Texture analysis using two-dimensional permutation entropy and amplitude-aware permutation entropy

Andreia Gaudêncio, Mirvana Hilal, João Cardoso, Anne Humeau-Heurtier, Pedro Vaz

Pattern Recognition Letters 159 (2022) 150–156

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Texture analysis using two-dimensional permutation entropy and amplitude-aware permutation entropy

Andreia S. Gaudêncio[a,b], Mirvana Hilal[b], João M. Cardoso[a], Anne Humeau-Heurtier[b], Pedro G. Vaz[a,*]

[a] LIBPhys, Department of Physics, University of Coimbra,Coimbra,P-3004,516,Portugal
[b] Univ Angers, LARIS, SFR MATHSTIC, Angers F-49000,France

ARTICLE INFO

Article history:
Received 22 February 2022
Revised 4 May 2022
Accepted 16 May 2022
Available online 18 May 2022

Edited by: Maria De MarsicoSCS103956

Keywords:
Bioinformatics
Entropy
Information theory
Texture,

ABSTRACT

Entropy algorithms have been applied extensively for time series analysis. The entropy value given by the algorithm quantifies the irregularity of the data structure. For higher irregular data structures, the entropy is higher. Both permutation entropy (PE) and amplitude-aware permutation entropy (AAPE) have been previously used to analyze time series. These two metrics have the advantage, over others, of being computationally fast and simple. However, fewer entropy measures have been proposed to process images. Two-dimensional entropy algorithms can be used to study texture and analyze the irregular structure of images. Herein, we propose the extension of AAPE for two-dimensional analysis ($AAPE_{2D}$). To the best of our knowledge, $AAPE_{2D}$ has never been proposed to analyze texture of images. For comparison purposes, we also study the two-dimensional permutation entropy (PE_{2D}) to analyze the effect of the amplitude consideration in texture analysis. In this study, we compare $AAPE_{2D}$ method with PE_{2D} in terms of irregularity discrimination, parameters sensitivity, and artificial texture differentiation. Both $AAPE_{2D}$ and PE_{2D} appear to be interesting entropy-based approaches for image texture analysis. When applied to a biomedical dataset of chest X-rays with healthy subjects and pneumonia patients, both methods showed to statistically differentiate both groups for $P < 0.01$. Finally, using a SVM model and multiscale entropy values as features, $AAPE_{2D}$ achieves an average of 75.7% accuracy which is slightly better than the results of PE_{2D}. Overall, both entropy algorithms are promising and achieve similar conclusions. This work is a new step towards the development of other entropy-based texture measures.

1. Introduction

Entropy-based metrics are known for being used to study the irregularity of biomedical signals. In the last few decades, several entropy-based algorithms were proposed, namely, approximate entropy [1], sample entropy [2], fuzzy entropy [3], and permutation entropy [4]. Although most entropy-based methods rely on the probability distribution of the data, they also require the adjustment of control parameters. Permutation entropy (PE) is a simple and fast method, with few adjustment parameters, used to estimate the entropy value based on the concept of counting *permutation patterns* [4]. Nonetheless, this promising technique presents one main issue: the amplitude fluctuations and variations within the signal are not accounted for [5]. This means that PE might not be able to discriminate different signal's behaviors by ignoring its structure [6]. Based on this, Azami and Escudero [5] proposed a new method denominated as amplitude-aware permutation entropy (AAPE) to solve this drawback for one-dimensional data.

When applied to images, these metrics allow to analyze different textures. An image associated with a more irregular texture will have a higher entropy value. Several entropy approaches have also been used to process two-dimensional (2D) data, such as: sample entropy 2D [7], fuzzy entropy 2D (FE_{2D}) [8], colored fuzzy entropy 2D [9], complexity-entropy causality plane 2D ($CECP_{2D}$) [10], and also permutation entropy 2D (PE_{2D}) [11]. These entropy algorithms can be used to study texture and analyze the irregular structure of images [12]. For example, FE_{2D} is known for extracting squared-patterns to establish comparisons between these patterns, which has been proven useful for texture characterization, even though it can be computationally expensive [8]. A different strategy for $CECP_{2D}$ is adopted where the image is considered as

* Corresponding author.
E-mail addresses: andreia.gaudencio@student.uc.pt (A.S. Gaudêncio), mirvana.hilal@univ-angers.fr (M. Hilal), jmrcardoso@uc.pt (J.M. Cardoso), anne.humeau@univ-angers.fr (A. Humeau-Heurtier), pvaz@uc.pt (P.G. Vaz).

https://doi.org/10.1016/j.patrec.2022.05.017

a 2D ordered array. The entropy is obtained through the estimation of the *ordinal* probability distribution, that relies on calculating the number of *ordinal patterns* which are obtained through the same process as *permutation patterns*. However, $CECP_{2D}$ obtains the normalized entropy with the normalized Shannon entropy equation and also assesses the statistical complexity value [10]. In [13], the concept of $CECP_{2D}$ is used to compare different paintings using the normalized permutation entropy and complexity concepts.

Based on the advantages of AAPE over PE for time series, we herein propose the extension of AAPE for the analysis of 2D data ($AAPE_{2D}$). To the best of our knowledge, $AAPE_{2D}$ algorithm has never been proposed to analyze texture of images. We also compare the behavior of $AAPE_{2D}$ with PE_{2D} to analyze the effect of the amplitude consideration in texture analysis. Finally, we apply both entropy algorithms to a chest X-rays dataset of healthy subjects and subjects diagnosed with pneumonia [14]. Pneumonia is a pulmonary infection whose diagnosis relies on tracking the symptoms and the signs of respiratory tract infection presence, radiological changes, the identification of a putative pathogen, and the treatment response consistent with pneumonia [15]. In Europe, mortality rates of hospitalized patients associated with community-acquired pneumonia are around 5–20% [15]. Usually, pneumonia is identified when lung consolidation, a radiological hallmark, is verified using chest X-rays [15,16]. Therefore, entropy algorithms can be helpful in the characterization of texture properties of lung consolidation verified in chest X-rays and CT scans when pneumonia is diagnosed.

The remaining of the paper is divided as follows: Section 2 describes the entropy algorithms used herein, details both the synthetic textures and biomedical images used, and describes the algorithms' validation tests; Section 3 discusses the results on PE_{2D} and the novel measure $AAPE_{2D}$; finally, Section 4 summarizes the main results, debates future work, and discusses the use of these entropy metrics as texture descriptors and features.

2. Methods and materials

Both PE and AAPE are based on the determination of the probability distribution of *permutation patterns*. The permutation patterns are obtained after re-arranging the positions of the initial patterns obtained from the signal/image. These approaches differ in how the probability is accounted for. The first approach, PE, is simpler as it only accounts for how many patterns exist for each permutation pattern. The second approach, AAPE, applies a correction factor by using an amplitude factor based on the mean and standard deviation of neighbouring points within the pattern. The two-dimensional definitions of PE and AAPE, PE_{2D} and $AAPE_{2D}$, respectively, are described below.

2.1. Entropy algorithms

First, consider an image $\mathbf{X}$ of width W and height H where a and b are the position of an element in $\mathbf{X}$ and $1 \le a \le W$ and $1 \le b \le H$. Considering the embedding dimension parameter m, we can define squared templates from the image as:

$$X_{a,b}^{m} = \begin{bmatrix} X_{a,b} & \cdots & X_{a,b+m-1} \\ X_{a+1,b} & \cdots & X_{a+1,b+m-1} \\ \cdots & \cdots & \cdots \\ X_{a+m-1,b} & \cdots & X_{a+m-1,b+m-1} \end{bmatrix} \quad (1)$$

where $1 \le a \le W-m+1$ and $1 \le b \le H-m+1$, making a total of $N_m = (W-m+1) \times (H-m+1)$ possible templates. For both approaches, PE_{2D} and $AAPE_{2D}$, the strategy herein used is to vectorize the squared templates as follows:

$$X_{a,b}^{'m} = \{X_{a,b}, \ldots, X_{a,b+m-1}, X_{a+1,b}, \ldots, X_{a+1,b+m-1}, \ldots, X_{a+m-1,b}, \ldots, X_{a+m-1,b+m-1}\}. \quad (2)$$

Similarly to the 1D approaches [4,5], the intensities of the templates must be rearranged in ascending order to determine the corresponding permutation pattern. As the template $X_{a,b}^{'m}$ has m^2 points, there will be $D = (m \times m)!$ permutation patterns, π_j^m (where $1 \le j \le D$). To fully understand this procedure, consider the example of the template shown in Fig. 1 for $m = 2$. The template is squared and has dimensions of 2×2. Step 1 illustrates the transformation of this template into a vector with 4 points. Afterwards, to obtain the corresponding permutation pattern, this vector $X_{a,b}^{'m}$ is rearranged according to its intensity in an ascending order (step 2). Finally, the order of the original positions of the intensity values give the corresponding permutation pattern π_j^m (step 3) shown in orange.

Then, for the PE_{2D} [11] approach, the probability of each permutation pattern π_j^m is defined as in Eq. (3).

$$p_p(\pi_j^m) = \frac{\#\{X_{ab}^{'m} \text{ produces } \pi_j^m\}}{N_m} \quad (3)$$

represents the *number of* [11]. Basically, this probability depends on the amount of times a certain permutation pattern is repeated within the image.

The PE algorithm does not consider the amplitude variations within the data structure. Furthermore, different templates with different amplitude variations or different mean values can lead to the equal permutation patterns discarding this information. Another issue is when equal intensity values exist within the template and a conflict in the ordering process emerges [5]. Bandt and Pompe [4] proposed adding noise to the data structure to avoid this but this can result in imprecise results [5]. Based on Azami and Escudero [5] approach to solve these issues, we herein propose the $AAPE_{2D}$ algorithm that uses a weighting factor A to obtain the probability of π_j^m in Eq. (4). This factor defines the weight of the mean of consecutive samples (first term of Eq. (5)) and the weight of their absolute difference values (second term of Eq. (5)). This weighting process is included in the probability definition as follows:

$$p_a(\pi_j^m) = \frac{\sum \delta_{ab} \text{ if } X_{ab}^{'m} \text{ produces } \pi_j^m}{\sum_{a=1}^{W-m+1} \sum_{b=1}^{H-m+1} \delta_{ab}} \quad (4)$$

where δ_{ab} is:

$$\delta_{ab} = \frac{A}{m^2} \sum_{k=1}^{m^2} \left| X_{ab}^{'}(k) \right| + \frac{1-A}{m^2-1} \sum_{k=1}^{m^2-1} \left| X_{ab}^{'}(k+1) - X_{ab}^{'}(k) \right| \quad (5)$$

where A is the adjusting coefficient between [0,1]. For $A = 0$, only the differences between consecutive samples are considered; for $A = 1$, only the mean value of consecutive samples is considered for the probability calculus; for $A < 0.5$, the differences between consecutive samples have more weight than the mean value of consecutive samples, and the opposite happens when $A > 0.5$ [5]. When considering the same example in Fig. 1, the first three steps are the same as for PE. However, to define the probability, a fourth step is required, corresponding to the determination of δ_{ab}.

Thus, we can define PE_{2D} and $AAPE_{2D}$ according to Eqs. (6) and (7), respectively:

$$PE_{2D} = -\sum_{j=1}^{m!} p_p(\pi_j^m) \times \log(p_p(\pi_j^m)) \quad (6)$$

$$AAPE_{2D} = -\sum_{j=1}^{m!} p_a(\pi_j^m) \times \log(p_a(\pi_j^m)). \quad (7)$$

These algorithms were applied using the *numba* package [17] of Python® v3.7 in a AMD Ryzen 7 5800H with Radeon Graphics 3.20 GHz.

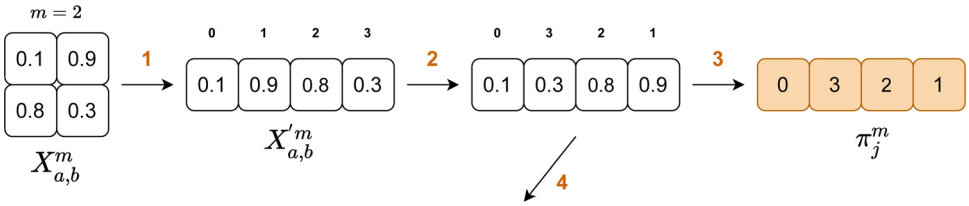


Fig. 1. Example illustrating the method to obtain the permutation pattern π^m_j from a template $X^m_{a,b}$ and δ_{ab}, using the weighting factor A.

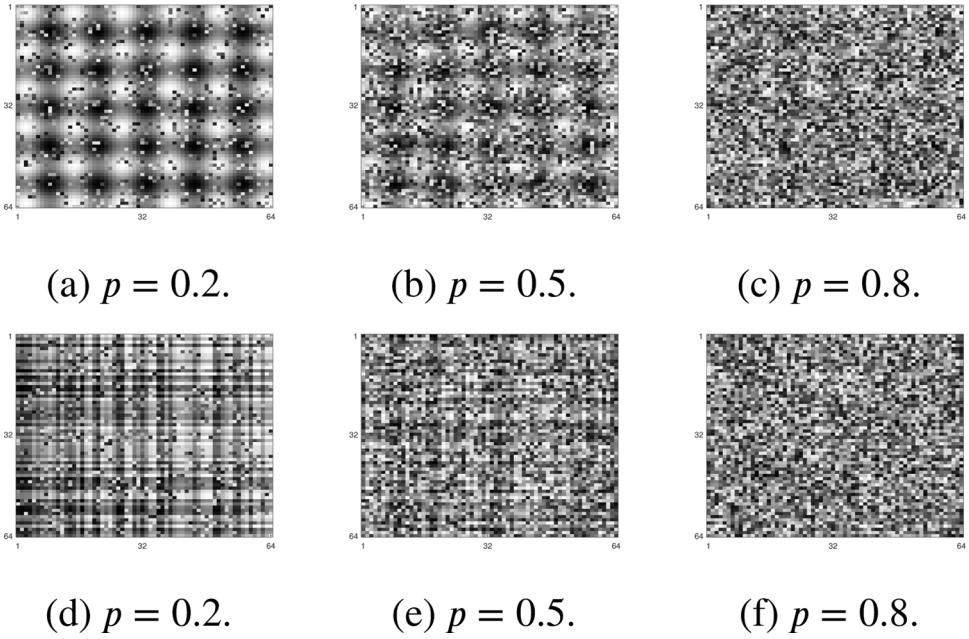


Fig. 2. $MIX_{2D}(p)$ images (2a-2c) and their shuffled versions (2a-2f) with a size of 64×64 pixels for $p = 0.2$, $p = 0.5$, and $p = 0.8$.

2.2. Synthetic textures and validation tests

First, to evaluate the effect of the embedding parameter on the irregularity discrimination performance, we used 5 $MIX_{2D}(p)$ images [7] of 64×64 pixels, where the p-level indicates the image irregularity. When p increases, so does the irregularity of the image. Figure 2 shows examples of $MIX_{2D}(p)$ images for $p = 0.2$, $p = 0.5$, and $p = 0.8$. Each image was tested by using $m = 2$ and $m = 3$ for both entropy algorithms and the p-level was varied within [0,1]. Besides, we assessed the influence of the coefficient A. The $AAPE_{2D}$ was tested for $A = 0.1$, $A = 0.5$, and $A = 0.9$.

Then, we compared the values of PE_{2D} and $AAPE_{2D}$ for $A = 0.1$, using the original $MIX_{2D}(p)$ images as shown in Fig. 2a–c and their shuffled versions in Fig. 2d–f. The process of shuffling the image pixels decreases the regular order present in the image, increasing, therefore, the irregularity.

We also assessed the computational cost of PE_{2D} and $AAPE_{2D}$ with 5 $MIX_{2D}(p = 0.6)$ images using $m = 2$ and $A = 0.1$. We present the average computational times and their standard deviation values.

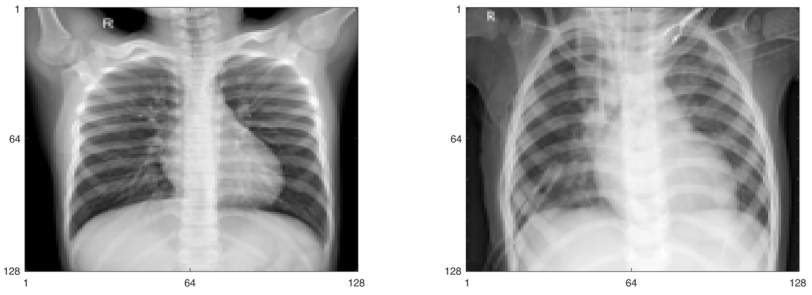
2.3. Biomedical dataset

Finally, we applied both algorithms, PE_{2D} and $AAPE_{2D}$, using $m = 2$ and $A = 0.1$ for $AAPE_{2D}$, to a public biomedical dataset of chest x-rays with a total of 5856 validated images of healthy subjects and subjects diagnosed with pneumonia [14]. The dataset is composed of 1583 normal chest X-rays and 4273 chest X-rays of patients diagnosed with pneumonia [14]. Figure 3 shows examples of such chest X-ray images (healthy person and a patient diagnosed with pneumonia). These images were resized to 128×128 pixels in order to have a consistent image size throughout the dataset for calculating the entropy values. The mean (μ) and standard deviation (σ) values of these images were also assessed for both groups. We verified that the images of the healthy group have a value of 122.6 ± 62.7 ($\mu \pm \sigma$) and that the images corresponding to the pneumonia group have a value of 122.9 ± 59.3 ($\mu \pm \sigma$). This allows us to conclude that the contrast of the whole dataset will not influence the texture analysis.

Based on multiscale analysis [18], we applied the downsampling (or *coarse-graining*) procedure to obtain different image structures for each scale factor (τ) according to the following equation:

$$\mathbf{Y}_{i,j}{}^{(\tau)} = \frac{1}{\tau^2} \sum_{\substack{l=(i-1)\tau+1 \\ m=(j-1)\tau+1}}^{\substack{j\tau \\ i\tau}} \mathbf{X}_{l,m}, \tag{8}$$

where $\mathbf{Y}^\tau$ is the new version of the image $\mathbf{X}$ having a size of $\frac{W}{\tau} \times \frac{H}{\tau}$ pixels. For $\tau = 1$, $\mathbf{Y}^1$ corresponds to the original image. For each coarse-grained image, $\mathbf{Y}^\tau$, we obtained the PE_{2D} and $AAPE_{2D}$



(a) Healthy. (b) Pneumonia.

Fig. 3. Resized chest X-rays images from a healthy individual and pneumonia patient from Kermany et al. [14].

entropy values. We have established a maximum scale factor of $\tau = 6$. Therefore, the smallest coarse-grained images have 21×21 pixels.

We first analyzed the mean entropy values given by PE_{2D} and $AAPE_{2D}$ for $1 \leq \tau \leq 6$. Afterwards, we assessed the normality for each scale factor ($1 \leq \tau \leq 6$) and each group (healthy and pneumonia) using the *Shapiro-Wilk* test for a significance level of $\alpha = 0.01$. For these conditions, both groups for all scale factors were verified to be normal. Then, we used a *t-student test* to differentiate the healthy and patient group through entropy values (for $1 \leq \tau \leq 6$) for $P < 0.01$.

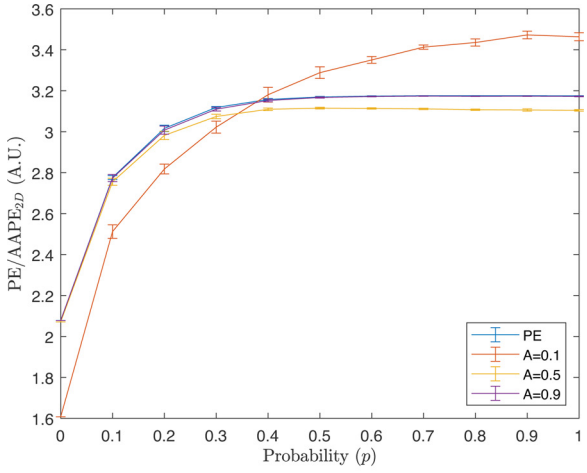
Since a large dataset ($N > 100$) is being used in this work, the P-values achieved during t-student tests are of difficult interpretation [19]. To overcome this artifact, we also verified the minimum percentage of samples required, for each scale factor, to statistically differentiate both groups using a t-student test ($P < 0.01$). The subset of entropy values for each percentage was randomly selected 5 times. Then, the mean P-value was determined for that percentage subset.

Afterwards, we tested a support vector machine (SVM) model with a radial basis function kernel to classify pneumonia and healthy subjects. We used 30% of the total dataset mentioned earlier as a test dataset. For the training dataset, a 5-fold cross-validation procedure was used, and a grid search for parameters optimization was also performed. The parameters to be optimized were C and γ: C was varied between 0.1 and 10000; and, γ was varied between 1×10^{-5} and 1. For this classification model, we used as features the entropy values of the 6 scale factors for each algorithm. We compared the results of the model using PE_{2D} and $AAPE_{2D}$ features in terms of accuracy, sensitivity, and specificity.

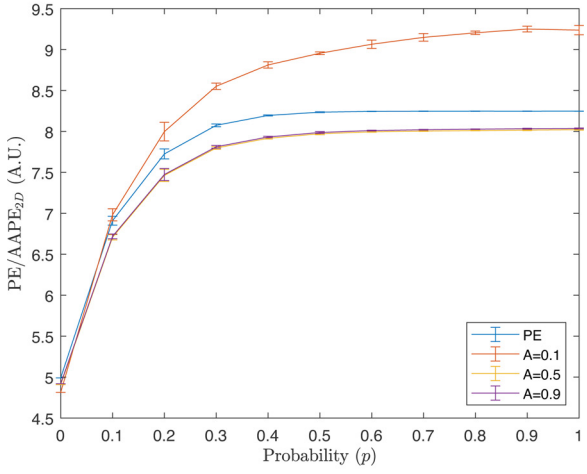
3. Results and discussion

3.1. Synthetic textures

The influence of the adjusting coefficient and embedding dimension can be observed in Fig. 4. Regardless of the parameters, both algorithms show the ability to discriminate the irregular behavior of $MIX_{2D}(p)$ images resulting in higher entropy when the p-level increases. Figure 4(a) shows that using $m = 2$ the curves of PE_{2D} and $AAPE_{2D}$ with $A = 0.5$ and $A = 0.9$ are very similar. In fact, between $p = 0.0$ and $p = 1$ (most regular and most irregular MIX_{2D} image, respectively), both PE_{2D} and $AAPE_{2D}$ using $A = 0.5$ show a relative increase of 33.3%. For $A = 0.9$, $AAPE_{2D}$ shows a relative increase of 34.5%. When using $A = 0.1$, $AAPE_{2D}$ achieves a relative increase of 53.6%. In Fig. 4(b), for $m = 3$, the relative increase of PE_{2D} for $0.0 \leq p \leq 1.0$ improved to 38.8%. In addition, for $A = 0.5$ and $A = 0.9$, $AAPE_{2D}$ also improved its relative increase to 39.5%. For m-values, the performance of $AAPE_{2D}$ is best when choosing $A = 0.1$. This means that weighting more the consecutive samples' differences instead of their mean value produces better results in discriminating irregularity of images. Therefore, for the following tests, PE_{2D} is compared with $AAPE_{2D}$ using $A = 0.1$.



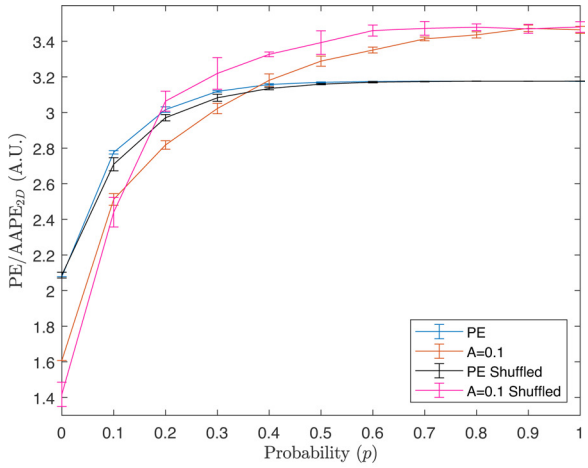
(a) $m = 2$.



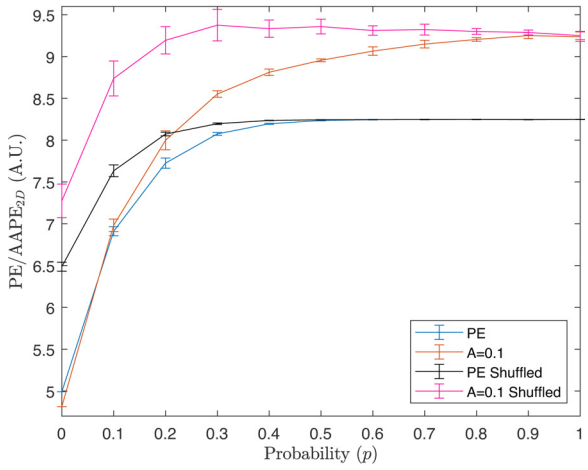
(b) $m = 3$.

Fig. 4. PE_{2D} (blue) and $AAPE_{2D}$ values using $A = 0.1$ (orange), $A = 0.5$ (yellow), and $A = 0.9$ (purple) for $MIX_{2D}(p)$ images of 64×64 pixels, $0 \leq p \leq 1$ (step of 0.1), $m = 2$ and $m = 3$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Figure 5 shows the results of shuffling the pixels of the images, *i.e.*, reordering randomly the pixels of the $MIX_{2D}(p)$ images. This

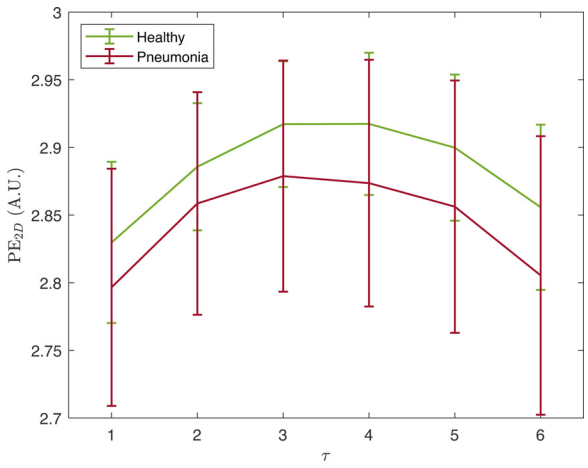


(a) $m = 2$.

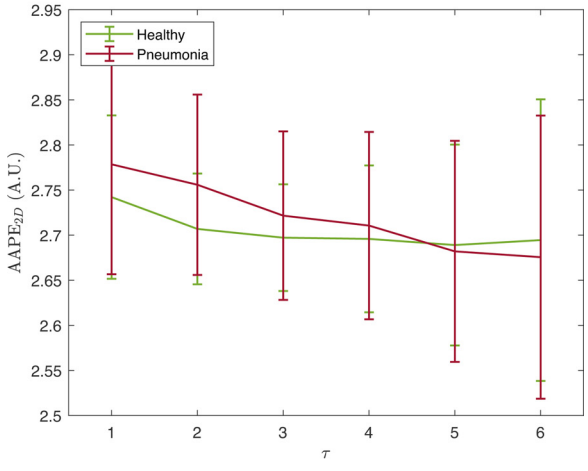


(b) $m = 3$.

Fig. 5. PE$_{2D}$ (blue) and AAPE$_{2D}$ using $A = 0.1$ (orange) for MIX$_{2D}(p)$ images and their shuffled versions (black and pink, respectively) of 64×64 pixels, using $0 \leq p \leq 1$ (step of 0.1) for $m = 2$ and $m = 3$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



(a) PE$_{2D}$.



(b) AAPE$_{2D}$.

Fig. 6. Multiscale PE$_{2D}$ and AAPE$_{2D}$ entropy values, using $A = 0.1$, $m = 2$, and $1 \leq \tau \leq 6$, for the healthy (green) and pneumonia (red) groups. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

reordering should reflect an increase of entropy values as the shuffling process increases the irregularity of the image, specially in more regular images.

In Fig. 5(a), when using $m = 2$, for PE$_{2D}$, the entropy values of the original images (blue) and their shuffled versions (in black) are very similar. Moreover, the PE$_{2D}$ values of the original MIX$_{2D}$ images is slightly higher than the shuffled images for $p \leq 0.5$.

For $m = 3$, an increase between PE$_{2D}$ for the original MIX$_{2D}$ images (blue) and their shuffled versions (black) can be observed (see Fig. 5(b)). In the previous test of irregularity discrimination (Fig. 4), we had also verified a slightly better performance of PE$_{2D}$ for $m = 3$. For AAPE$_{2D}$, there is a substantial increase of entropy for the shuffled versions (pink) of the MIX$_{2D}$ images. However, previously, when analyzing the irregularity discrimination, AAPE$_{2D}$ achieved a higher irregularity discrimination for $m = 2$. Based on this, the best $m-$value for AAPE$_{2D}$ is 2. For PE$_{2D}$, the performance using $m = 2$ or $m = 3$ is similar. Therefore, for comparison purposes, for both algorithms, the $m-$value was fixed to $m = 2$ for the following tests.

In terms of computational cost, when analysing the time required to compute a MIX$_{2D}(p)$ image ($p = 0.6$) of 64×64 pixels, PE$_{2D}$ takes an average of 14.3 ± 0.2 ms and AAPE$_{2D}$ consumes 14.5 ± 0.1 ms, having only a difference of 1.4%. We can conclude that these algorithms are extremely fast in processing images, which is usually a concern for other entropy algorithms.

3.2. Biomedical dataset

In Fig. 6, we observe that for PE$_{2D}$, the entropy values are lower for the patients group (red) than for healthy (green) subjects regardless of the scale factor. However, for AAPE$_{2D}$, the entropy is higher for the patients group for $1 \leq \tau \leq 4$. For $\tau \geq 5$, this tendency inverts and the entropy for the healthy subjects is higher. Furthermore, we can observe that the PE$_{2D}$ curves behavior is similar for both groups. First, the entropy increases between $1 \leq \tau \leq 3$ and then, it decreases between $4 \leq \tau \leq 6$. For AAPE$_{2D}$, overall, both curves show a decreasing behavior but the curves cross each other. In Table 1, we verify that PE$_{2D}$ is able to statistically differentiate both groups for all the scale factors considered. However, AAPE$_{2D}$ only differentiates 5 out 6 scale factors. Moreover, in most cases, the P-value for PE$_{2D}$ is considerably smaller than for AAPE$_{2D}$. Therefore, both methods seem to be good options when characterizing these different biomedical textures.

Given the number of samples within the dataset, the low P-values observed can also be derived from the dataset's size. In [19],

Table 1
P-values for multiscale PE_{2D} and $AAPE_{2D}$ with $1 \leq \tau \leq 6$, $m = 2$, and $A = 0.1$.*- statistical significance for $P < 0.01$.

Scale Factors	PE_{2D}	$AAPE_{2D}$
1	$2.76 \times 10^{-43*}$	$5.32 \times 10^{-27*}$
2	$7.66 \times 10^{-35*}$	$2.05 \times 10^{-72*}$
3	$3.67 \times 10^{-63*}$	$4.33 \times 10^{-22*}$
4	$1.07 \times 10^{-70*}$	$3.59 \times 10^{-7*}$
5	$3.42 \times 10^{-67*}$	0.05
6	$7.71 \times 10^{-73*}$	$4.12 \times 10^{-5*}$

Table 2
P-values for multiscale PE_{2D} and $AAPE_{2D}$ for $1 \leq \tau \leq 6$, $m = 2$, and $A = 0.1$, for the minimum percentage of samples from the original dataset that allows to statistically differentiate the healthy and the pneumonia groups (except for $AAPE_{2D}$ when $\tau = 5$). *- number of samples percentage lower than < 10%. POS - percentage of samples used.

τ	PE		AAPE	
	POS (%)	P-value	POS (%)	P-value
1	5*	1.71E-04	8*	4.28E-04
2	8*	5.89E-03	6*	7.37E-06
3	2*	2.65E-03	12.5	2.25E-04
4	5*	6.24E-04	45	3.37E-03
5	3*	2.47E-04	100	0.04
6	2*	9.98E-03	50	5.20E-03

Table 3
Classification values of multiscale PE_{2D} and $AAPE_{2D}$ features.

	PE_{2D}	$AAPE_{2D}$
Accuracy	$75.5 \pm 0.4\%$	$75.7 \pm 0.7\%$
Precision	$75.6 \pm 0.6\%$	$77.7 \pm 0.7\%$
Sensitivity	$95.6 \pm 1.3\%$	$93.3 \pm 0.6\%$
Specificity	$26.8 \pm 3.2\%$	$28.8 \pm 1.9\%$

the large number of samples is indicated to possibly influence the statistical analysis as the low *P*-values can be artifacts of the large sample size. Therefore, it is suggested, for example, to verify the dependence of the *P*-values with different sample sizes. Based on this, we show, in Table 2, the statistical significance (for $P < 0.01$) between the healthy and pneumonia groups according to different percentages of samples from the original dataset. We verified that the multiscale PE_{2D} method performs better for every scale factor as it can statistically differentiate both groups for less than 8% of number of samples. For $AAPE_{2D}$, this can only be verified for the first 2 scale factors.

Finally, the SVM model was able to achieve a mean accuracy value of 75.5% and 75.7% for multiscale PE_{2D} and $AAPE_{2D}$ features, respectively, with $AAPE_{2D}$ achieving a slightly higher accuracy. Moreover, $AAPE_{2D}$ achieved higher precision, and specificity (see Table 3). However, PE_{2D} shows a slightly higher sensitivity, *i.e.*, it has a better ability to classify pneumonia cases when the images are from pneumonia patients. Hence, one can say that both $AAPE_{2D}$ and PE_{2D} have similar performances in classifying pneumonia, resulting in two reliable and promising texture descriptors for this particular biomedical imaging application.

4. Conclusion and final remarks

In this analysis, it was proven that $AAPE_{2D}$ and PE_{2D} are interesting methods for image analysis and texture discrimination. When analysing the two main parameters, the embedding dimension, m, and the adjusting coefficient, A, we observed that the best parameters for comparison purposes were $m = 2$ and $A = 0.1$. These parameters allow to have a relatively good irregularity discrimination between several images with increasing irregularity behavior. Moreover, when differentiating several artificial textures, overall, $AAPE_{2D}$ shows a slightly better precision and ability to differentiate several textures and their synthesized versions. However, we can say that both methods, PE_{2D} and $AAPE_{2D}$, perform similarly and can discriminate irregularity clearly.

For the biomedical chest X-rays dataset, it was verified that PE_{2D} leads, for most scale factors, to lower *P*-values than the ones obtained with $AAPE_{2D}$. Furthermore, the PE_{2D} curves behavior is similar for both groups with increasing entropy between $1 \leq \tau \leq 3$, followed by a decrease between $4 \leq \tau \leq 6$. Nevertheless, $AAPE_{2D}$ curves show a decreasing behavior. First, with the pneumonia group having higher entropy than the healthy group, and then, for $\tau \geq 4$, with the healthy group having higher entropy values. Therefore, one can conclude that both algorithms differentiate the healthy from the pneumonia group. However, the curves are different according to the chosen algorithm.

Afterwards, we observed the effect size of the dataset in the statistical analysis by reducing the number of samples used for verifying the statistical significance using $P < 0.01$ for both groups. We concluded that PE_{2D} outperforms $AAPE_{2D}$ as it can statistically differentiate pneumonia patients from healthy individuals by using less than 8% of samples for all the scale factors tested.

When analysing the multiscale PE_{2D} and $AAPE_{2D}$ features performances, using a simple SVM classification model, we observe that $AAPE_{2D}$ has a slightly better performance in terms of accuracy, and precision. However, PE_{2D} features achieve a slightly better sensitivity. This leads to the conclusion that both entropy metrics are reliable and promising texture descriptors for this particular biomedical imaging application. Furthermore, the fact that most or even all scale factors entropy values were able to differentiate both groups with a 99% confidence level, can indicate that these texture descriptors as features can be relevant for similar texture biomedical applications such as other pulmonary diseases diagnosed with chest X-rays or CT scans.

For this particular biomedical dataset, both algorithms achieve similar results. Therefore, to further discuss the advantage of using the amplitude-aware method applied to permutation entropy, it could be interesting to obtain a dataset composed of images with several amplitudes using an histogram equalization technique.

PE_{2D} and $AAPE_{2D}$ can be proven advantageous for image analysis as they are fast and allow to obtain multiple texture and irregularity features when using multiscale analysis. In addition, they are easy to interpret in terms of irregularity content of the image and allow to construct a more understandable and easy classification model than the state-of-the-art deep learning models which use the complete image as a feature input. This can allow a more simple bridge between the image processing method and the medical analysis and increase the medical and patient confidence in diagnosis aid algorithms.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Andreia S. Gaudêncio: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing. **Mirvana Hilal:** Writing – review & editing. **João M. Cardoso:** Methodology, Writing – review & editing, Supervision. **Anne Humeau-Heurtier:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Pedro G. Vaz:** Conceptualization, Methodology, Writing – review & editing, Supervision.

Acknowledgements

This work was supported by FCT (Fundação para a Ciência e Tecnologia) under the projects UIDP/04559/2020 and UIDB/04559/2020 to fund human resources and activities of Laboratory for Instrumentation, Biomedical Engineering and Radiation Physics, and under the project PTDC/EMD-TLM/30295/2017 of European Regional Development Fund (PT-COMPETE 2020). This work was also support by both FCT and the ESF (European Social Fund) under the scholarship UI/BD/152802/2022. The authors also acknowledge the Laboratory for Advanced Computing at University of Coimbra for providing computing resources that have contributed to the research results reported within this paper (https://www.uc.pt/lca).

References

[1] S.M. Pincus, Approximate entropy as a measure of system complexity, Proc. Natl. Acad. Sci. 88 (6) (1991) 2297–2301, doi:10.1073/pnas.88.6.2297.
[2] J.S. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am. J. Physiol.-HeartCirc. Physiol. 278 (6) (2000) H2039–H2049, doi:10.1152/ajpheart.2000.278.6.h2039.
[3] W. Chen, Z. Wang, H. Xie, W. Yu, Characterization of surface EMG signal based on fuzzy entropy, IEEE Trans. Neural Syst. Rehabil. Eng. 15 (2) (2007) 266–272, doi:10.1109/tnsre.2007.897025.
[4] C. Bandt, B. Pompe, Permutation entropy: a natural complexity measure for time series, Phys. Rev. Lett. 88 (2002) 174102, doi:10.1103/PhysRevLett.88.174102.
[5] H. Azami, J. Escudero, Amplitude-aware permutation entropy: illustration in spike detection and signal segmentation, Comput. Methods Programs Biomed. 128 (2016) 40–51, doi:10.1016/j.cmpb.2016.02.008.
[6] Z. Zhang, Z. Xiang, Y. Chen, J.X. and, Fuzzy permutation entropy derived from a novel distance between segments of time series, AIMS Math. 5 (6) (2020) 6244–6260, doi:10.3934/math.2020402.
[7] L.E.V. Silva, A.C.S.S. Filho, V.P.S. Fazan, J.C. Felipe, L.O.M. Junior, Two-dimensional sample entropy: assessing image texture through irregularity, Biomed. Phys. Eng. Express 2 (4) (2016) 045002, doi:10.1088/2057-1976/2/4/045002.
[8] M. Hilal, C. Berthin, L. Martin, H. Azami, A. Humeau-Heurtier, Bidimensional multiscale fuzzy entropy and its application to pseudoxanthoma elasticum, IEEE Trans. Biomed. Eng. 67 (7) (2020) 2015–2022, doi:10.1109/tbme.2019.2953681.
[9] M. Hilal, A.S.F. Gaudêncio, C. Berthin, P.G. Vaz, J. Cardoso, L. Martin, A. Humeau-Heurtier, Bidimensional colored fuzzy entropy measure: a cutaneous microcirculation study, in: 2019 Fifth International Conference on Advances in Biomedical Engineering (ICABME), IEEE, 2019, pp. 1–4.
[10] L. Zunino, H.V. Ribeiro, Discriminating image textures with the multiscale two-dimensional complexity-entropy causality plane, Chaos Solitons Fractals 91 (2016) 679–688.
[11] C. Morel, A. Humeau-Heurtier, Multiscale permutation entropy for two-dimensional patterns, Pattern Recognit. Lett. 150 (2021) 139–146, doi:10.1016/j.patrec.2021.06.028.
[12] A. Humeau-Heurtier, Texture feature extraction methods: a survey, IEEE Access 7 (2019) 8975–9000, doi:10.1109/ACCESS.2018.2890743.
[13] H.Y. Sigaki, M. Perc, H.V. Ribeiro, History of art paintings through the lens of entropy and complexity, Proc. Natl. Acad. Sci. 115 (37) (2018) E8585–E8594.
[14] D. Kermany, K. Zhang, M. Goldbaum, et al., Labeled optical coherence tomography (OCT) and chest X-ray images for classification, Mendeley Data 2 (2) (2018).
[15] W.S. Lim, Pneumonia–overview, Ref. Module Biomed. Sci. (2020).
[16] C. Cillóniz, S. Ewig, E. Polverino, C. Muñoz-Almagro, F. Marco, A. Gabarrús, R. Menéndez, J. Mensa, A. Torres, Pulmonary complications of pneumococcal community-acquired pneumonia: incidence, predictors, and outcomes, Clin. Microbiol. Infect. 18 (11) (2012) 1134–1142.
[17] I. Anaconda, et al., Numba: Numba makes python code fast, 2021, accessed:29.11.2021. https://numba.pydata.org/.
[18] M. Costa, A.L. Goldberger, C.-K. Peng, Multiscale entropy analysis of biological signals, Phys. Rev. E 71 (2) (2005) 021906.
[19] H. Lucas, G. Shmueli, et al., Too big to fail: large samples and the p-value problem, Inf. Syst. Res. 24 (4) (2013) 906–917.