Spatial Uncertainty Modeling of Fuzzy Information in Images for Pattern Classification

Tuan D. Pham*
Aizu Research Cluster for Medical Engineering and Informatics, Center for Advanced Information Science and Technology, The University of Aizu, Aizu-Wakamatsu, Fukushima, Japan

Abstract
The modeling of the spatial distribution of image properties is important for many pattern recognition problems in science and engineering. Mathematical methods are needed to quantify the variability of this spatial distribution based on which a decision of classification can be made in an optimal sense. However, image properties are often subject to uncertainty due to both incomplete and imprecise information. This paper presents an integrated approach for estimating the spatial uncertainty of vagueness in images using the theory of geostatistics and the calculus of probability measures of fuzzy events. Such a model for the quantification of spatial uncertainty is utilized as a new image feature extraction method, based on which classifiers can be trained to perform the task of pattern recognition. Applications of the proposed algorithm to the classification of various types of image data suggest the usefulness of the proposed uncertainty modeling technique for texture feature extraction.

Introduction
There are different types of images used in the diverse applications of image classification; to name a few, medical images, biological images, remote-sensing images, scene images, and so on. The information contents of different image types are different from each other, but they may share some common properties. To classify categorical images, the numerical description of the images, which is known as an image feature, is necessary in order to capture their distinctive characteristics that can be used for training classifiers or as labeled samples. A critical challenge in the discriminative quantification of image properties of various regions of interest is that they are usually subject to noise and vague boundaries between the objects and background, which are often found in medical, life-science, and natural data [1]–[7]. Consequently, these factors adversely affect the classification performance. To deal with the uncertainty of image information, statistical measures of sets of images are often utilized to construct probability models of images, in which they are considered as random variables. An approach for handling imprecision rather than randomness in images is to consider them as fuzzy events so that some non-probabilistic measure of uncertainty can be established.

Based on the example made to interpret the definition of the entropy of fuzzy sets [8], the above notions of uncertainty in images can be elucidated by considering examples of the outcomes of image segmentation and image enhancement. Figure 1 shows an original part of an MRI image of the brain, in which the bright areas are the white matter hyperintensities of the brain. The image intensity levels are in the range [0, 1]. Figures 2 and 3 are the two binary segmentation results of the original image (Figure 1) using the gray-level threshold of 0.1647 obtained from the Otsu’s image segmentation method [9] (there are well-known public-domain software packages for MRI brain segmentation, which include prior anatomical information, such as SPM [10] and its latest version SPM8, and FSL [11]; this segmentation method used here to only provide a simple example of a segmented section of the brain on MRI for the conceptual discussion in this study), and by setting the threshold at 0.25, respectively. Figures 4–5 shows an enhancement result of the original image using an adjustment method that increases the contrast of the image by mapping the values of the original intensity image to new values such that 1% of the data is saturated at low and high intensities of the input data. Figure 5 is another enhancement result of the original image using histogram equalization that enhances the contrast of the image by transforming the values in the original image so that the histogram of the output image approximately matches the uniform distribution histogram. The pixel intensity in Figures 2 and 3 are either black (background) or white (object). In other words, the uncertainty involves in the segmentation outcomes refer to the presence or absence of object pixels, which can be modeled as a random variable taking on the values of either 1 (white object) or 0 (black background). The uncertainty associated with the outcomes shown in Figures 4 and 5 can be expressed in terms of the degree of grayness of the pixels that are subject to imprecision or subjectiveness. Therefore, such a description of imprecise pixel values can be considered as a fuzzy set. In fact, image enhancement is the process that attempts to improve an image appearance so that it looks subjectively better. There exist no standards that can guide how the image should look, but one can...
tell if it has been improved or not by considering, for example, the
detailed contents or contrast of the image [12].

The treatment of uncertainty in images have been largely
discussed using the theory of fuzzy sets [13,14] and geostatistics
[15,16]. Fuzzy logic addresses uncertainty that is caused by
imprecision or vagueness in provided information and models such
uncertainty with the admission of degrees of possibility. By using
the notion of fuzzy sets, an event, which is a set of outcomes of an
experiment to which a probability is assigned, can be extended to
define fuzzy events to significantly enhance the applications of the
theory of probability in the fields in which uncertainty being due to
fuzziness is inherently pervasive [17]. Although the integration of
the theories of probability and fuzzy sets is natural in image
analysis and pattern recognition, little effort has been spent on
exploring the potential application of this idea since a work on
binary image thresholding carried out by extending the probability
measure of fuzzy events to calculate the fuzzy measure of similarity
between two sets [18].

Furthermore, the concept of modeling uncertainty associates
different sources of uncertainty with various deterministic and
non-deterministic laws of physical and dynamical processes,
depending on the purpose for which the models are applied.
Many pattern classification problems, such as image analysis,
involve spatial modeling that naturally calls for techniques covered
in geostatistics. In fact, the variability in images with alternating
high and low pixel values is evident in many practical domains of
medicine, science and engineering. Modeling spatial continuity in
images is therefore critical to providing a solution to the question of
addressing uncertainty, since a spatial model of image properties
will constitute to a different assessment of uncertainty compared
with the assumption that everything is random [19]. Research
works on geostatistics have been reported in literature, including
ordinary kriging and indicator kriging, applied to image processing
[16,20] and classification of remote-sensing images [21]–[23].

The combinations of geostatistics, fuzzy sets and fuzzy cluster
analysis have been developed for image enhancement [15] and
image segmentation [24,25]. However, it appears that, for the first
time, this paper presents an integrated approach to measuring
spatial uncertainty in images by modulating the calculi of
probability and fuzzy sets to incorporate their interdependencies.
The mathematical development of the proposed method is based
on the notion of the probability measure of fuzzy sets and the
definition of the entropy of a fuzzy event with respect to a
probability distribution that is derived from the theory of
geostatistics. The novelty of this mathematical model is that the
inherently fuzzy partitions of the image space can be used to
model the spatial uncertainty of the image by coding the data as
probability values at different degrees of membership of belonging
to the fuzzy partitions. As a result, the derivations of the fuzzy sets
and spatial probability measures of the image uncertainty allow
the quantification of the entropy of fuzzy pixels with respect to
their probability distributions. This type of a measure of
uncertainty or entropy can be readily utilized as a pattern feature
for image classification.

Methods

Entropy of a fuzzy event

Consider a Euclidean $n$-space $R^n$ and a probability space
represented by a triplet $(R^n, B, P)$, where $B$ is the $\sigma$-field of Borel
sets (Borel sets in a topological space are the $\sigma$-algebra generated
by the open sets. An algebra of sets which is closed under
countable unions is known as a $\sigma$-algebra, $\sigma$-field or Borel field

Figure 1. Image showing part of an MRI of the brain: bright
patches are the white matter hyperintensities or lesions of the
brain, and other parts are brain tissues and fluid.
doi:10.1371/journal.pone.0105075.g001

Figure 2. Binary segmentation of the brain MRI using gray-
level threshold of 0.3.
doi:10.1371/journal.pone.0105075.g002
in $R^n$ and $P$ is a probability measure over $R^n$. Also, let $x \in R^n$ and $B \in S$. The probability of $B$, $P(B)$, can be expressed as

$$P(B) = \int_{R^n} \beta_B(x) \, dP = E(\beta_B).$$

$$P(A) = \int_{R^n} \mu_A(x) \, dP = E(\mu_A).$$

As Eq. (1) defines the probability of a crisp event as the expectation of its characteristic function, so the probability of a fuzzy event defined in Eq. (2) is the expectation of its membership function. The presented definitions of a fuzzy event and its probability constitute to the generalized framework of the theories of fuzzy sets, probability, and information [17]. Further study has also shown that the theory of probability is of a rich structure for incorporating fuzzy events within its framework to logically generate probabilities of fuzzy events so that the uncertainty of outcomes and of imprecision can be successfully unified [27].

To specifically explore how the concepts of fuzzy sets, probability and information can be made in a coherent framework; we turn the discussion to the notion of the entropy of a probability distribution, which is mathematically expressed as

$$H(x) = - \sum_{i=1}^{n} p_i \log(p_i).$$

$$H^P(A) = - \sum_{i=1}^{n} \mu_A(x_i) p_i \log(p_i)$$

where $H(x)$ is the entropy of the distribution $P = \{p_1, \ldots, p_n\}$, and for $p_i = 0$, $\log(p_i) = 0$, indicating the non-feasibility of obtaining information about an impossible event.

The definition expressed in Eq. (3) suggests that the entropy of fuzzy event $A = \{x_1, \ldots, x_n\}$ with respect to a probability distribution $P = \{p_1, \ldots, p_n\}$ can be defined as [17]

The above definition of the entropy of a fuzzy set forms the basis for our modeling of the spatial uncertainty of the intensity distribution in an image. It has been widely known that categorical information is inherently imprecise in images, particularly in the context of biology and medicine [14]–[31]. First, we apply the theory of fuzzy sets as a calculus for the treatment of uncertainty associated with the classification of image intensity distributions. Second, we use geostatistical tools to study the spatial continuity as a transition probability in an image to evaluate its spatial uncertainty. These two types of uncertainty measures allow fuzzy
sets, probability and information to work in concert for the identification of image classes.

Modeling image-content imprecision with the FCM algorithm

Uncertainty in an image, which is inherently due to the imprecise description of the image content, can be mathematically modeled by the partition of the image space using the fuzzy c-means algorithm [32]. Let an image \( I \) of size \( L \times W \) be arbitrarily partitioned into a number of imprecise clusters using the fuzzy c-means algorithm [32]. Mathematically, let \( M_k \) be a fuzzy c-partition space, \( X \) be a subset of the real \( g \)-dimensional vector space \( R^g \) : \( X = \{x_1, x_2, \ldots, x_n\} \subseteq R^g \) where \( x_k = (x_{k1}, x_{k2}, \ldots, x_{kg}) \in R^g \). The fuzzy c-means (FCM) clustering is based on the minimization of the fuzzy objective function \( J_F : M_k \times R^g \rightarrow R^+ \), which is defined as [32]

\[
J_F(U,v) = \sum_{i=1}^{N} \sum_{j=1}^{c} (\mu_{ij})^q [d(x_i, v_j)]^2
\]

where \( q \in ]1, \infty[ \) is the fuzzy weighting exponent, \( U \in M_k \), \( v = (v_1, v_2, \ldots, v_c) \in R^{cg} \), and \( d(x, v_j) \) is any inner-product norm metric induced on \( R^g \).

The fuzzy objective function expressed in Eq. (5) is subject to the following constraints:

\[
\sum_{j=1}^{c} \mu_{ij} = 1, i = 1, \ldots, L \times W
\]

where

\[
\mu_{ij} \in [0,1], i = 1, \ldots, L \times W, j = 1, \ldots, c.
\]

The objective function \( J_F(U,v) \) is a squared error clustering criterion and to be minimized to optimally determine \( U \) and \( v \). A solution to the minimization of the objective function is by a process of iteratively updating \( U \) and \( v \) until some convergence is reached [32]. Thus, given fuzzy \( c \) partitions, the FCM assigns each pixel to the \( c \) clusters with its respective membership grades. In other words, the FCM-based cluster analysis can be utilized to construct the modeling of uncertainty in an image in the context of imprecise boundaries or ill-defined classes.

Modeling spatial uncertainty of imprecision with indicator kriging formalism

The uncertainty involved with imprecision in image intensity has been addressed with the notion of fuzziness; whereas the uncertainty regarding to pixel locations refers to spatial randomness, which can be modeled with the indicator kriging formalism of geostatistics [33]. The indicator kriging has been applied as a natural tool for determining a non-parametrical conditional probability distribution of categorical data [34]. An interest in this study is to utilize indicator kriging to construct local spatial distributions of uncertainty in an image, which can be incorporated within the framework of the probability measure of fuzzy information. Let \( f_i \) be the intensity value of a pixel located at \( i, i = 1, \ldots, L \times W \), \( L \times W \) is the size of the image. Here, the purpose of applying the indicator formalism is to estimate the probability distribution of uncertainty at unsampled location \( i \). The cumulative distribution function is usually estimated with a set of cutoff thresholds \( z_t, t = 1, \ldots, T \); and the probabilities are then determined by coding the data as binary indicator values. The indicator coding at location \( i \) is defined as follows [34]

\[
I(i,T) = P(f_i \leq z_t) = \begin{cases} 1 : & f_i \leq z_t, \\ 0 : & \text{otherwise}. \end{cases}
\]

Using thresholds \( z_t, t = 1, \ldots, T \), as values in the range of the image intensity does not conveniently offer a procedure for modeling spatial uncertainty in an image (for example, the thresholds can be chosen from the histogram of the data to represent percentiles, which are the values below the percentanges of the observations [35]). Therefore, instead of using \( z_t \), we make use of the previously discussed fuzzy image partitions or clusters \( v_j, j = 1, \ldots, c \), which allows every pixel to belong to every partition with different fuzzy membership grades, and apply a series of \( x \)-level cuts, \( z \in [0,1] \), to code the categorical pixels as follows:

\[
I_j(i, z) = \begin{cases} 1 : & \mu_{ij} \geq z, \\ 0 : & \text{otherwise}, \end{cases}
\]

where \( I_j(i, z) \) is the indicator that codes the assignment of \( f_i \) to cluster \( v_j \), having a fuzzy membership grade being equal to or greater than \( z \). Here, \( z \) can be selected as a set of the fuzzy membership grades to represent the degrees of imprecision that indicate the possibility being higher than the most fuzzy value of 0.5.

The next step of the indicator kriging formalism is the determination of the cumulative distribution function (CDF), which characterizes the probability of \( f_i \) belonging to \( v_j \) with a membership value of being greater or equal to \( z \), and can be
mathematically expressed as

\[
F_y(i, x) = P(\mu_i \geq x) \tag{9}
\]

Taking advantage of the available information of \(M\)-neighboring data, the conditional CDF is

\[
F_y(i, x|M) = P(\mu_i \geq x|M) \tag{10}
\]

where \(M\) is the number of neighboring pixels of \(i\).

The CDF according to the indicator expressed in Eq. (8) can be estimated using the ordinary kriging [35], and the result of indicator kriging is a model of spatial uncertainty at the pixel at location \(i\):

\[
\hat{F}_y(i, x|M) = \sum_{m=1}^{M} w_m(i, x) I_m(x, m) \tag{11}
\]

where \(w_m(i, x)\) is the ordinary kriging weight that indicates the influence of neighboring pixel \(m\) over pixel \(i\) with respect to level cut \(x\). These weights can be obtained by the ordinary kriging system of equations [35]:

\[
\begin{align*}
\sum_{m=1}^{M} w_m(i, x) \gamma_{m, m'} + \lambda = & \gamma_{i, m'}', \quad m' = 1, \ldots, M, \\
\sum_{m=1}^{M} w_m(i, x) = & 1,
\end{align*} \tag{12}
\]

where \(\lambda\) is a Lagrange multiplier, \(\gamma_{i, m'}',\) with a lag of absolute difference \(|m - m'| = h\), is the semivariogram of the indicator \(I_y(i, x)\) expressed in Eq. (8), and is defined as the expected value [34]:

\[
\gamma_h = \frac{1}{2} E[(I_y(x, m) - I_y(x, m+h, x))^2] \tag{13}
\]

The indicator semi-variogram that is experimentally calculated for lag distance \(h\) is defined as the average squared difference of values separated by \(h\):

\[
\tilde{\gamma}_h = \frac{1}{2N(h)} \sum_{N(h)} [(I_y(x, m) - I_y(x, m+h, x))^2] \tag{14}
\]

where \(N(h)\) is the number of pairs for lag \(h\).

Alternatively, the ordinary kriging system can be expressed in a matrix form as

\[
Gw = g \tag{15}
\]

where \(G\) is the square and symmetrical matrix that represents the semi-variogram value between the known neighboring values \(I_y(m, x), m = 1, \ldots, M; w\) is the vector of kriging weights; and \(g\) is the vector representing the semi-variogram values between \(I_y(m, x)\) and \(I_y(i, x), m = 1, \ldots, M, m \neq i\). These terms are defined as [35]

\[
\begin{bmatrix}
\gamma_{11} & \gamma_{12} & \cdots & \gamma_{1M} & 1 \\
\vdots & \ddots & \ddots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\gamma_{M1} & \gamma_{M2} & \cdots & \gamma_{MM} & 1 \\
1 & 1 & \cdots & 1 & 0
\end{bmatrix}
\]

where \(\gamma_{ij}\) is the semi-variance of \(I_y(m = 1, x)\) and \(I_y(m = 2, x)\);

\[
w = [w_1 \ w_2 \ \cdots \ w_M - \lambda]^T
\]

where \(w_1, \ldots, w_M\) are the kriging weights, \(\lambda\) is a Lagrange multiplier, and \(\sum_{m=1}^{M} w_m = 1;\) and

\[
g = [\gamma_{11} \ \gamma_{12} \ \cdots \ \gamma_{1M} \ 1]^T
\]

where \(\gamma_{1M}\) is the semi-variance of \(I_y(i, x)\) and \(I_y(M, x)\).

Given that \(G^{-1}\) exists, the kriging weights can be obtained by solving:

\[
w = G^{-1}g \tag{16}
\]

An implicit assumption of the ordinary kriging system having presented in Eq. (12) is that the underlying statistics is invariant in space under translation. Such a property is known as statistically stationary. However, statistical stationarity is a property of a random function, but not an inherent property of real data [36]. This nonstationary property is also true for medical images in which different internal organs can have different variations of the image intensity and the mean of the image changes locally. Here, kriging with a nonstationary mean is applied in order to enhance the reliability of the estimate of the kriging weights. This technique is called universal kriging (UK) [37,38].

In ordinary kriging, the estimation is carried out with the error variable from a stationary mean that must be known at all positions and can be set as the global mean or modeled with a drift or a local trend. A local mean with a drift \(m(u)\) can be modeled as a linear combination of the geometric coordinates of the pixels with a local neighborhood as

\[
m(u) = \sum_{k=0}^{K} a_k L_k(u), \tag{17}
\]

where \(a_k\) are unknown drift coefficients, \(L_0(u) = 1\) (constant function for the constant-mean case), and \(L_k(u), k = 1, \ldots, K\), are the polynomials or basis functions, which can be modeled as the first-degree or second-degree terms as follows, respectively.

\[
m(u) = m_0 + \sum_{k=1}^{K} (a_1 x_{1k} + a_2 x_{2k}), \tag{18}
\]

where \(x_{1k}\) and \(x_{2k}\) are the pixel coordinates in row-wise and column-wise of \(f_k\), respectively.

The drift effect can be incorporated into the ordinary kriging system to find kriging weights as additional constraints. Solving this extended set of simultaneous equations, a set of universal kriging weights that model the drift within the local neighbors around the location of the unknown value. In general, the UK system can be expressed with the following matrix structure.
where \( G^*, w^*, \) and \( g^* \) are the \( G \) without the last row and last column, \( w^* \) without the last row, and \( g \) without the last row as defined for the ordinary kriging system, respectively; 
\[
\begin{bmatrix}
G^* \\ L \\ L^T \\ 0 \\
\end{bmatrix}
\begin{bmatrix}
w^* \\ m \\ 1 \\
\end{bmatrix} =
\begin{bmatrix}
g^* \\ \lambda \\
\end{bmatrix}
\tag{20}
\]

It is noted that the above decision rule expressed in Eq. (21) is general and can be applied to any type of pattern classifiers.

### Results and Discussion

#### Detection of Mitochondria in Microscope Images

The proposed method was applied for the detection of mitochondria in microscope images. The mitochondrion is a membrane-bound organelle found in most eukaryotic cells. Mitochondria are considered as the powerhouse of the cell because they function as the platform for generating the production of chemical energy. The visual information of mitochondria revealed by the recent advanced technology in nanoimaging opens doors to life-science researchers to gain insights into its spatial structure and its spatial distribution within the cell. In order to simulate and model mitochondria using a large amount of images, the first task in image processing is the automated detection of this organelle. In fact, the classification of molecular images has been a long-pursued research in the disciplinary field of engineering and computer science in life sciences [39,40]. However, there is always a strong demand for exploring appropriate feature extraction methods for the automated identification of particular types of objects or regions of interest in cell biology with different levels of technical challenge, ranging from the detection of cells nuclei [41] to subcellular patterns [2]. If different types of the images can be automatically distinguished by computerized methods, such an ability can help researchers to quickly and accurately study cell function to discover mechanisms underlying complex diseases, and carry out spatial modeling and simulation of biological signaling pathways, which may identify critical organelles attributing to the regulation of the cellular process within the intracellular space.

The cells were imaged using scanning electron microscopy (SEM) and focused ion beam (FIB) technology with Helios NanoLab 650, which is one of the most recent advances in field emission SEM and FIB technologies and their combined use, and designed to access extremely high resolution characterization, and higher quality sample preparation. Figure 6 is a typical SEM-FIB image showing half of the intracellular space of a cancer cell line that was derived from a human head and neck squamous cell carcinoma (SCC-61) parental line [42]. Figure 7 shows a typical FIB-SEM image of the same cancer cell in which the ground-truth mitochondria were manually identified and marked by a cell biologist. The interest here is to detect image regions of interest that contain the mitochondria. Such detected regions will greatly alleviate the difficulty in the image segmentation of the mitochondria [43] to facilitate the spatial modeling and simulation of the role of this major organelle for studying human complex diseases such as cancer [44,45].

The detection of the mitochondria in the intracellular space was carried out with a window of 53 by 60 pixels, which is the average size of the mitochondria in the images. The number of clusters \( c \) and the exponent \( q \) expressed in Eq. (5) were selected to be three (to approximately represent the number of intensity groups in the images) and two (commonly specified in many applications), respectively. The values of the \( \alpha \)-cut used for the indicator expressed in Eq. (8) are 0.5, 0.6, 0.7, 0.8 and 0.9. The numbers of the neighboring pixels \( M \) used in Eq. (11) are 5 and 7. The detection of the mitochondria was performed by moving the 53-by-60 window along the horizontal and vertical directions of the

Image classification with integrated uncertainty modeling

The integrated framework for modeling uncertainty due to both imprecision and randomness has been formulated. In other words, a new type of image feature has been introduced in terms of the probability of a fuzzy event for pattern classification. In this context, a fuzzy event can be an imprecise object or sample to be categorized identified. In application, the next task is to decide which class that best matches the feature extracted from the unknown sample. This is pattern classification that associates the appropriate class label with the test sample by using the descriptive features. A general way is to use a function or a classifier, such as a distance measure, to find the class with features that differ the least from the features of the unknown sample. This discussion completes with a decision or classification procedure for a computed set of entropy features for \( o_a, s = 1, \ldots, K \), classes as follows.

Let fuzzy cluster centers \( v_1 > \cdots > v_c \) and \( z \)-level cuts \( z_1 > \cdots > z_b \) (if such orders do not exist, then the orders are rearranged). Also, let \( H^P(A,z_1,v), r = 1, \ldots, b, j = 1, \ldots, c \) be the entropy of fuzzy event \( A \) with respect to a probability distribution \( P \) defined in Eq. (4), obtained by using \( z \)-level cut \( z_r \) and fuzzy partition \( v_j \) for class \( o_a \):

\[
L_1(o_a) = [H^P(A,z_1,v_1), \ldots, H^P(A,z_1,v_c)], \ldots, \\
L_c(o_a) = [H^P(A,z_c,v_1), \ldots, H^P(A,z_c,v_c)].
\]

Given \( \phi_j(L) = f(\omega_jL), j = 1, \ldots, c \), where \( f(\cdot) \) is a monotonically increasing discriminant function (the larger the value of the function the better the match); the decision for classification is carried out as follows:

Classify \( L \) in \( o_a \) if \( \phi_j(L) > \phi_s(L), \forall s \neq s. \)
image to extract different features for training. Twenty scans of the FIB-SEM images of the cancer single cell were available in this study. To show the effectiveness of various feature extraction methods, only one image was used for training. The training was performed by extracting the proposed probabilistic entropy measure of the fuzzy information, expressed in Eq. (4), of the mitochondrial and non-mitochondrial regions using OK, denoted as PEFI1, and expressed in Eq. (12), and using UK, denoted as PEFI2, expressed in Eq. (20), respectively. To compare with other feature extraction methods, the same images were used to obtain the gray-level co-occurrence matrix (GLCM), fractal dimension (FD), semi-variogram (SV), semi-variogram exponent (SVE), and the indicator-kriging co-occurrence matrix (IKCM) for the mitochondrial and non-mitochondrial objects, described in [46]. Ten lags were used to extract the semi-variogram values of each image window. If the image window contained whole or part of a mitochondrion, it was labelled as a mitochondrial object. This is designed to capture all small regions of interest containing the mitochondria in order to maximize the sensitivity (true positive rate), while the specificity (true negative rate) can be first reasonably obtained and then maximized in the localized image segmentation task performed window by window. To validate the effectiveness of the extracted features, two simple measures that are the Euclidean and Mahalanobis distances were used to calculate the similarity between the unknown (test) samples and the trained prototypes of the mitochondrial and non-mitochondrial objects.

Sensitivity and specificity are statistical measures of the performance of a binary classification test. In this study, the sensitivity (true positive rate) is the percentage of the mitochondrial regions that are correctly identified; whereas specificity (true negative rate) is the percentage of the non-mitochondrial regions that are correctly identified. Tables 1 and 2 show the sensitivity and specificity of the experiment obtained from several feature extraction methods using the Euclidean and Mahalanobis distances, respectively. In general, the SV, SVE, IKCM, PEFI1 and PEFI2 performed consistently using either the Euclidean distance or Mahalanobis distance. The PEFI2 yields the best
results in both sensitivity and specificity in both distance measures: sensitivity = 100% and specificity = 93%, using the Euclidean distance; sensitivity = 100% and specificity = 97%, using the Mahalanobis distance. The PEFI1 yields the second best: sensitivity = 100% and specificity = 91%, using the Euclidean distance; sensitivity = 100% and specificity = 94%, using the Mahalanobis distance. Using the Euclidean distance, the GLCM achieved 100% for sensitivity, but its specificity is lowest (5%) in comparisons with the other features. On the other hand, the FD performed well with the specificity (90%) but poorly with the sensitivity (35%), using the Euclidean distance. In general, the use of the Mahalanobis distances improved all the detection results provided by all the features, in which the specificity obtained by the GLCM (65%) is significantly higher than using the Euclidean distance.

Identification of Abdominal Tissues on Computed Tomography

The proposed method was also tested for abdominal wall hernia mesh tissue classification on computed tomography (CT), which was recently carried out in [47]. The data were obtained from the abdominal and pelvic CT scans of patients. Eight types of mesh were studied in this experiment: alloderm (M1), marlex (M2), parietex (M3), proceed (M4), strattice (M5), surgimend (M6), and surgisis (M7), and permacol (M8). The available numbers of M1 = 10, M2 = 9, M3 = 30, M4 = 34, M5 = 18, M6 = 15, and M7 = 7, M8 = 54. Half of the samples of each mesh type were used for training and the other half for testing. The classification was carried out by the \( k \)-nearest neighbor method [48], where \( k = 3 \) to decide which type of mesh was present. Ten Monte-Carlo iterations were used for the random selection of training and testing data to enhance the statistics of the experimental results.

The results obtained from the feature extracted by the proposed spatial uncertainty modeling (SUM) were compared with other features extracted by wavelets, the gray-level co-occurrence matrix entropy (GLCME), geostatistical entropy (GE), probabilistic Fusion (PF), and entropy fusion (EF) models, which were carried out in [47]. Further details about the CT data and implementations of GLCME, GE, and EF methods were described in [47]. For the implementation of the FCM, the number of clusters \( c \) and the exponent \( q \) expressed in Eq. (5) were selected to be three (to approximately represent the organs of gray and white intensities and the background) and two (commonly specified in many applications), respectively. The values of the \( \alpha \)-cut used for the indicator expressed in Eq. (8) are 0.5, 0.6, 0.7, 0.8 and 0.9. The numbers of the neighboring pixels \( M \) used in Eq. (11) are 5 and 7. The total average results obtained from the current technique, using OK (PEFI1) and UK (PEFI2), and other methods are shown in Table 3. The results show that the performance of proposed feature extracted by PEFI1 (94.50%) and PEFI2 (94.92%) are the best among the other features for the classification task. The performance of PEFI2 is only slightly better than that of PEFI1. It should be noted that the proposed feature not only outperforms the other individual features (wavelets, GLCME, and GE), but also yields better classification rates than the combinations of the

| Table 1. Sensitivity and Specificity of Mitochondrial Detection Obtained from Various Features using Euclidean Distance. |
|-----------------|--------|--------|--------|--------|--------|--------|--------|
| GLCM            | FD     | SV     | SVE    | IKCM   | PEFI1  | PEFI2  |
| Sensitivity     | 100    | 35     | 95     | 100    | 100    | 100    |
| Specificity     | 5      | 90     | 70     | 50     | 87     | 91     | 93     |

doi:10.1371/journal.pone.0105075.t001
GLCM and geostatistical models in terms of entropy (92.81%) and probability (90.76%) measures.

### Classification of Logos on Document Imaging

Furthermore, the proposed spatial uncertainty modeling approach was tested for the classification of logos on document images. Ten sets of logos were obtained from the public-domain logo database of the University of Maryland, which consist of 105 intensity logo images. Fifty other logo images were also included, embedded in several document formats including letters, faxed documents, and billing statements [49,50]. All logo images were also generated subject to translation, scaling, orientation and degradation to create different sets of images [49,50]. Image rotations include 2-degree, 4-degree, 6-degree, 8-degree, 10-degree orientations. The images were shrunk by the factor of two using the bicubic interpolation and anti-aliasing. For the translation, all images were shifted left (x-shifted) by 50 pixels and up (y-shifted) by 30 pixels. All images were degraded with Gaussian noise of zero mean and 0.02 variance.

The features extracted from the logo images are: 1) semi-variograms, 2) Zernike moments, 3) wavelets, 4) Gabor features, and 5) SUM-based feature. The first four features were studied in [50–52]. These features were equally divided into training and test datasets. For the implementation of the FCM expressed in Eq. (5), the number of clusters \( c = 2 \) to represent the object and background, and the fuzzy exponent \( q = 2 \). The values of the \( \alpha \)-cut used for the indicator expressed in Eq. (8) are 0.5, 0.6, 0.7, 0.8, and 0.9. The numbers of the neighboring pixels \( M \) used in Eq. (11) are 5 and 7.

The datasets were equally split into half for training and the other half for testing. Furthermore, the data were randomly selected for 10 times to repeat the training and testing in order to establish statistically meaningful results of the experiment. The \( k \) nearest neighbor (\( k\)-NN) method was applied for the task of classification, with \( k = 3, 5, \) and 7. The total average classification results shown in Table 4 suggest that the proposed SUM-based feature outperforms the use of the other four features, with the order of performance from the lowest to highest classification rates as follows: wavelet feature, Gabor features, Zernike moments, semi-variograms, and proposed feature (PEFI1 and PEFI2, where both algorithms achieved an equal classification rate).

### Conclusion

A modeling of spatial uncertainty in images for pattern classification using the theories of fuzzy sets and geostatistics has been presented and discussed. The proposed model has been implemented as a new feature extraction method for the classification of image patterns.

The entropy of a fuzzy image information with respect to a probability distribution is calculated as an integrated spatial uncertainty of the image, which can be used for characterizing categorical images. Simple classifiers were trained with this new feature for comparisons with other related existing features. The training of the proposed feature with advanced classifiers can be expected to enhance the results. In particular, the applications of the proposed approach for automated detection of mitochondria in the real intracellular imaging of a cancer cell line, tissue identification and logo classification have been carried out. The comparative results suggest the usefulness of the proposed mathematical framework for image feature extraction. Being similar to the use of the probabilities of the gray-level co-occurrence matrix, the indicator-kriging probabilities can be utilized to construct other statistical features of an image for pattern classification.

The model developed in this study can be further improved by selecting effective strategies for selecting the number of fuzzy clusters and adding additional spatial constraints to the fuzzy objective function. In particular, as the constrained independent component analysis has been developed to reduce ambiguity in studying fMRI data by imposing temporal and spatial constraints to the mathematical model [53]; the fuzzy objective function defined in Eq. (5) can be modified by adding similar temporal and spatial constraints [54,55] to improve the modeling of uncertainty in the setting of geostatistics.

### Table 3. Total Average Classification Rates Provided by Different Methods.

| Classifier                        | Accuracy (%) |
|-----------------------------------|--------------|
| Wavelets                          | 45.29        |
| Gray-Level Co-Occurrence Entropy (GLCME) | 56.61        |
| Geostatistical Entropy (GE)       | 76.72        |
| Probabilistic Fusion (PF)         | 90.76        |
| Entropy Fusion (EF)               | 92.81        |
| PEFI1                             | 94.50        |
| PEFI2                             | 94.92        |

As follows: wavelet feature, Gabor features, Zernike moments, semi-variograms, and proposed feature (PEFI1 and PEFI2, where both algorithms achieved an equal classification rate).
Acknowledgments

The author thanks the Academic Editor, Karl Herbolz, and the two Anonymous Reviewers for their constructive comments that helped improve the manuscript. The MRI image was obtained from Klaus Berger, Instituto de Epidemiology and Social Medicine, The University of Muenster, Germany. The cell image data were obtained from Professor Kazuhisa Ichikawa, The Institute of Medical Science, The University of Tokyo, Japan. The abdominal CT data were provided by Dinh Le, Division of General Surgery, Oregon Health and Science University, USA.

References

1. Uchida S (2013) Image processing and recognition for biological images. Dev Growth Differ. 55:523–549.
2. Zhang B, Pham TD (2011) Phenotype recognition with combined features and random subspace classifier ensemble. BMC Bioinformatics 12:128. DOI:10.1186/1471-2105-12-128.
3. Nixon MS, Aguado AS (2012) Feature Extraction & Image Processing for Computer Vision, 3rd ed. London: Academic Press.
4. Theodoridis S, Koutroambas K (2009) Pattern Recognition, 4th ed. London: Academic Press.
5. Duela RO, Hart PE, Stork DG (2001) Pattern Classification, 2nd ed. New York: John Wiley & Sons.
6. Castaneda CAB, Fraja JS, Fernandes S, Gruber A, Costa LF (2007) Biological image processing for automatic image recognition and diagnosis of protozoan parasites of the genus Eimeria. Pattern Recognition 40:1899-1910.
7. Gonzalez-Rufino E, Carrion P, Cernadas E, Fernandez-Delgado M, Domínguez-Pérez R (2013) Exhaustive comparison of colour texture features and classification methods to discriminate cell categories in histological images of fish ovary. Pattern Recognition 46:2391-2407.
8. De Luca A, Termini S (1972): A definition of nonprobabilistic entropy in the setting of fuzzy sets theory. Information and Control 20:301-312.
9. Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Trans. Systems, Man, and Cybernetics 9:62-66.
10. Friston K, Ashburner J, Kiebel S, Nichols T, Penny W (2007) Statistical Parametric Mapping: The Analysis of Functional Brain Images. Amsterdam: Elsevier.
11. Jenkinson M, Beckman C, Behrens T, Woolrich MW, Smith SM (2012) FSL. Neuroimage 62:782-790.
12. Petrou M, Bosdogianni P (1999) Image Processing: The Fundamentals. Chichester: John Wiley & Sons.
13. Bezdek JC, Pal SK (1992) Fuzzy Models for Pattern Recognition: Methods that Search for Structures in Data. IEEE Press: New York.
14. Chi Z, Yan H, Pham T (1996) Fuzzy Algorithms: With Applications to Image Processing and Pattern Recognition. Singapore: World Scientific.
15. Pham TD, Wagner M (2000) Image enhancement by kriging and fuzzy sets. Int. J. Pattern Recognition 14:1025-1038.
16. Sanchez-Bera LM, Bernabeu E (2006) Uncertainty estimation by convolution setting of fuzzy sets theory. Information and Control 20:301-312.
17. Castellano G, Bonilha L, Li LM, Cendes F (2004) Texture analysis of medical images. Computer Methods and Programs in Biomedicine 113:153-161.
18. Clark ES, Whigham AS, Yarbrough WG, Weaver AM (2007) Cortactin is an essential regulator of matrix metalloproteinase secretion and extracellular matrix degradation in invadopodia. Cancer Research 67:4227-4235.
19. Pham TD (2011) Image segmentation using probabilistic fuzzy c-means clustering. Proc. IEEE Int. Conf. Image Processing 1:722-725.
20. Howson AG (1976): A Handbook of Terms used in Algebra and Analysis. Cambridge: Cambridge University Press.
21. Siringatva ND, Booker JM (2004) Membership functions and probability measures of fuzzy sets. Journal of the American Statistical Association 99:867-877.
22. Cheng HD, Chen JR, Li J (1998) Threshold selection based on fuzzy c-partition entropy approach. Pattern Recognition 31:857-870.
23. Szczeniak PS, Liskien PJG, Kacprzyk J (2000) Fuzzy Systems in Medicine. Heidelberg: Springer.
24. Pham TD, Crane D, Tran TH, Nguyen TH (2004) Extraction of fluorescent cell puncta by adaptive fuzzy segmentation. Bioinformatics 20:2189-2196.
25. Pham TD, Crane D, Tran TH, Nguyen TH (2004) Extraction of fluorescent cell puncta by adaptive fuzzy segmentation. Bioinformatics 20:2189-2196.
26. Assal MI, Aci R (2005) Reliability analysis of microarray data using fuzzy c-means and normal mixture modeling based classification methods. Bioinformatics 21:644-649.
27. Bezdek JC (1981) Pattern Recognition with Fuzzy Objective Function Algorithms. New York: Plenum Press.
28. Creesie N, Wilde GK (2011) Statistics for Spatio-Temporal Data. New Jersey: Wiley.
29. Deutsch CV (2002) Geostatistical Reservoir Modeling. New York: Oxford University Press.
30. Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Trans. Systems, Man, and Cybernetics 9:62-66.
31. Friston K, Ashburner J, Kiebel S, Nichols T, Penny W (2007) Statistical Parametric Mapping: The Analysis of Functional Brain Images. Amsterdam: Elsevier.
32. Jenkinson M, Beckman C, Behrens T, Woolrich MW, Smith SM (2012) FSL. Neuroimage 62:782-790.
33. Cressie N, Wikle CK (2011) Statistics for Spatio-Temporal Data. New Jersey: Wiley.
34. Deutsch CV (2002) Geostatistical Reservoir Modeling. New York: Oxford University Press.
35. Leunghoong O, Khan KD, Deutsch CV (2008) Solved Problems in Geostatistics. New Jersey: Wiley.
36. Davis JC (2000) Statistics and Data Analysis in Geology, 3rd Edition. New York: Wiley, New York.
37. Chiles JP, Delﬁner P (2012) Geostatistics: Modeling Spatial Uncertainty, 2nd edition. New Jersey: Wiley.
38. Asyali MH, Alci M (2005) Membership functions and probability measures of fuzzy sets. Journal of the American Statistical Association 99:867-877.
39. Iannaccone PM, Khokha M (1995) Fractal Geometry in Biological Systems: An Analytical Approach. Boca Raton: CRC Press.
40. Castellano G, Bonilha L, Li LM, Cendes F (2004) Texture analysis of medical images. Clinical Radiotherapy 59:1061–1069.
41. Ploisi ME, Nikos C, Charchantzi A (2013) Combining shape, texture and intensity features for cell nuclei extraction in Pap smear images. Pattern Recognition Letters 32:838-853.
42. Guer S, Xiang AS, Yarborough WG, Weaver AM (2007) Cortactin is an essential regulator of matrix metalloproteinase secretion and extracellular matrix degradation in invadopodia. Cancer Research 67:4227-4235.
43. Nguyen-Thanh N, Pham TD, Ichikawa K (2013) Segmentation of mitochondria in intracellular space. Proc. IEEE Symposium Series on Computational Intelligence: 218–221.
44. Pham TD, Ichikawa K (2013) Spatial chaos and complexity in the intracellular space of cancer and normal cells. Theoretical Biology and Medical Modelling 10:62. DOI: 10.1186/1742-4692-10-62.
45. Pham TD, Ichikawa K, Nakajima K, Nakajima K (2013) Establishment and Validation of Computational Model for MT1-MMP Dependent ECM Degradation and Intervention Strategies. PLoS Comput Biol 9(4): e1002479. DOI:10.1371/journal.pcbi.1002479.
46. Pham TD (2014): Automated identiﬁcation of mitochondrial regions in complex intracellular space by texture analysis. Proc. 5th Int. Conf. Graphics and Image Processing 90600G: doi:10.1117/12.2050102.
47. Pham TD, Le DTP, Xu J, Nguyen DT, Martinugale RD, Deveney CW (2014) Personalised identiﬁcation of abdominal wall hernia meshes on computed tomography. Computer Methods and Programs in Biomedicine 113:153-161.
48. Nixon M, Aguado A (2008) Feature Extraction & Image Processing, 2nd Edition. Amsterdam: Academic Press.

Table 4. Total Average Logo-Image Classification Rates (%).

| Method       | Semi-variogram | Zernike | Wavelets | Gabor | PEI1 | PEI2 |
|--------------|----------------|---------|----------|-------|------|------|
|              | 69.77          | 61.50   | 47.40    | 50.21 | 72.89| 72.89|

doi:10.1371/journal.pone.0105075.t004

The logo data of the University of Maryland are publically available [56]. The FCM algorithm was implemented using the Fuzzy Logic Matlab toolbox. Matlab codes for kriging algorithms can be obtained at: http://mgstat.sourceforge.net/ (A Geostatistical Matlab Toolbox).
49. Pham TD (2003) Unconstrained logo detection in document images. Pattern Recognition 36: 3023–3025.
50. Pham TD (2003) Applications of geostatistics and Markov models for logo recognition. Proc. SPIE Conf. Document Recognition and Retrieval X 5010: 20–27.
51. Pham TD (2014) Pattern recognition by active visual information processing in birds. Information Sciences 270: 134–142.
52. Pham TD (2003) Variogram-based feature extraction for neural-network recognition of logos. Proc. SPIE Conf. Applications of Artificial Neural Networks in Image Processing VIII 5015: 12–22.
53. Wang Z, Xia M, Jin Z, Yao L, Long Z (2014) Temporally and spatially constrained ICA of fMRI data analysis. PLoS ONE 9: e94211. DOI:10.1371/journal.pone.0094211
54. Pham TD, Berger K (2011) Automated detection of white matter changes in elderly people using fuzzy, geostatistical, and information combining models. IEEE Trans Information Technology in Biomedicine 15: 242–250.
55. Liu J, Pham TD (2012) A spatially constrained fuzzy hyper-prototype clustering algorithm. Pattern Recognition 45: 1759–1771.
56. DATASET: Logo Database for Research. Available: http://lampsrv02.umiacs.umd.edu/projdb/project.php?id = 47. Accessed 2003 January 20.