–777).
- Zhang, L., Zhang, Y., Yan, H., Gao, Y., & Wei, W. (2018). Salient object detection in hyperspectral imagery using multi-scale spectral-spatial gradient. *Neurocomputing*, 291, 215–225. <https://doi.org/10.1016/j.neucom.2018.02.070>.
- Zhou, S., Sun, L., Xing, W., Feng, G., Ji, Y., Yang, J., & Liu, S. (2020). Hyperspectral imaging of beet seed germination prediction. *Infrared Physics & Technology*, 108, 103,363. <https://doi.org/10.1016/j.infrared.2020.103363>.
- Zlatintsi, A., Iosif, E., Marago, P., & Potamianos, A. (2015). Audio salient event detection and summarization using audio and text modalities. In *2015 23rd european signal processing conference (EUSIPCO)* (pp. 2311–2315). <https://doi.org/10.1109/EUSIPCO.2015.7362797>.

## Affiliations

**Annalisa Appice**<sup>1,2</sup>  · **Angelo Cannarile**<sup>1</sup> · **Antonella Falini**<sup>1</sup>  · **Donato Malerba**<sup>1,2</sup>  · **Francesca Mazzia**<sup>1</sup>  · **Cristiano Tamborrino**<sup>3</sup>

Angelo Cannarile  
a.cannarile2@studenti.uniba.it

Antonella Falini  
antonella.falini@uniba.it

Donato Malerba  
donato.malerba@uniba.it

Francesca Mazzia  
francesca.mazzia@uniba.it

Cristiano Tamborrino  
cristiano.tamborrino@uniba.it

<sup>1</sup> Dipartimento di Informatica, Università degli Studi di Bari Aldo Moro, via Orabona, 4, 70125 Bari, Italy

<sup>2</sup> Consorzio Interuniversitario Nazionale per l'Informatica - CINI, Bari, Italy

<sup>3</sup> Dipartimento di Matematica, Università degli Studi di Bari Aldo Moro, via Orabona, 4, 70125, Bari, Italy