Bidimensional ensemble entropy: Concepts and application to emphysema lung computerized tomography scans

Andreia Gaudêncio, Hamed Azami, João Cardoso, Pedro Vaz, Anne Humeau-Heurtier

▶ **To cite this version:**

Andreia Gaudêncio, Hamed Azami, João Cardoso, Pedro Vaz, Anne Humeau-Heurtier. Bidimensional ensemble entropy: Concepts and application to emphysema lung computerized tomography scans. Computer Methods and Programs in Biomedicine, 2023, 242, pp.107855. 10.1016/j.cmpb.2023.107855 . hal-04232582

HAL Id: hal-04232582

https://univ-angers.hal.science/hal-04232582v1

Submitted on 23 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Computer Methods and Programs in Biomedicine 242 (2023) 107855

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

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

Bidimensional ensemble entropy: Concepts and application to emphysema lung computerized tomography scans ☆

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

[a] LIBPhys, Department of Physics, University of Coimbra, Coimbra, P-3004 516, Portugal
[b] Univ Angers, LARIS, SFR MATHSTIC, F-49000 Angers, France
[c] Centre for Addiction and Mental Health, Toronto Dementia Research Alliance, Univ Toronto, Toronto, ON, Canada

ARTICLE INFO

Keywords:
Bioinformatics
Computed tomography
Emphysema
Ensemble
Entropy
Texture

ABSTRACT

Background and Objective: Bidimensional entropy algorithms provide meaningful quantitative information on image textures. These algorithms have the advantage of relying on well-known one-dimensional entropy measures dedicated to the analysis of time series. However, uni- and bidimensional algorithms require the adjustment of some parameters that influence the obtained results or even findings. To address this, ensemble entropy techniques have recently emerged as a solution for signal analysis, offering greater stability and reduced bias in data patterns during entropy estimation. However, such algorithms have not yet been extended to their two-dimensional forms.

Methods: We therefore propose six bidimensional algorithms, namely ensemble sample entropy, ensemble permutation entropy, ensemble dispersion entropy, ensemble distribution entropy, and two versions of ensemble fuzzy entropy based on different models or parameters initialization of an entropy algorithm. These new measures are first tested on synthetic images and further applied to a biomedical dataset.

Results: The results suggest that ensemble techniques are able to detect different levels of image dynamics and their degrees of randomness. These methods lead to more stable entropy values (lower coefficients of variations) for the synthetic data. The results also show that these new measures can obtain up to 92.7% accuracy and 88.4% sensitivity when classifying patients with pulmonary emphysema through a *k*-nearest neighbors algorithm.

Conclusions: This is a further step towards the potential clinical deployment of bidimensional ensemble approaches to detect different levels of image dynamics and their successful performance on emphysema lung computerized tomography scans. These bidimensional ensemble entropy algorithms have potential to be used in various imaging applications thanks to their ability to distinguish more stable and less biased image patterns compared to their original counterparts.

1. Introduction

Several bidimensional entropy algorithms have been proposed to study two-dimensional data which are of great interest for medical imaging [1–13]. Entropy is linked to irregularity or uncertainty in the data's structure. When an image is more irregular, its entropy is higher [1,2].

Two-dimensional approximate entropy ($ApEn_{2D}$) [3], derived from univariate approximate entropy [14], was proposed to differentiate cancer cells from healthy ones. Nonetheless, approximate entropy is a biased method (because it takes into account self-match patterns in its algorithm) with possible unreliable results for small-sized images [15]. To alleviate the limitation of self-match patterns in $ApEn_{2D}$, two-dimensional sample entropy ($SampEn_{2D}$) [1] was developed, but

☆ 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. This work was also supported by both FCT and the ESF (European Social Fund) under the scholarship UI/BD/152802/2022. The authors would also like to thank to the Laboratory of Advanced Computing (LCA) at the University of Coimbra for providing computing resources that have contributed to the research results reported within this paper (https://www.uc.pt/lca).

* Corresponding author at: LIBPhys, Department of Physics, University of Coimbra, Coimbra, P-3004 516, Portugal.
E-mail address: andreia.gaudencio@fis.uc.pt (A.S. Gaudêncio).

https://doi.org/10.1016/j.cmpb.2023.107855
Received 27 March 2023; Received in revised form 1 October 2023; Accepted 8 October 2023
Available online 12 October 2023
0169-2607/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

its values may still be undefined or unreliable for small-sized images [1].

To overcome the limitation of undefined SampEn_{2D} values, two-dimensional fuzzy entropy (FuzEn_{2D}) was developed [9]. FuzEn_{2D} has already been used to assist in the diagnosis of pseudoxanthoma elasticum, a skin condition disorder [10], and to assess skin lesions, like melanoma, using colored dermoscopic images [8,12]. To address the limitation of undefined SampEn_{2D} values, two-dimensional permutation entropy (PerEn_{2D}) [6] was also developed based on two-dimensional permutation patterns and Shannon entropy. PerEn_{2D} is relatively simple algorithmically though sensitive to noise [2]. Nevertheless, it has been used to study mammary tissue density [6], synthetic textures [11], breast histopathologic images [11], and chest X-rays images [13].

To overcome the limitation of PerEn_{2D} in estimating entropy of noisy data and the shortcoming of undefined SampEn_{2D} values, Azami et al. [4] proposed a new method called bidimensional distribution entropy (DistEn_{2D}) based on the distance between patterns and their distribution through a histogram. DistEn_{2D} was found to be less sensitive to noise and less dependent on the setting parameters, demonstrating good performance even for small-sized images [4]. However, based on the DistEn algorithm, new data created simply by random permutations of an original image (shuffling image) have DistEn_{2D} values close to that for the original image. For instance, if the elements of an image are sorted, its DistEn_{2D} value is not changed considerably. However, as expected theoretically and intuitively, sorting leads to a lower entropy value (less irregularity). To alleviate this shortcoming, two-dimensional dispersion entropy (DispEn_{2D}) [7] was proposed based on mapping pixels within an image into different c classes, leading to a reliable, fast, and stable method [7]. DispEn_{2D} was applied to synthetic and real medical datasets and was found to be more stable, faster, and without undefined values compared to SampEn_{2D} [7].

Recently, to address the sensitivity of one-dimensional entropy to their parameters, the concept of ensemble entropy was introduced based on generating stable and low bias data patterns for entropy estimation [16]. Information theory-based entropy and its ensemble form have the potential to bridge various disciplines, including biology, biomedicine, chemistry, and quantum mechanics [17–20]. For instance, the concept of information entropy was employed to describe the structural (topological) complexity of chemical compounds and reactions [19,20]. While the conventional approach involves summing individual entropies, often leading to counterintuitive outcomes, a novel method has emerged for cases where multiple molecules form an ensemble. This method treats molecular ensembles as holistic entities, thereby enhancing accuracy in estimating complexity changes during reactions [19,20].

Ensemble entropy is based on using a range of parameter settings, in opposition to other methods that use a single set of parameters. As an extension of the one-dimensional ensemble entropy algorithms already developed, we propose, for the first time, two-dimensional ensemble sample entropy (EnsSampEn_{2D}), two-dimensional ensemble permutation entropy (EnsPerEn_{2D}), and two-dimensional ensemble dispersion entropy (EnsDispEn_{2D}). Furthermore, we propose the use of the ensemble concept for entropy algorithms that were never tested in an ensemble fashion, namely the two-dimensional ensemble fuzzy entropy, with multiple embedding dimensions m (EnsFuzEn^{M}_{2D}) and with multiple tolerance r (EnsFuzEn^{R}_{2D}), and two-dimensional ensemble distribution entropy (EnsDistEn_{2D}).

Herein, we apply these ensemble entropy techniques to differentiate pulmonary emphysema. Pulmonary emphysema occurs when the destruction of the lung parenchyma is verified [21,22]. This condition is derived from chronic obstructive pulmonary disease (COPD) where the patient's airflow is limited and is not fully reversible. Based on the lobular anatomy, pulmonary emphysema can be classified as centrilobular emphysema (CLE), panlobular emphysema (PLE), and paraseptal emphysema (PSE) [21]. In this study, three types of pulmonary tissue were analyzed: normal, CLE, and PSE. Computed tomography (CT) is established as an *in-vivo* method for emphysema analysis as it can be useful for both diagnosing and quantifying pulmonary emphysema, providing morphological information [22,23]. When using high-resolution CT, emphysema is identified by the presence of low attenuation areas contrasting with the surrounding parenchyma of normal attenuation values [22].

The remaining paper is divided as follows: section 2 describes the original and ensemble entropy methods, and details the synthetic data and biomedical dataset used; section 3 shows and discusses the results obtained for the ensemble and original entropy techniques. Finally, section 4 recapitulates the main results, and proposes future works.

2. Methods and materials

Based on the ensemble approaches recently published [16], we propose the two-dimensional extensions of ensemble dispersion entropy, ensemble permutation entropy, and ensemble sample entropy for two-dimensional data (respectively named EnsDispEn_{2D}, EnsPerEn_{2D}, and EnsSampEn_{2D}). Moreover, we also introduce the use of two-dimensional ensemble distribution entropy (EnsDistEn_{2D}), and two forms of two-dimensional ensemble fuzzy entropy (EnsFuzEn_{2D}).

First, we describe the ensemble techniques and summarize the corresponding original entropy metrics. Later, we describe the synthetic data for validation tests and the biomedical dataset used herein. Finally, we present the parameter settings used for both the synthetic and biomedical datasets. The MATLAB® code used in this paper is available at https://github.com/HamedAzami.

2.1. Two-dimensional ensemble dispersion entropy

Two-dimensional dispersion entropy (DispEn_{2D}) [7] is based on mapping pixel values into different classes. The simplest mapping procedure is the linear mapping but other techniques can be used like normal cumulative distribution function (NCDF), tansig, logsig, and sorting mappings. Therefore, DispEn_{2D} has three main parameters to be set: the mapping procedure, the embedding size, m-value, of the templates, and the number of classes C for pixel mapping. For more details regarding DispEn_{2D} please see Ref. [7].

Herein, we describe the methods behind EnsDispEn_{2D}. This low bias technique is based on the simultaneous use of the five different mappings mentioned above: linear, NCDF, tansig, logsig, and sorting. For an image $\{\mathbf{X}_{i,j}\}$, the elements – with $1 \le i \le H$ and $1 \le j \le W$ (H-height, W-width) – are mapped into integers from 1 to C classes according to the mapping procedure. This results in five mapped image versions: linear ($\boldsymbol{\eta}^c$), NCDF ($\boldsymbol{\xi}^c$), tansig ($\boldsymbol{\kappa}^c$), logsig (λ^c), and sorting ($\boldsymbol{\omega}^c$), with the same size as the original image.

Then, based on the mapped versions, templates are obtained according to the desired embedding parameter m. The templates for $\boldsymbol{\eta}^c$ are defined in equation (1). For the remaining mapped versions, the templates $\boldsymbol{\xi}^{m,c}_{i,j}$, $\boldsymbol{\kappa}^{m,c}_{i,j}$, $\boldsymbol{\lambda}^{m,c}_{i,j}$, and $\boldsymbol{\omega}^{m,c}_{i,j}$ are obtained similarly. The number of possible templates N_t obtained for each mapping is $N_t = (H-m)\times(W-m)$.

$$\boldsymbol{\eta}^{m,c}_{i,j} = \begin{bmatrix} \eta_{i,j} & \cdots & \eta_{i,j+m-1} \\ \cdots & \cdots & \cdots \\ \eta_{i+m-1,j} & \cdots & \eta_{i+m-1,j+m-1} \end{bmatrix}. \tag{1}$$

Afterwards, the templates for each mapping are vectorized into dispersion patterns $\pi^{m,c}_{i,j}$. In the next step, we obtain the probability associated with each dispersion pattern by evaluating all the mappings at once. Let $\Pi^{m,c}$ be the set composed by all the unique dispersion patterns in the mapped images. Then, $C(\Pi^{m,c}_k)$ corresponds to the number of occurrences of $\Pi^{m,c}_k$ on the dispersion patterns extracted from the five image mapped versions. The probability of each occurrence is then computed as:

$$p(\Pi^{m,c}_k) = \frac{C(\Pi^{m,c}_k)}{5N_t}. \tag{2}$$

Finally, EnsDispEn_{2D} is defined as:

$$\text{EnsDispEn}_{2D}(\mathbf{X}, m, C) = -\frac{1}{\log(N_d)} \sum_{1}^{K} p(\Pi_k) \times \log(p(\Pi_k)) , \quad (3)$$

where $N_d = C^{m \times m}$ is the number of possible dispersion patterns.

2.2. Two-dimensional ensemble distribution entropy

Two-dimensional distribution entropy (DistEn_{2D}) [4] is based on the empirical probability density function (ePDF), estimated through a histogram [4]. DistEn_{2D} depends on two parameters: the number of bins (B), used in the histogram, and the embedding dimension (m). DistEn_{2D} is detailed in Ref. [4].

EnsDistEn_{2D} is based on the mean of multiple DistEn_{2D} values by using several m-values to obtain the templates, *i.e.*, $[m_1, \ldots, m_K]$. To the best of our knowledge, an ensemble distribution entropy approach has not yet been proposed either for one- or two-dimensional data.

Let's consider an image $\mathbf{X}$ of dimensions $H \times W$. First, the templates $\mathbf{X}_{i,j}^{m_k}$ for a specific m_k-value are obtained using equation (4). These squared templates have $m_k \times m_k$ points. Similarly to DispEn_{2D}, the total number of templates is $N_t = (H - m_k) \times (W - m_k)$.

$$\mathbf{X}_{i,j}^{m_k} = \begin{bmatrix} X_{i,j} & \cdots & X_{i,j+m_k-1} \\ \cdots & \cdots & \cdots \\ X_{i+m_k-1,j} & \cdots & X_{i+m_k-1,j+m_k-1} \end{bmatrix} . \quad (4)$$

Afterwards, the distance between templates $\mathbf{X}_{i,j}^{m_k}$ and $\mathbf{X}_{a,b}^{m_k}$ is obtained using the following definition:

$$d_{i,j,a,b}^{m_k} = \max |\mathbf{X}_{i,j}^{m_k} - \mathbf{X}_{a,b}^{m_k}| , \quad (5)$$

with $1 \le a \le H - m_k$, $1 \le b \le W - m_k$, and $(i,j) \ne (a,b)$. The histogram of the distance values is determined using B bins, and the probability of each bin p_b, $b = 1, \ldots, B$, is used to compute the entropy value according to Shannon's definition:

$$\text{DistEn}_{2D}(\mathbf{X}, m_k) = -\frac{\sum_{b=1}^{B} p_b \times \log_2(p_b)}{\log_2(B)} . \quad (6)$$

EnsDistEn_{2D} is obtained by using different m_k-values as follows:

$$\text{EnsDistEn}_{2D}(\mathbf{X}, M) = \frac{1}{K} \sum_{k=1}^{K} \text{DistEn}_{2D}(\mathbf{X}, m_k) . \quad (7)$$

2.3. Two-dimensional ensemble permutation entropy

Permutation entropy (PerEn) [24] was introduced as a robust, simple, and computationally low-cost entropy algorithm for signal analysis. Later on, two-dimensional permutation entropy (PerEn_{2D}) was proposed by Morel et al. [11]. This method relies on the determination of the probability associated with each permutation pattern. PerEn was designed to avoid the need for multi-parameter tuning [24]. This characteristic is extended to PerEn_{2D} such that only the parameter m has to be set.

EnsPerEn_{2D} is defined as the average of PerEn_{2D} for multiple m-values, *i.e.*, $M = [m_1, \ldots, m_K]$, with $m_1 = 2$ and K the number of models. Based on PerEn recommendation [24], the maximum m-value should be defined based on the limit $((m_K + 1)^2)! \le (W \times H)$.

As the methods before, templates are extracted from the image with a specific embedding dimension ($\mathbf{X}_{i,j}^{m_k}$). The maximum number of templates for each m_k-value is defined as $N_t = (H - m_k) \times (W - m_k)$. Then, each template is vectorized line by line. The intensities of the vectorized templates must be rearranged in ascending order to determine the corresponding permutation pattern $\pi_{i,j}^{m_k}$.

Let's consider that one of these vectors has the following intensities $[0.8, 0.1, 0.5, 0.9]$ and their positions are defined as $\{0, 1, 2, 3\}$. When rearranging the intensity values in ascending order, the vector becomes $[0.1, 0.5, 0.8, 0.9]$. Therefore, the initial positions are sorted as $\{1, 2, 0, 3\}$. The corresponding permutation pattern $\pi_{i,j}^2$ would be $[1, 2, 0, 3]$ [11]. The vector has m_k^2 points, and there are $N_p^{m_k} = (m_k \times m_k)!$ possible different permutation patterns.

Considering Π^{m_k} the set of all unique permutation patterns extracted from the original image, and $C(\Pi_c^{m_k})$ the number of occurrences of $\Pi_c^{m_k}$ on the set of all the permutation patterns extracted from the image, we can compute the following probability.

$$p(\Pi_c^{m_k}) = \frac{C(\Pi_c^{m_k})}{N_t} . \quad (8)$$

For each m_k-value within M and based on Shannon's definition, the PerEn_{2D} is defined as:

$$\text{PerEn}_{2D}(\mathbf{X}, m_k) = -\sum_{c=1}^{C} p(\Pi_c^{m_k}) \times \ln(p(\Pi_c^{m_k})) . \quad (9)$$

Then, we define normalized PerEn_{2D} as follows:

$$\text{NPerEn}_{2D}(\mathbf{X}, m_k) = \frac{\text{PerEn}_{2D}(\mathbf{X}, m_k)}{\ln(N_p^{m_k})} . \quad (10)$$

Finally, EnsPerEn_{2D} can be defined as:

$$\text{EnsPerEn}_{2D}(\mathbf{X}, M) = \frac{1}{K} \sum_{k=1}^{K} \text{NPerEn}_{2D}(\mathbf{X}, m_k) . \quad (11)$$

2.4. Two-dimensional ensemble sample entropy

Two-dimensional sample entropy (SampEn_{2D}) [1] relies on verifying if squared templates of $(m \times m)$-points remain similar for $(m+1) \times (m+1)$. SampEn_{2D} was shown to be robust and to be able to differentiate distinct textures but it can be inaccurate for small images [4].

SampEn_{2D} depends on two main parameters, the size of templates (m) and the tolerance value (r) used to define similarity between templates. The r balances the estimation of entropy. When the tolerance is small, templates are more difficult to be considered similar, which can lead to inaccurate (or even undefined) entropy values. In contrast, when r is large, many templates are defined as similar and the details of the image can be lost, associating biased high entropy values to the images [1]. Therefore, EnsSampEn_{2D} is based on using different r-values, where $R = [r_1, \ldots, r_K]$, with K corresponding to the number of tolerance values used, hence, the number of considered models.

Once again, we consider an image $\mathbf{X}$ of dimensions $H \times W$, and a fixed embedding dimension m. In addition, we define the tolerance parameter as $r_k = p \times \text{SD}(\mathbf{X})$, where p is a percentage level, SD stands for standard deviation, and $1 \le k \le K$.

First, the templates are defined as the other methods ($X_{i,j}^m$). Similarly, there will be $N_t = (H - m) \times (W - m)$ possible templates within the image $\mathbf{X}$. Then, we can calculate the distance between templates $X_{i,j}^m$ and $X_{a,b}^m$ of the same embedding size as shown in equation (5).

Considering $C(d_{i,j,a,b}^m \le r_k)$ the number of times that the distance is less than or equal to the threshold r_k, we can defined the probability as:

$$\phi^m(r_k) = \frac{C(d_{i,j,a,b}^m \le r_k)}{N_t} , \quad (12)$$

where $1 \le k \le K$. Subsequently, the average $\Phi^m(r_k)$ is determined as follows:

$$\Phi^m(r_k) = \frac{\sum_1^K \phi^m(r_k)}{N_t} . \quad (13)$$

Thereupon, the average of all $\Phi^m(r_k)$ is obtained (equation (14)), and after repeating the previous steps for templates of size $(m+1) \times (m+1)$, EnsSampEn_{2D} is defined as in equation (15).

$$\Phi^m = \frac{\sum_1^{N_r} \Phi^m(r_k)}{N_r} . \quad (14)$$

$$\text{EnsSampEn}_{2D}(\mathbf{X}, m, R) = \ln \frac{\Phi^m}{\Phi^{m+1}} . \quad (15)$$

2.5. Two-dimensional ensemble fuzzy entropy

Hilal et al. [9,10] proposed the use of two-dimensional fuzzy entropy (FuzEn_{2D}) based on the original concept of fuzzy entropy (FuzEn) [25]. FuzEn_{2D} was used to characterize both the local and global characteristics of an image, in contrast with FuzEn that analyzed only local properties of the signal. FuzEn_{2D} is an entropy metric in which similarity between templates is obtained continuously. In addition, FuzEn_{2D} remains consistent for short data, when compared to SampEn_{2D}, although it has a higher computational cost. This metric is dependent on three parameters: the embedding dimension of the templates m, tolerance r, and fuzzy power n (usually set to $n = 2$) [9,10].

Herein, we describe the two derived ensemble approaches of FuzEn_{2D}: the two-dimensional ensemble fuzzy entropy with multiple tolerances (EnsFuzEn^{R}_{2D}) and the two-dimensional ensemble fuzzy entropy with multiple embedding dimensions (EnsFuzEn^{M}_{2D}).

EnsFuzEn^{R}_{2D} will first be described. For an image $\mathbf{X}$ of dimensions $H \times W$, a fixed embedding dimension m, and multiple r_k-values $R = [r_1, \ldots, r_K]$ with K being the number of models, the templates are defined as in the other cases with $N_t = (H - m) \times (W - m)$ possible templates within the image $\mathbf{X}$.

Afterwards, the distance $d^{m}_{i,j,a,b}$ between the templates $X^{m}_{i,j}$ and $X^{m}_{a,b}$ for $1 \le i, a \le H - m$, $1 \le j, b \le W - m$, and $(i, j) \neq (a, b)$ is obtained as in equation (5) ($m_k = m$).

Next, the similarity degree $D^{m}_{i,j,a,b}(r_k)$ is obtained as follows:

$$D^{m}_{i,j,a,b}(r_k) = \exp\left(-\left(\frac{d^{m}_{i,j,a,b}}{r_k}\right)^{n}\right), \tag{16}$$

where n is the fuzzy power. Then, the average of the similarity degree values between templates is defined as:

$$\Phi^{m}(r_k) = \frac{\sum_{1}^{N_t} D^{m}_{i,j,a,b}(r_k)}{{N_t}^2}. \tag{17}$$

The average of $\Phi^{m}(r_k)$ is obtained as:

$$\Phi^{m} = \frac{\sum_{k=1}^{K} \Phi^{m}(r_k)}{K}. \tag{18}$$

Subsequently, the previous steps are repeated for templates of size $(m + 1) \times (m + 1)$, and EnsFuzEn^{R}_{2D} is defined as:

$$\text{EnsFuzEn}^{R}_{2D}(\mathbf{X}, m, R, n) = \ln \frac{\Phi^{m}}{\Phi^{m+1}}. \tag{19}$$

For EnsFuzEn^{M}_{2D}, we now consider a fixed tolerance and a variable m-value, for a certain number of models K where $M = [m_1, \ldots, m_K]$. The process is similar to the EnsFuzEn^{R}_{2D} but the average similarity degree for the set of m-values is defined as:

$$\Phi^{M}(r) = \frac{\sum_{k=1}^{K} \Phi^{m_k}(r)}{K}. \tag{20}$$

Finally, the EnsFuzEn^{M}_{2D} can be computed by using a new set of m-values where $M' = [m_1 + 1, \ldots, m_K + 1]$.

$$\text{EnsFuzEn}^{M}_{2D}(\mathbf{X}, M, r, n) = \ln \frac{\Phi^{M}(r)}{\Phi^{M'}(r)}. \tag{21}$$

2.6. Synthetic data

To evaluate if two-dimensional ensemble algorithms are able to detect the degree of randomness, $\text{MIX}_{2D}(p)$ images [1] were used. $\text{MIX}_{2D}(p)$ images are defined by sinusoidal regular behavior and a random distribution (irregular behavior). The level of irregularity is determined by the parameter p. Fig. 1 shows 64×64 pixels-sized images defined by $\text{MIX}_{2D}(p)$ processes.

$\text{MIX}_{2D}(p)$ images of 64×64 pixels were used to show the behavior of entropy metrics upon increasing irregularity. For this test, we varied the parameter p between 0.0 and 1.0 with a step of 0.05 and calculated

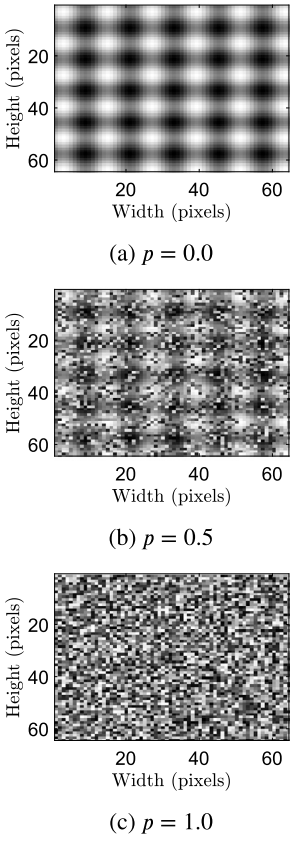


Fig. 1. Examples of $\text{MIX}_{2D}(p)$ images with 64×64 pixels.

their entropy values. This test was repeated 10 times for each p value. In addition, to assess the stability of these metrics in comparison with their corresponding original entropy values, we applied the ensemble and original entropy methods to $\text{MIX}_{2D}(p)$ images of 32×32, 64×64, and 128×128 pixels with fixed p to $p = 0.5$. For each size, we used 10 randomly generated images.

2.7. Biomedical application

The biomedical dataset used in this study is publicly available at https://lauge-soerensen.github.io/emphysema-database [26]. This public dataset has 168 patches of high-resolution computed tomography scans from 39 subjects. These subjects were either never-smokers individuals (9), smokers (10), or smokers diagnosed with chronic obstructive pulmonary disease (COPD). The patches were manually segmented by an experienced radiologist and experienced pulmonologist. These patches were divided into 3 categories regarding their leading pattern: normal tissue (NT), centrilobular emphysema (CLE), and paraseptal emphysema (PSE).

CLE is usually distributed in the upper lobe of the lung or in the superior segment of the lower lobe. When CLE is present, an enlargement of the centriacinar airspace is verified mainly in proximal respiratory bronchioles. CLE is the type of emphysema most associated with smoking habits [22]. On the other hand, PSE leads to an enlarged airspace in the peripheric acini and is usually limited in extent and present along the dorsal surface of the upper lung. Reduced alveoli is also verified [21].

The patches have a size of 61×61 pixels. The ensemble entropy algorithms and their corresponding original forms were used to characterize the texture of these patterns. Afterwards, the normality of each group was assessed with the Shapiro Wilk test for a significance level of $\alpha = 0.05$. Moreover, we performed a Kruskal-Wallis analysis regarding the statistical significance between the three groups and we tested

the statistical difference between each groups' pair (NT vs CLE, CLE vs PSE, and NT vs PSE) using Tukey's honestly test.

Finally, through a multiclass KNN classifier we assessed the ability of each entropy metric to classify NT texture, CLE texture, and PSE texture. The classification metrics chosen are: area under the curve (AUC) for the three groups (NT, CLE, and PSE), accuracy, sensitivity (macro and micro), precision (macro and micro), and F1-score (macro and micro). Sensitivity and precision macro were obtained by averaging the sensitivity and precision obtained for each class. F1-score macro was obtained with these macro values. Sensitivity and precision micro were obtained with the true positives, the false negatives, and false positives values of each class. F1-score micro was then obtained with sensitivity and precision micro.

2.8. Parameters settings

First, the effects of parameter settings and the performance of ensemble entropy approaches was compared to the original entropy methods in the validation tests. For this purpose the methods were tested with the following parameter settings:

- EnsDispEn_{2D}:
  $n_c = 5$ and five mappings;
- EnsDistEn_{2D}:
  $M = [2, 3]$ and $B = 1024$;
- EnsSampEn_{2D}:
  $R = [0.1, 0.15, 0.2, 0.25, 0.3] \times \text{SD}(\mathbf{X})$ and $m = 2$;
- EnsFuzEn^{M}_{2D}: $M = [2, 3, 4]$, $r = 0.2 \times \text{SD}(\mathbf{X})$, and $n = 2$;
- EnsFuzEn^{R}_{2D}:
  $R = [0.1, 0.15, 0.2, 0.25, 0.3] \times \text{SD}(\mathbf{X})$, $m = 2$, and $n = 2$.

The ensemble entropy algorithms were also tested, on the biomedical dataset, with the parameters above.

For the original entropy algorithms, the parameters were fixed to $m = 2$ for DistEn_{2D}, FuzEn_{2D}, PerEn_{2D}, and SampEn_{2D}, $r = 0.2 \times \text{SD}(\mathbf{X})$ for FuzEn_{2D}, and SampEn_{2D}, $n_c = 6$ and NCDF mapping for DispEn_{2D}, $B = 1024$ for DistEn_{2D}, and $n = 2$ for FuzEn_{2D}.

For the statistical tests, we used Matlab's® *kruskalwallis* function for the *Kruskal Wallis* test, and the *multcompare* function for *Tukey's honestly test*. In addition, we also used the function *fitcknn* for the k-nearest neighbor (KNN) classifier with automatic hyperparameters optimization option.

3. Results and discussion

3.1. Validation tests

First, we show the variation of entropy for ensemble and original entropy techniques according to different irregularity p levels of $\text{MIX}_{2D}(p)$ images. As the irregularity of images increases, we expect to observe higher entropy values.

In Fig. 2, comparing both DispEn_{2D} (green) approach and the ensemble version, both methods show increasing entropy with the increasing of image irregularity. DispEn_{2D} is slightly better than EnsDispEn_{2D} when discriminating different irregular images because it shows a higher entropy variation ($45.9 \pm 0.2\%$ vs. $41.5 \pm 0.1\%$). DispEn_{2D} has better differentiation than EnsDispEn_{2D} for more irregular images as they have higher frequency components. Based on this, DispEn_{2D} might perform better for higher frequency based images than EnsDispEn_{2D}. For more regular images composed by lower frequencies, both algorithms should perform similarly.

For DistEn_{2D} and EnsDistEn_{2D} methods (see Fig. 3), we observe that DistEn_{2D} achieves a variation of entropy of $53.0 \pm 0.0\%$. EnsDistEn_{2D} shows a variation of $51.6 \pm 0.0\%$ and a curve profile very similar to DistEn_{2D}. Moreover, DistEn_{2D} and EnsDistEn_{2D} have an increase of around 2% between $p = 0.95$ and $p = 1.0$ which occurs due to the nature of the method. Both approaches are determined by using the histogram of the distances between templates.

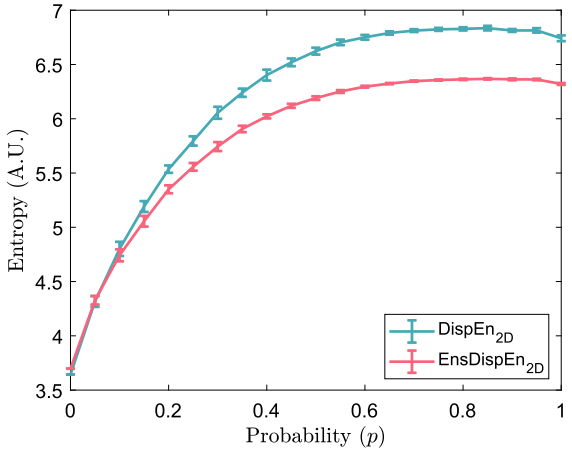


Fig. 2. DispEn_{2D} and EnsDispEn_{2D} entropy values for MIX_{2D}(p) images of 64×64 pixels and $0 \leq p \leq 1$.

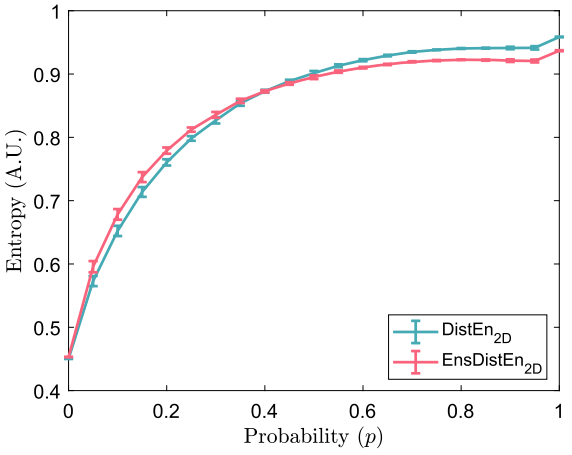


Fig. 3. DistEn_{2D} and EnsDistEn_{2D} entropy values for MIX_{2D}(p) images of 64×64 pixels.

Fig. 4 shows a similar variation for EnsPerEn_{2D} and PerEn_{2D}. EnsPerEn_{2D} shows a relative increase of entropy of $36.5 \pm 0.0\%$ which is better compared to the $34.6 \pm 0.0\%$ increase of PerEn_{2D}. Both measures show stable curves for $p \geq 0.4$, with little or none increase of entropy after that.

For SampEn_{2D} (Fig. 5), the relative increase observed is $89.0 \pm 1.2\%$. However, this value cannot be compared directly with the remaining measures since entropy for $0.55 \leq p \leq 1.0$ was undefined. The same happens for EnsSampEn_{2D} but, in this case, the entropy was undefined for $p \geq 0.6$, with an increase of $89.2 \pm 0.9\%$. Both curves completely overlap leading to the conclusion that both parameter settings have similar behaviors in entropy evaluation. However, EnsSampEn_{2D} is more stable since it can determine entropy for $p \leq 0.6$. For this size (64×64 pixels) and given the irregularity ($p \geq 0.6$ and $p \geq 0.5$), EnsSampEn_{2D} and SampEn_{2D} are not able to find similar templates within the image, leading to undefined entropy values.

In addition, the performances of FuzEn_{2D} and EnsFuzEn^{M}_{2D} are similar for every value of p. Their entropy variation is $82.9 \pm 0.1\%$ and $87.7 \pm 0.1\%$, respectively. EnsFuzEn^{R}_{2D} shows a lower relative entropy increase of $81.5 \pm 0.1\%$.

Overall, EnsFuzEn^{M}_{2D} (see Fig. 6) shows the best increase of entropy (87.7%) between the most regular ($p = 0$) and the most irregular ($p = 1$) images. The remaining methods and their different parameter settings

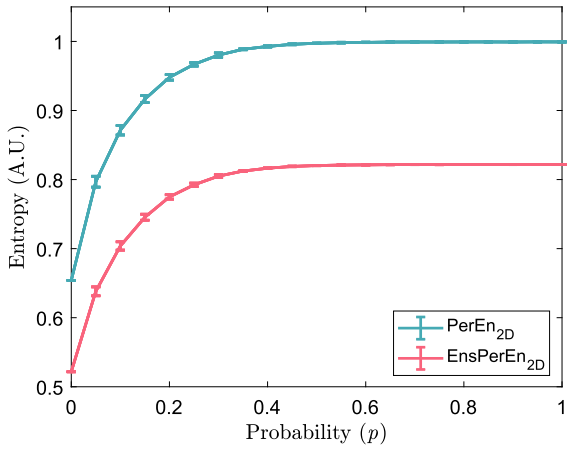


Fig. 4. PerEn$_{2D}$ and EnsPerEn$_{2D}$ entropy values for MIX$_{2D}$(p) images of 64×64 pixels.

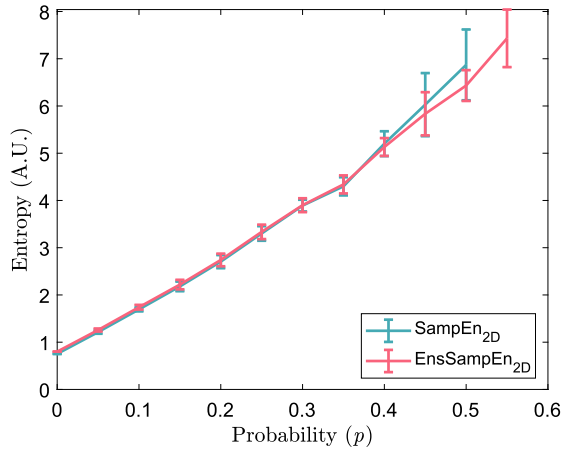


Fig. 5. SampEn$_{2D}$ and EnsSampEn$_{2D}$ entropy values for MIX$_{2D}$(p) images of 64×64 pixels.

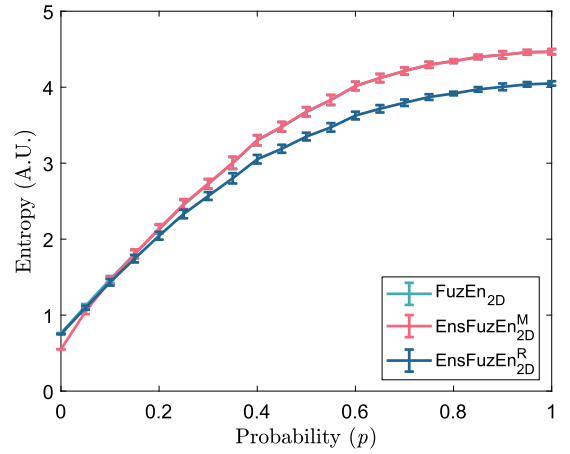


Fig. 6. FuzEn$_{2D}$, EnsFuzEn$^{M}_{2D}$, and EnsFuzEn$^{R}_{2D}$ entropy values for MIX$_{2D}$(p) images of 64×64 pixels.

have a lower variation of entropy values when considering images with increasing irregularity such as MIX$_{2D}(p)$ images.

Summarizing the previous results, we observe that:

- DispEn$_{2D}$ and DistEn$_{2D}$ show very similar behaviors compared to their corresponding ensemble techniques;
- EnsPerEn$_{2D}$ and PerEn$_{2D}$ have a similar irregularity discrimination;
- EnsSampEn$_{2D}$ is more stable than the original algorithm, being able to determine entropy for $0.0 \leq p \leq 0.6$;
- EnsFuzEn$^{M}_{2D}$ has a higher entropy variation, followed by FuzEn$^{M}_{2D}$, and finally, EnsFuzEn$^{R}_{2D}$.

Table 1
Mean ensemble and original entropy values for MIX$_{2D}(p = 0.5)$ images of 32×32, 64×64, and 128×128 pixels.

Entropy algorithms	32×32	64×64	128×128
DispEn$_{2D}$	6.13 ± 0.04	6.63 ± 0.02	6.78 ± 0.01
EnsDispEn$_{2D}$	6.09 ± 0.03	6.19 ± 0.01	6.21 ± 0.01
DistEn$_{2D}$	0.90 ± 0.00	0.90 ± 0.00	0.90 ± 0.00
EnsDistEn$_{2D}$	0.89 ± 0.00	0.89 ± 0.00	0.90 ± 0.00
PerEn$_{2D}$	0.99 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
EnsPerEn$_{2D}$	0.76 ± 0.00	0.82 ± 0.00	0.87 ± 0.00
SampEn$_{2D}$	-	6.57 ± 0.77	6.82 ± 0.22
EnsSampEn$_{2D}$	-	6.44 ± 0.37	6.58 ± 0.13
FuzEn$_{2D}$	3.68 ± 0.08	3.68 ± 0.07	3.70 ± 0.02
EnsFuzEn$^{M}_{2D}$	3.69 ± 0.08	3.68 ± 0.10	3.70 ± 0.02
EnsFuzEn$^{R}_{2D}$ (2)	3.36 ± 0.07	3.35 ± 0.06	3.37 ± 0.01

Table 2
Coefficient of variation (CV) for the values obtained by the ensemble and original entropy methods using MIX$_{2D}(p = 0.5)$ images of 32×32, 64×64, and 128×128 pixels.

Entropy algorithms	32×32	64×64	128×128
DispEn$_{2D}$	0.006	0.003	0.002
EnsDispEn$_{2D}$	0.005	0.002	0.001
DistEn$_{2D}$	0.006	0.003	0.001
EnsDistEn$_{2D}$	0.005	0.002	0.001
PerEn$_{2D}$	0.002	0.001	0.000
EnsPerEn$_{2D}$	0.001	0.000	0.000
SampEn$_{2D}$	-	0.117	0.033
EnsSampEn$_{2D}$	-	0.057	0.021
FuzEn$_{2D}$	0.022	0.018	0.004
EnsFuzEn$^{M}_{2D}$	0.022	0.018	0.004
EnsFuzEn$^{R}_{2D}$	0.021	0.017	0.004

Table 1 shows the mean entropy values of MIX$_{2D}(p = 0.5)$ images with different sizes. This allows to observe how different entropy algorithms remain stable based on the image's size. First, we can observe that DistEn$_{2D}$, EnsDistEn$_{2D}$, PerEn$_{2D}$, and EnsPerEn$_{2D}$ show the lowest SD values, indicating higher precision. Nonetheless, DispEn$_{2D}$, EnsDispEn$_{2D}$, FuzEn$_{2D}$, EnsFuzEn$^{M}_{2D}$, and EnsFuzEn$^{R}_{2D}$ also show very low SD values. Table 2 reflects the coefficient of variation (CV = SD/μ) of these algorithms for the different image sizes.

The ensemble version of DispEn$_{2D}$ has an improved consistency (Table 1) when compared with DispEn$_{2D}$. The original algorithm has a relative difference (RD) of about 9.6% when using images of 128×128 and 32×32 pixels, and EnsDispEn$_{2D}$ has a RD of 2.1%. Table 2 also shows that EnsDispEn$_{2D}$ is more stable since its corresponding CV is lower.

RD between the smallest and the largest image was 0.0% for DistEn$_{2D}$ and 1.1% for EnsDistEn$_{2D}$. However, in terms of CV, it is possible to observe that EnsDistEn$_{2D}$ is more stable (see Table 2).

On the other hand, PerEn$_{2D}$ and EnsPerEn$_{2D}$ are both precise but PerEn$_{2D}$ only varies around 0.5% between the smallest and the largest images, in opposition to 12.8% for EnsPerEn$_{2D}$.

SampEn$_{2D}$ and EnsSampEn$_{2D}$ are the least precise entropy algorithms and the least reliable ones for small images analysis. These algorithms were not able to determine entropy for images of 32×32

Table 3
Kruskal-Wallis test for the ensemble and original entropy values of normal tissue, centrilobular emphysema, and paraseptal emphysema. The symbol * indicates statistical significance for $P < 0.01$.

Metrics	H(2)	P-value
$DispEn_{2D}$	77.81	1.27E−17*
$EnsDispEn_{2D}$	64.46	1.01E−14*
$DistEn_{2D}$	81.60	1.91E−18*
$EnsDistEn_{2D}$	75.72	3.61E−17*
$PerEn_{2D}$	15.20	5.00E−04*
$EnsPerEn_{2D}$	21.36	2.30E−05*
$SampEn_{2D}$	81.81	1.72E−18*
$EnsSampEn_{2D}$	76.22	2.81E−17*
$FuzEn_{2D}$	70.46	5.02E−16*
$EnsFuzEn^{M}_{2D}$	61.47	4.48E−14*
$EnsFuzEn^{R}_{2D}$	71.86	2.50E−16*

pixels. Nevertheless, if we compare the entropy between images of 64×64 and 128×128 pixels, $EnsSampEn_{2D}$ methods are more consistent than the original method. $SampEn_{2D}$ and $EnsSampEn_{2D}$ show a RD of 3.8% and 2.2%, respectively. Besides, when analyzing their corresponding CVs, $SampEn_{2D}$ has the highest CV reported (including when comparing with other entropy techniques), and, although it decreases with size, this algorithm is the least stable. $EnsSampEn_{2D}$ is able to improve the algorithm stability when compared to $SampEn_{2D}$.

Finally, $FuzEn_{2D}$ is consistent with a RD of 0.3% (between images of 32×32 and 128×128 pixels). $EnsFuzEn^{R}_{2D}$ shows a RD of 0.0% and $EnsFuzEn^{M}_{2D}$ shows a RD of 0.2%, both below $FuzEn_{2D}$ RD (Table 1). In fact, $EnsFuzEn^{R}_{2D}$ and $EnsFuzEn^{M}_{2D}$ are more stable than $FuzEn_{2D}$ when analyzing their corresponding CVs (Table 2).

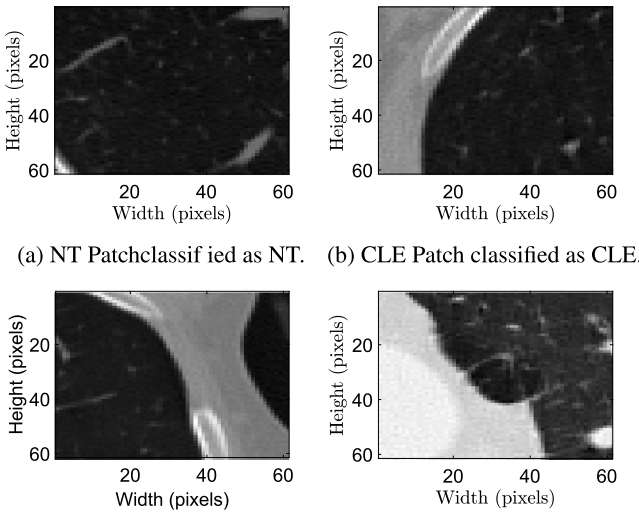
3.2. Biomedical application

As mentioned above, this study analyzes the texture of images from three different pulmonary conditions: normal tissue (NT), centrilobular emphysema (CLE), and paraseptal emphysema (PSE). Both original and ensemble entropy methods with the aforementioned parameter settings were used to characterize these textures.

We evaluated the ability of the ensemble and original methods to differentiate the three groups based on the Kruskal-Wallis test. The proposed methods are compared in Table 3 (99% confidence level). The results show that $EnsPerEn_{2D}$ and $EnsFuzEn^{R}_{2D}$ give lower P-values than their corresponding original entropy algorithms.

Furthermore, when comparing two groups with each other (NT vs. CLE, NT vs. PSE, and CLE vs. PSE) through the Tukey's honestly test (Table 4), only $EnsDistEn_{2D}$, $DistEn_{2D}$, $EnsPerEn_{2D}$, and $PerEn_{2D}$ are not able to statistically differentiate all the three possible comparisons for $P < 0.05$. Differences in $EnsDistEn_{2D}$, $DistEn_{2D}$, and $EnsPerEn_{2D}$ are only statistically significant for the cases NT vs. PSE and CLE vs. PSE. $PerEn_{2D}$ is only able to statistically differentiate CLE and PSE groups. On the other hand, the remaining ensemble and original entropy techniques are able to statistically differentiate the three groups between each other.

In addition, Table 4 shows the effect size (d) between groups by using the Cohen's test [27]. The effect size allows us to quantify the experimental effect of the relationship between variables. When evaluating the entropy algorithms and the three possible comparisons, the $|d|$ values were lower when comparing the NT and CLE groups within the same entropy algorithm. Within the algorithms achieving statistical significance, $EnsFuzEn^{R}_{2D}$ shows the weakest difference, with $|d| = 0.18$. The remaining entropy techniques show relative medium effect size when comparing NT and CLE ($0.5 \leq |d| < 0.8$).



(a) NT Patchclassif ied as NT. (b) CLE Patch classified as CLE.

(c) NT Patch classified as CLE. (d) NT Patch classified as PSE.

Fig. 7. Examples of patches correctly (a-b) and incorrectly (c-d) labeled by the KNN classifier using $EnsFuzEn^{M}_{2D}$ values.

When comparing the NT and PSE groups, the effect size is relatively high with $|d| > 0.8$, indicating a strong difference between the entropy obtained for NT and that for PSE. $EnsDispEn_{2D}$ leads to the lowest $|d|$ while $EnsPerEn_{2D}$ leads to the highest one for the NT vs. PSE groups. In addition, between the CLE and PSE groups, the effect size is even higher ($|d| > 1.5$) for all the entropy metrics, indicating that the strongest difference is obtained between these two groups. When differentiating these last two groups, the obtained P-values are also lower compared to the P-values obtained for the other groups' comparisons.

When applying a KNN classifier to differentiate the three types of tissue, shown in Table 5 and Table 6, $EnsFuzEn^{M}_{2D}$ is the best ensemble method in terms of accuracy, precision (macro and micro), F1-score (macro and micro), sensitivity (macro and micro), AUC of NT, and AUC of PSE (see Table 5). When analyzing sensitivity micro, $EnsDistEn_{2D}$ is slightly better than $EnsFuzEn^{M}_{2D}$ with 88.3% of sensitivity. However, considering F1-score which balances the sensitivity and precision, $EnsFuzEn^{M}_{2D}$ is clearly the best method. The alternative $EnsFuzEn^{R}_{2D}$ gives the best AUC value for the classification of CLE. The ensemble method with the lowest accuracy is $EnsPerEn_{2D}$(69.8%).

Table 5 shows that $EnsFuzEn^{M}_{2D}$ has the best accuracy value within ensemble and original entropy methods. Overall, the ensemble methods show better accuracy than the original techniques except for $EnsSampEn_{2D}$ and $EnsFuzEn^{R}_{2D}$. However, $EnsFuzEn^{R}_{2D}$ is only slightly lower than $FuzEn_{2D}$ in terms of accuracy.

When comparing original and ensemble methods, $EnsDispEn_{2D}$ and $EnsFuzEn^{R}_{2D}$ have worse F1-scores than the corresponding original methods $DispEn_{2D}$ and $FuzEn_{2D}$, respectively. However, the remaining ensemble methods ($EnsDistEn_{2D}$, $EnsPerEn_{2D}$, $EnsSampEn_{2D}$, and $EnsFuzEn^{M}_{2D}$) all perform better than the original techniques, demonstrating their practical application. Moreover, in terms of accuracy, all of these ensemble methods (except for $EnsFuzEn^{R}_{2D}$) achieve higher accuracy values than the corresponding original entropy algorithms (see Table 6).

In Figs. 7a-7b, we showed four examples of patches that were correctly classified as normal tissue (NT) and CLE, respectively, using the best entropy method for classification, $EnsFuzEn^{M}_{2D}$, with a KNN classifier. On the other hand, Conversely, in Figs. 7c-7d, we can observe two instances of misclassified patches, where NT was incorrectly classified as CLE and PSE, respectively.

In Ref. [28], one of the latest works using the same database, the authors proposed two neural networks with data augmentation obtaining a total of 5520 CT images (512×512 pixels) to classify emphysema.

Table 4
Tukey's honestly test's P-value and Cohen's $|d|$ effect size for ensemble entropy techniques to compare normal tissue (NT), centrilobular emphysema (CLE), and paraseptal emphysema (PSE) groups. The symbol * indicates statistical significance for $P < 0.05$.

Metrics	NT vs CLE		NT vs PSE		CLE vs PSE	
	P-value	$\|d\|$	P-value	$\|d\|$	P-value	$\|d\|$
$DispEn_{2D}$	1.40E-04*	0.83	2.55E-06*	1.02	2.01E-18*	2.84
$EnsDispEn_{2D}$	1.38E-03*	0.65	8.87E-06*	0.81	3.48E-15*	2.45
$DistEn_{2D}$	3.74E-01	0.24	1.64E-12*	1.36	3.38E-16*	2.04
$EnsDistEn_{2D}$	4.10E-01	0.23	1.07E-11*	1.32	4.955E-15*	1.83
$PerEn_{2D}$	1.84E-01	0.34	6.88E-02	0.45	3.11E-04*	0.67
$EnsPerEn_{2D}$	9.76E-02	0.21	2.24E-02*	1.43	1.30E-05*	3.59
$SampEn_{2D}$	1.05E-03*	0.83	6.84E-08*	1.10	5.16E-19*	1.93
$EnsSampEn_{2D}$	7.59E-04*	0.70	5.59E-07*	1.32	7.66 E-18*	1.92
$FuzEn_{2D}$	1.59E-02*	0.71	5.20E-08*	1.31	6.67E-16*	1.81
$EnsFuzEn^{M}_{2D}$	3.16E-03*	0.73	7.94E-06*	0.93	2.00E-14*	2.70
$EnsFuzEn^{R}_{2D}$	1.44E-02*	0.18	3.86E-08*	1.30	3.21E-16*	1.60

Table 5
Multiclass classification using KNN classifier for ensemble entropy techniques to evaluate normal tissue, centrilobular emphysema, and paraseptal emphysema (* represents the best value).

Metrics (%)		$EnsDispEn_{2D}$	$EnsDistEn_{2D}$	$EnsPerEn_{2D}$	$EnsSampEn_{2D}$	$EnsFuzEn^{M}_{2D}$	$EnsFuzEn^{R}_{2D}$
	Accuracy	84.9	86.5	69.8	79.0	92.7*	81.0
	Sensitivity	55.6	72.2	71.9	70.0	86.5*	70.1
Macro	Precision	59.4	89.5	72.8	73.3	91.5*	81.9
	F1-score	57.4	80.0	72.3	71.6	88.9*	75.6
	Sensitivity	53.9	74.4	71.4	75.0	88.4*	73.9
Micro	Precision	84.9	86.5	69.8	79.0	92.7*	81.0
	F1-score	65.9	80.0	70.6	76.9	90.5*	77.3
	AUC of NT	50.1	62.2	58.7	47.4	72.4*	67.2
	AUC of CLE	80.9	69.8	72.7	61.1	80.9	85.3*
	AUC of PSE	78.0	89.8	71.8	76.1	93.7*	83.5

Table 6
Multiclass classification using KNN classifier for original entropy techniques to evaluate normal tissue, centrilobular emphysema, and paraseptal emphysema (* represents the best value).

Metrics (%)		$DispEn_{2D}$	$DistEn_{2D}$	$PerEn_{2D}$	$SampEn_{2D}$	$FuzEn_{2D}$
	Accuracy	73.2	81.6	56.8	89.2*	86.7
	Sensitivity	87.5	63.8	70.7	56.7	87.9*
Macro	Precision	77.1	76.9	63.9	63.2	86.6*
	F1-score	82.0	69.7	67.1	59.8	87.2*
	Sensitivity	90.9	75.6	75.0	68.8	88.6*
Micro	Precision	73.2	81.6	56.8	89.2*	86.7
	F1-score	81.1	78.5	64.6	77.7	87.6*
	AUC of NT	39.1	75.1*	57.1	65.3	70.9
	AUC of CLE	82.3	67.0	74.1	87.1*	77.1
	AUC of PSE	89.9	91.0*	56.2	87.1	88.5

When classifying each class, their best network, the enhanced multiscale residual network with data augmentation (EMS-ResNet-DA), obtained an accuracy over 95% for PSE. The remaining classes obtained accuracy values between 90 and 95%. Overall, the average accuracy was 94.6%. The sensitivity and precision values of each class were all around 95%. The F1-score (macro) was approximately 93% for CLE, around 95% for PSE, and 94% for NT [28]. However, it is important to mention that these results were obtained for a substantially higher number of data. Our results for our best model, $EnsFuzEn^{M}_{2D}$, are only slightly inferior using only 168 patches. Furthermore, our results are obtained based on a single ensemble-based entropy feature to classify emphysema patches retrieved from CT scans.

Furthermore, the pre-trained Visual Geometry Group with 16 layers depth (VGG16) model was also used to classify the 512×512 pixels-sized CT images [29] using the dataset of [26]. After pre-processing the images, they divided the 115 CT images into 2 classes: normal and emphysema subjects. In that work, the accuracy, precision, sensitivity, and F1-score were verified to be 88% [29]. In our case, when using only a single feature $EnsFuzEn^{M}_{2D}$, the sensitivity value of the classifier was slightly lower (86.5%). Isaac et al. [30] also proposed a model to classify the same emphysema dataset based on a competitive co-evolution model with feature selection, reporting an accuracy of 93.7%, precision of 90.61%, recall of 90.61%, and specificity of 95.3% [30].

On the other hand, Barnacles Mating-based Butterfly Optimization Algorithm hybrid classifier (BM-BOA-HC) [31] was also proposed to study this dataset [26]. In this case, the accuracy obtained was 93.8%. Moreover, the precision was 92.3%, the sensitivity was 82.8%, and the F1-score was 82.8% [31]. Even though this method achieved slightly better accuracy than EnsFuzEn^{M}_{2D}, when analyzing the F1-score, which balances the sensitivity and the precision, we observe that KNN EnsFuzEn^{M}_{2D} F1-score was 88.9% compared to the 82.8% of BM-BOA-HC method. Overall, our classification values based on an ensemble-based entropy were similar to the machine and deep learning approaches.

In our study, the introduced ensemble entropy techniques are not directly correlated; instead, they target distinct aspects of information entropy. Each introduced ensemble entropy serves a specific purpose based on its original form, aiming to capture unique characteristics of 2D data, with its own set of advantages and disadvantages. For instance, EnsDistrEn_{2D}, EnsPerEn_{2D}, and EnsDispEn_{2D} measure the information learned within a system, while EnsSampEn_{2D} and EnsFuzEn_{2D} are rooted in conditional entropy, addressing the rate of information production in a system. However, it is important to note that EnsPerEn_{2D} exhibits sensitivity to artifacts, EnsDistrEn_{2D} demands storage for a large number of elements, and EnsDispEn_{2D}'s performance can be affected by its parameters. Additionally, the computation time for both EnsSampEn_{2D} and EnsFuzEn_{2D} is relatively high.

In spite of the promising findings, we suggest selecting a subset of two-dimensional entropy models to distinguish various features for an ensemble method. Using all the models might provide redundant information instead of helping to improve the performance of an ensemble method for an application. Thus, we may need to exclude redundant models and choose a subset of all the models based on the mutual information between the data obtained by these models. Such a selection strategy aims to select better two-dimensional entropy models among original entropy approaches. In fact, the key aim of the entropy ensemble selection is the selection of an appropriate subset of base entropy approaches and forms a smaller entropy ensemble that performs better than the set of all of the base entropy methods. The ensemble forms of multiscale two-dimensional and three-dimensional ensemble entropy can also be developed to produce a larger set of textural features [10,32,33].

4. Conclusions

In this study, we developed the low bias ensemble forms of two-dimensional SampEn, DispEn, FuzEn, and DistEn entropy methods to quantify the image irregularity or uncertainty. The results on the synthetic data showed that the ensemble approaches are able to detect the degrees of randomness and periodicity in images. The ensemble process led to more stable results (lower coefficient of variations). Even though these approaches were tested with synthetic images, additional tests of images under noisy situations should be evaluated to test the robustness of these algorithms under different noise conditions. Finally, classifiers based on most of the ensemble approaches distinguished normal tissue, centrilobular emphysema, and paraseptal emphysema with high accuracy and outperformed classifiers based on their corresponding original entropy forms. Together, thanks to its ability to detect low bias and stable patterns of an image, the present findings support further study of ensemble entropy in various applications of image processing.

CRediT authorship contribution statement

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

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Andreia S. Gaudencio reports a relationship with Foundation for Science and Technology that includes: funding grants. Andreia S. Gaudencio reports a relationship with European Social Fund that includes: funding grants.

Acknowledgements

This work was supported by FCT (Foundation for Science and Technology) 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. This work was also supported by both FCT and the ESF (European Social Fund) under the scholarship UI/BD/152802/2022. The authors would also like to thank to the Laboratory of Advanced Computing (LCA) at the 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] L.E.V. Silva, A.C.S. Senra Filho, V.P.S. Fazan, J.C. Felipe, L.O. Murta Junior, Two-dimensional sample entropy: assessing image texture through irregularity, Biomedical Physics & Engineering Express 2 (4) (jul 2016) 045002.

[2] Luciano Zunino, Haroldo V. Ribeiro, Discriminating image textures with the multiscale two-dimensional complexity-entropy causality plane, Chaos Solitons Fractals 91 (oct 2016) 679–688.

[3] Christopher John Moore, A threshold structure metric for medical image interrogation: the 2d extension of approximate entropy, in: 2016 20th International Conference Information Visualisation (IV), IEEE, 2016, pp. 336–341.

[4] Hamed Azami, Javier Escudero, Anne Humeau-Heurtier, Bidimensional distribution entropy to analyze the irregularity of small-sized textures, IEEE Signal Process. Lett. 24 (9) (sep 2017) 1338–1342.

[5] Luiz Fernando Segato dos Santos, Leandro Alves Neves, Guilherme Botazzo Rozendo, Matheus Gonçalves Ribeiro, Marcelo Zanchetta do Nascimento, Thaína Aparecida Azevedo Tosta, Multidimensional and fuzzy sample entropy (SampEnMF) for quantifying h&e histological images of colorectal cancer, Comput. Biol. Med. 103 (dec 2018) 148–160.

[6] Adriana Antonelli, Gustavo Meschino, Virginia Ballarin, Mammographic density estimation through permutation entropy, in: IFMBE Proceedings, Springer, Singapore, may 2018, pp. 135–141.

[7] Hamed Azami, Luiz Eduardo Virgilio da Silva, Ana Carolina, Mieko Omoto, Anne Humeau-Heurtier, Two-dimensional dispersion entropy: an information-theoretic method for irregularity analysis of images, Signal Process. Image Commun. 75 (jul 2019) 178–187.

[8] Mirvana Hilal, Andreia Sofia, F. Gaudêncio, Clémence Berthin, Pedro G. Vaz, Joao Cardoso, Ludovic Martin, Anne 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.

[9] Mirvana Hilal, Anne Humeau-Heurtier, Bidimensional fuzzy entropy: principle analysis and biomedical applications, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2019, pp. 4811–4814.

[10] Mirvana Hilal, Clemence Berthin, Ludovic Martin, Hamed Azami, Anne Humeau-Heurtier, Bidimensional multiscale fuzzy entropy and its application to pseudoxanthoma elasticum, IEEE Trans. Biomed. Eng. 67 (7) (jul 2020) 2015–2022.

[11] Cristina Morel, Anne Humeau-Heurtier, Multiscale permutation entropy for two-dimensional patterns, Pattern Recognit. Lett. 150 (2021) 139–146.

[12] Mirvana Hilal, Andreia S. Gaudêncio, Pedro G. Vaz, João Cardoso, Anne Humeau-Heurtier, Colored texture analysis fuzzy entropy methods with a dermoscopic application, Entropy 24 (6) (2022) 831.

[13] Andreia S. Gaudêncio, Mirvana Hilal, João M. Cardoso, Anne Humeau-Heurtier, Pedro G. Vaz, Texture analysis using two-dimensional permutation entropy and amplitude-aware permutation entropy, Pattern Recognit. Lett. (2022).

[14] S.M. Pincus, Approximate entropy as a measure of system complexity, Proc. Natl. Acad. Sci. 88 (6) (mar 1991) 2297–2301.

[15] Joshua S. Richman, J. Randall Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am. J. Physiol., Heart Circ. Physiol. 278 (6) (jun 2000) H2039–H2049.

[16] Hamed Azami, Saeid Sanei, Tarek K. Rajji, Ensemble entropy: a low bias approach for data analysis, Knowl.-Based Syst. (2022) 109876.
[17] Michael A. Nielsen, Isaac L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, 2010.
[18] Matthias Dehmer, Abbe Mowshowitz, A history of graph entropy measures, Inf. Sci. 181 (1) (2011) 57–78.
[19] Denis Sh Sabirov, Information entropy changes in chemical reactions, Comput. Theor. Chem. 1123 (2018) 169–179.
[20] Denis Sh Sabirov, Information entropy of mixing molecules and its application to molecular ensembles and chemical reactions, Comput. Theor. Chem. 1187 (2020) 112933.
[21] Masashi Takahashi, Junya Fukuoka, Norihisa Nitta, Ryutaro Takazakura, Yukihiro Nagatani, Yoko Murakami, Hideji Otani, Kiyoshi Murata, Imaging of pulmonary emphysema: a pictorial review, Int. J. Chron. Obstruct. Pulmon. Dis. 3 (2) (2008) 193.
[22] Diana Litmanovich, Phillip M. Boiselle, Alexander A. Bankier, Ct of pulmonary emphysema-current status, challenges, and future directions, Eur. Radiol. 19 (3) (2009) 537–551.
[23] Afarine Madani, Caroline Keyzer, Pierre-Alain Gevenois, Quantitative computed tomography assessment of lung structure and function in pulmonary emphysema, Eur. Respir. J. 18 (4) (2001) 720–730.
[24] Christoph Bandt, Bernd Pompe, Permutation entropy: a natural complexity measure for time series, Phys. Rev. Lett. 88 (Apr 2002) 174102.
[25] Weiting Chen, Zhizhong Wang, Hongbo Xie, Wangxin Yu, Characterization of surface EMG signal based on fuzzy entropy, IEEE Trans. Neural Syst. Rehabil. Eng. 15 (2) (jun 2007) 266–272.
[26] Lauge Sorensen, Saher B. Shaker, Marleen De Bruijne, Quantitative analysis of pulmonary emphysema using local binary patterns, IEEE Trans. Med. Imaging 29 (2) (2010) 559–569.
[27] James Algina, H.J. Keselman, Randall D. Penfield, An alternative to Cohen's standardized mean difference effect size: a robust parameter and confidence interval in the two independent groups case, Psychol. Methods 10 (3) (2005) 317.
[28] T. Manikandan, S. Maheswari, Automated classification of emphysema using data augmentation and effective pixel location estimation with multi-scale residual network, Neural Comput. Appl. (2022) 1–16.
[29] Sweta Parui, Debanjan Parbat, Monisha Chakraborty, A deep learning paradigm for computer aided diagnosis of emphysema from lung hrct images, in: International Conference on Computing in Engineering & Technology, Springer, 2022, pp. 198–207.
[30] Anisha Isaac, H. Khanna Nehemiah, Snofy D. Dunston, V.R. Elgin Christo, Arputharaj Kannan, Feature selection using competitive coevolution of bio-inspired algorithms for the diagnosis of pulmonary emphysema, Biomed. Signal Process. Control 72 (2022) 103340.
[31] Sumita Mondal, Anup K. Sadhu, Pranab Kumar Dutta, Automated diagnosis of pulmonary emphysema using multi-objective binary thresholding and hybrid classification, Biomed. Signal Process. Control 69 (2021) 102886.
[32] Andreia Sofia, F. Gaudêncio, Pedro G. Vaz, Mirvana Hilal, João M. Cardoso, Guillaume Mahé, Mathieu Lederlin, Anne Humeau-Heurtier, Three-dimensional multiscale fuzzy entropy: validation and application to idiopathic pulmonary fibrosis, IEEE J. Biomed. Health Inform. 25 (1) (2020) 100–107.
[33] Delphine Lebret, Andreia S. Gaudêncio, Mirvana Hilal, Sonia Saib, Rakelle Haidar, Michel Nonent, Anne Humeau-Heurtier, Three-dimensional dispersion entropy for uterine fibroid texture quantification and post-embolization evaluation, Comput. Methods Programs Biomed. 215 (2022) 106605.