IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 14, 2020, accepted January 12, 2021, date of publication January 14, 2021, date of current version January 25, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3051758

A Competitive Generalized Gamma Mixture Model for Medical Image Diagnosis

SAMI BOUROUIS[1], HASSEN SALLAY[2], AND NIZAR BOUGUILA[3], (Senior Member, IEEE)

[1]College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia
[2]College of Computer and Information Systems, Umm Al-Qura University, Makkah 21955, Saudi Arabia
[3]Concordia Institute for Information Systems Engineering (CIISE), Concordia University, Montreal, QC H3G 1T7, Canada

Corresponding author: Hassen Sallay (hmsallay@uqu.edu.sa)

This work was supported in part by the Research and Development Grants Program for National Research Institutions and Centers (GRANTS), the Target Research Program, the Infectious Diseases Research Grant Program, and the King Abdulaziz City for Science and Technology (KACST), Kingdom of Saudi Arabia, under Grant 5-20-01-007-0010.

ABSTRACT Parametric family of statistical distributions are of great importance for several applications. In particular, we propose to investigate the generalized Gamma mixture model (gΓMM) for modeling and classifying medical imaging (Chest x-ray and CT-scans). The main advantage of this mixture over some existing Gaussian models is that it allows additional flexibility in shape modeling, which is crucial for classification systems. In order to capture accurately the intrinsic nature of medical images, we propose to derive some efficient measures based on Fisher, Kullback-Leibler and Bhattacharyya distances for the mixtures of generalized Gamma distributions. Indeed, the main idea is to investigate these distances effectively via the statistical model parameters in order to make our proposed scheme particularly appropriate for image classification problem. The proposed approach involves the extraction of robust texture descriptors, the learning of mixture model gΓMM via the expectation-maximization (EM) and Newton-Raphson algorithms, and the classification of images using the derived mixtures-based distances. We evaluate our model against the challenging problem of early diagnosis of pneumonia diseases. Experimental results on different datasets show the merits of our developed framework compared with the other methods.

INDEX TERMS Mixture models, generalized Gamma distribution, Fisher distance, Kullback-Leibler distance, Bhattacharyya distance, chest x-ray images (CXR) classification.

I. INTRODUCTION

Chest X-ray (CXR) imagery is one of the cheapest and affordable procedures in the majority of hospitals and clinical centers. It is proved to be fast, to provide sensitive diagnostic standards and suitable for identifying lung abnormalities caused by many pneumatic diseases. In particular, [1] reported the efficiency of using CXR for identifying nowadays Corona virus (Covid-19) pandemic disease [2]. As known, the first wave of new Covid-19 locked most countries and caused about a million deaths in the world, the second wave comes with more violent and greater infection spreading and mortality rates worldwide. While several techniques were developed to detect the infection, they remain expensive and not affordable by several middle-and-low income countries. Investigating these kinds of images is of high importance and can help in the early detection of the infection and the efficient triage of patients mainly in the very crowded cities with cost-effective manner and therefore slow down its spreading worldwide. To do so, it is necessary to develop smart tools based on medical images since the accurate classification of input images as normal or infected and the recognition of specific disease symptoms' is a very valuable step for infection detection. However, analyzing accurately and automatically a huge number of medical images in real time is too difficult due to noisy images, the lack of contrast between tissues and the lack of experience of many radiologists. Moreover, the appearance symptoms in images is not the same for different stages of disease (i.e. early, middle, or advanced). For example, for the case of an advanced stage, the covid-19 symptoms may be exhibited clearly compared to the early and intermediate stages. In other words, these symptoms can resemble other viral infections like RSV pneumonia at early stage.

The associate editor coordinating the review of this manuscript and approving it for publication was Yin Zhang.

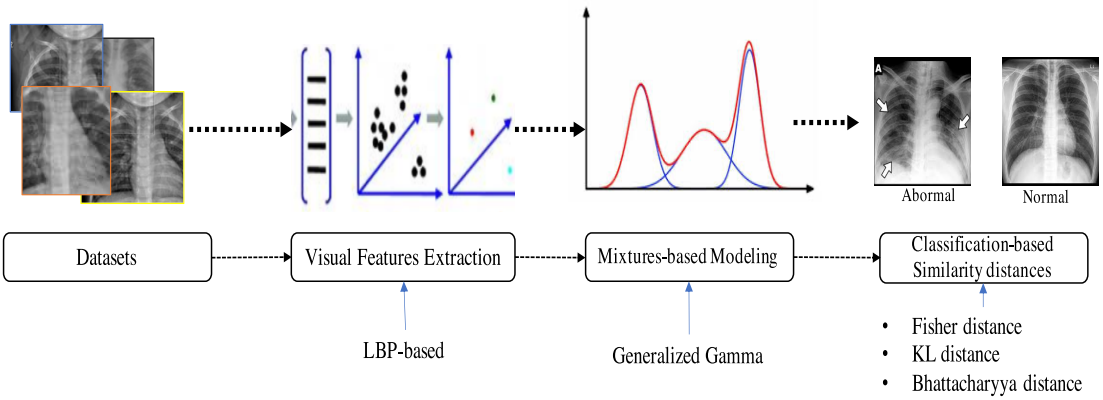


FIGURE 1. The pipeline of proposed method. First, robust visual features are extracted from each CXR image. Then, features are modelled using the mixture model (gΓMM). Finally, images are classified on the basis of different distances between mixtures.

Several promising biomedical image-based techniques have been developed in the past decades to deal especially with pneumonia infection detection [3]. In particular, many image processing-based systems have been developed to deal with infection localization, detection and quantification. Among the statistical based systems, the ''generalized Gamma mixture model (gΓMM)'' has been proven to be competitive, and works better than many other formerly proposed statistical models especially for complex pattern characterized images. The main advantage of the generalized Gamma distribution is that it is very attractive given its compact analytical form and its ability to model different family of distributions. In our case, gΓMM can be shown as a suitable choice for statistically describing the visual features of chest x-ray images thanks to its high flexibility and its ability to represent a rich family of distributions and models. Our work here is motivated by the interesting results obtained previously with the generalized Gamma mixture for data clustering and characterization. Thanks to these merits, we propose to exploit this mixture model (gΓMM) for covid-19 detection using medical imaging. It is noteworthy that generative probabilistic models have many benefits and offer principled way for incomplete data modelling, but the classification rate is not always higher especially for complex data. To deal with these limitations, we propose in this work to capture accurately the intrinsic nature of medical images by deriving some efficient measures based on Fisher, Kullback-Leibler and Bhattacharyya distances for the mixtures of generalized Gamma distributions. Indeed, the main idea is to investigate these distances effectively via the gΓMM model's parameters in order to make our proposed scheme particularly suitable for image classification systems. The proposed pipeline is shown in Fig.1.

The rest of this paper is organized as follows. Section 2 presents some related work. Section 3 describes the proposed generalized Gamma mixture model as well as its parameters estimation. Section 4 presents our image classification based on similarity measurement between mixtures of generalized Gamma mixtures. Section 5 evaluates the performance of the proposed model, and shows the experimental results using different real datasets. Finally, the last Section is devoted to the conclusion and future works.

II. RELATED WORKS

In the context of image processing and pattern recognition, non-Gaussian data modelling plays an essential role for accurate data clustering and classification. This problem can be addressed with finite mixture models (FMM) which are designed as well principled and successful statistical approaches in view of its ability to describe multimodal distributions [4]–[14]. For instance, finite Gaussian mixture models (GMM) are employed to detect lung cancer in CT images [15]. FMM have shown to be helpful for medical diagnosis and can assist to identify a variety of infections and hence enhance the health care [13], [16], [17]. Obviously, the main issue related to the exploiting of mixture models for data classification is the choice of the component's densities which determines the flexibility of mixture models. For instance, Laplace and Normal densities have applied with success in the past especially for the case of low or medium-resolution medical images [18]–[22]. However, these densities suffer from several drawbacks and cannot be always the best approximations for complex data modelling such as X-rays and CT scan images. In fact, there are many medical phenomena for which these densities cannot be realistic. In addition, they cannot guarantee the convergence to optimal solution, thus resulting in the difficulty of effectively classifying such complex images. Furthermore, more advanced statistical models have been shown to be more suitable for complex-real images. This is the case of non-Gaussian mixture models such as Gamma distribution that performs substantially better than conventional models [23], [24] especially for the case of SAR images analysis [7], [12], [25], [26]. However, Gamma distribution fails when describing the heavier tails caused by specific complex patterns in such images and when

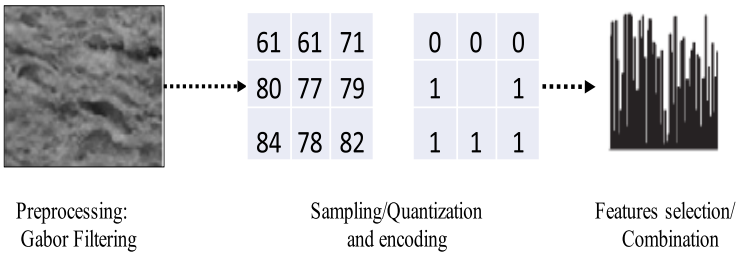


FIGURE 2. The basic pipeline for Local Binary Pattern (LBP) features extracted from gray scale images.

dealing with multi-dimensional data like the classification of high-resolution SAR images [27] and the characterization of ultrasonic tissue [28]. In light of the previous limitations which could affect the accuracy of the multidimensional data modeling, we are motivated by (gΓMM) flexible model for classifying x-ray images. This passes through investigating the robustness of this non-Gaussian model for multidimensional data modelling and then to exploit it in the context of discriminative framework for accurate medical image classification.

III. LBP-BASED FEATURES EXTRACTION

A crucial step for computerized analysis of medical image is the informative visual features extraction. Recently, this step has become one of the main subjects in diagnostic procedures. It is noted that some image modalities like chest x-rays are difficult to interpret by radiologists, consequently, it is required to identify important local or global patterns to understand better their content. For the case of medical images, the characterization of appearing infections and the accurate extraction of relevant details in x-ray images is of great importance for further medical analysis steps and for making decision. Textures are one of the most essential visual features used in the context of medical image analysis thanks to important information they contain. In particular, texture-based classification plays a great part in computer vision and pattern recognition applications. In this work we proceed with texture features which is one of the main factors of human visual perception. Some texture-based descriptors are developed and are invariant to many geometric transformations. It is possible to extract texture features from images with different techniques, including statistical methods that often rely on higher order statistics which allow different measurements to be accurately calculated [29]–[31]. In this study, we are primarily motivated by local image information that describes image in more details. One of the popular and widely used scale invariant local texture descriptor is named Local Binary Pattern (LBP) [32]. LBP has the advantage to define the local contrast and spatial structure of a part of an image. Unlike GLCM-based texture features (such as Haralick) which take into account the global texture representation and describe the entire image in an average manner, LBP quantifies local information (extracted from small sub-images) based on the current pixel value and its neighborhood. As a result, LBP will be more robust to invariant texture analysis for many applications notably medical image analysis. This descriptor has undergone extensive development and application in the context of pattern recognition and image processing. Many researchers have considered LBP as a commonly applied texture descriptor since it is able to help classifiers give results with high accuracy thanks to its simple implementation. The key idea of LBP operator is to transform the input image into a list of integer labels describing its small-scale appearance. In practice, it comes out to apply a circular neighborhood strategy for each central point (*CP*) with different size. Then, we calculate the LBP value that encodes the local structure of this pixel by comparing the value of the *CP* with its neighbors. In other word, this results in testing whether the neighbors are less than or greater than the current *CP* (i.e. the results is a binary value 0 or 1). This process is repeated for all key points of interest in the input image. Basically, most LBP-based feature extraction techniques [33] follow the steps summarized in Fig. 2.

IV. GENERALIZED GAMMA MIXTURE LEARNING

In this section, we present the developed deterministic learning algorithm based on the generalized gamma mixture model. Indeed, in mixture modeling, we suppose that our observed data is generated from a mixture of sub populations. Given an observed data set $\mathcal{Y}$ with N data instances $\mathcal{Y} = \{\vec{Y}_1, \ldots, \vec{Y}_N\}$, where each $\vec{Y}_i = (Y_{i1}, Y_{i2}, \ldots, Y_{i,D})$ is a D -dimensional positive vector that follows a mixture of multi-dimensional generalized gamma distributions.

A. STATISTICAL MODEL SPECIFICATION

Let M denotes the number of clusters, N the number of feature vector, and a set of features vectors $\mathcal{Y}$ where each Y_i $(i = 1, \ldots, N)$ is generated from a generalized Gamma mixture model with parameter Θ, then, the mixture model is expressed as follows:

$$p(\vec{Y}_i \mid \Theta) = \sum_{j=1}^{M} \pi_j p(\vec{Y}_i \mid \theta_j) \tag{1}$$

where $\Theta = \{\theta_1, \theta_2, \ldots, \theta_M, \pi_1, \ldots, \pi_M\}$ and $\theta_j = \{\alpha_j, \beta_j, \lambda_j\}$ is the set of parameters of the distribution related to the class j. Here π_j denotes the mixing proportion parameter where $0 \leqslant \pi_j \leqslant 1$, and $\sum_{j=1}^{M} \pi_j = 1$. For a selected cluster j, each $\vec{Y}_i$ follows a generalized Gamma distribution with the following probability density function:

$$p(\vec{Y}_i \mid \theta_j) = \prod_{d=1}^{D} p(X_{id} \mid \theta_{jd}) = \prod_{d=1}^{D} \frac{\frac{\lambda_{jd}}{\beta_{jd}} X_{id}^{\beta_{jd}-1} \exp -(\frac{X_{id}}{\alpha_{jd}})^{\lambda_{jd}}}{\Gamma(\frac{\beta_{jd}}{\lambda_{jd}})} \quad (2)$$

where $\Gamma(.)$ is the Gamma function and the parameter $\theta_{jd} = (\alpha_{jd}, \beta_{jd}, \lambda_{jd})$ $)d = 1, \ldots, D)$. Now, for M components, we can express the mixture of generalized Gamma distributions as:

$$p(\mathcal{Y} \mid \Theta) = \prod_{i=1}^{N} \sum_{j=1}^{M} \pi_j p(\vec{Y}_i \mid \theta_j) \quad (3)$$

$Z_i = (Z_{i1}, \ldots, Z_{iM})$ are the latent membership vectors that indicates if the vector $\vec{Y}_i$ belongs to cluster j ($Z_{ij} = 1$) or not ($Z_{ij} = 0$). Thus, the complete-data likelihood of our mixture model is given as:

$$p(\mathcal{Y}, Z \mid \Theta) = \prod_{i=1}^{N} \prod_{j=1}^{M} \pi_j p(\vec{Y}_i \mid \theta_j)^{z_{ij}} \quad (4)$$

B. PARAMETERS ESTIMATION

In this section, we estimate the parameters of the statistical mixture model based on the well known Expectation–Maximization algorithm [34]. These parameters are estimated using the maximum likelihood estimator. Indeed, setting the derivate of the complete log-likelihood with respect to the parameters will lead to find the appropriate equation of each parameter. Thus, we need to maximize the following log-likelihood function:

$$\log p(\mathcal{Y}, Z \mid \Theta) = \sum_{i=1}^{N} \sum_{j=1}^{M} Z_{ij} \log \pi_j p(\vec{Y}_i \mid \theta_j) \quad (5)$$

As a result we obtain the log likelihood function given by.

$$\log p(\mathcal{Y}, Z \mid \Theta) = \sum_{i=1}^{N} \sum_{j=1}^{M} Z_{ij} \log \pi_j p(\vec{Y}_i \mid \theta_j) + \Lambda(1 - \sum_{j=1}^{M} \pi_j) \quad (6)$$

By setting the derivate of the complete log-likelihood with respect to the mixing weight and the model's parameters $\pi_j, \alpha_{jd}, \beta_{jd}, \lambda_{id}$, respectively, we obtain the update equation of the mixing weight as follows:

$$\pi_j = \frac{\sum_i^N p(j \mid \vec{Y}_i)}{N} \quad (7)$$

where $p(j \mid \vec{Y}_i)$ is the posterior distribution that defines the probability of affecting each vector X_i to the appropriate cluster j. In this case, the posterior probability function $\hat{Z}_{ij} = p(j \mid \vec{Y}_i)$ is given as:

$$\hat{Z}_{ij} = p(j \mid \vec{Y}_i) = \frac{\pi_j p(\vec{Y}_i \mid \theta_j)}{\sum_{j=1}^{M} \pi_j p(\vec{Y}_i \mid \theta_j)} \quad (8)$$

Regarding the estimation of the other parameters, we compute the second derivative of $logp(\mathcal{Y} \mid \Theta)$ with respect to α_{jd} and β_{jd}, and we obtain respectively the following:

$$\alpha_{jd} = \left(\frac{\lambda_{jd}}{\beta_{jd}} \frac{\sum_{i=1}^{N} p(j \mid \vec{Y}_i) X_{id}^{\lambda_{jd}}}{N \pi_j} \right)^{\frac{1}{\lambda_{jd}}} \quad (9)$$

$$\beta_{jd} = \lambda_{id} \Psi^{-1} \left(\lambda_{jd} \left(\frac{\sum_{i=1}^{N} p(j \mid \vec{Y}_{id}) \log(X_{id})}{\sum_{i=1}^{N} p(j \mid \vec{Y}_{id})} - \log(\alpha_{jd}) \right) \right) \quad (10)$$

where $\Psi^{-1}(.)$ is the inverse digamma function.

Now, regarding the parameter λ, it has not a closed form because it is coupled with the other parameters. To solve this issue, we apply an iterative algorithm based on Newton-Raphsen method. The new expression of$\lambda_{jd}$is determined as:

$$\lambda_{jd}^{new} = \lambda_{jd}^{old} - \gamma \frac{\partial \log p(\mathcal{Y} \mid \Theta)}{\partial \lambda_{jd}} \left(\frac{\partial^2 \log p(\mathcal{Y} \mid \Theta)}{\partial^2 \lambda_{jd}} \right)^{-1} \quad (11)$$

where γ is the constant step size.

C. THE COMPLETE LEARNING ALGORITHM

In the following, we summarize our algorithm used for the learning and that take into account for model complexity. Indeed, the determination of the model complexity M is a crucial step in mixture modeling problems allowing us to provide better generalisation capabilities. For this purpose, we use the Minimum Message Length criterion (MML) that has been used successfully in several approaches [35]–[37]. This criterion is based on information theory that is used to evaluate statistical models according to compress a message containing the data. MML has the advantage to estimate the accurate optimal number of components in the mixture through minimizing the following function:

$$MML(\Theta, Y) = -log(p(\Theta)) - L(\Theta, Y) + \frac{1}{2} log|F(\Theta)| + \frac{N_p}{2} - \frac{1}{2} log(12) \quad (12)$$

where $p(\Theta)$ is the prior probability, $|F(\Theta)|$ is the determinant of the Fisher information matrix, and N_p is the number of parameters.

In the following, we choose a prior distribution, $p(\Theta)$, and derive an expression for $|F(\Theta)|$. It is noted that the parameters of different mixture' components are independent since there is no knowledge from a parameter of one class to provide for other classes. Thus, we have the following:

$$p(\Theta) = p(\alpha)p(\beta)p(\lambda)p(\pi) \quad (13)$$

Algorithm 1 Proposed Learning Model

Input : Data set $\mathcal{Y}$, M_{max}
Output: K^*, Θ^*
foreach $1 \leq K \leq M_{max}$ **do**
  Initialization algorithm :
  Apply K-Means to initialize the parameters of each component.
  repeat
    for $0 \leq j \leq M$ **do**
      E-step :
      Compute posterior probabilities using the following: $\hat{Z}_{ij} = p(j \mid \vec{Y}_i) = \frac{\pi_j p(\vec{Y}_i|\theta_j)}{\sum_{j=1}^{M} \pi_j p(\vec{Y}_i|\theta_j)}$
      M-step :
      Update the parameter α_j using Eq (9).
      Update the parameter β_j using Eq (10).
      Update the parameter λ_j using Eq (11).
    end
  until *Convergence of Likelihood*
  Calculate the associated message length using Eq(12).
  Save Θ, M and the MML of each model
end
Return the optimal model M_{opt} with the minimum MML criterion.

For the prior distribution of $p(\pi)$, we take a Dirichlet prior. This prior is given by:

$$p(\pi) = Dir(\pi|u_0) = \frac{\Gamma(\sum_{i=1}^{M} u_0)}{\prod_{i=1}^{M} \Gamma(u_0)} \prod_{i=1}^{M} \pi^{u_0-1} \tag{14}$$

For the parameter α, β, and λ, we adopt the Gamma prior distributions as follow:

$$\begin{aligned} \alpha &\sim Gamma(a_0, b_0) \\ \beta &\sim Gamma(a_1, b_1) \\ \lambda &\sim Gamma(a_2, b_2) \end{aligned} \tag{15}$$

where $Gamma(x; a, b) = \frac{1}{\Gamma(a)} b^a x^{a-1} e^{-bx}$

The Fisher information matrix is the determinant of the Hessian matrix of $-L(Y|\Theta)$. Given the complicated analytical form of this quantity for the case of mixtures, we proceed with an approximate form of the determinant of Fisher Information matrix which is determined as follows:

$$|F(\Theta)| = |F(\pi)| \prod_{j=1}^{M} |F(\alpha_j)||F(\beta_j)||F(\lambda_j)| \tag{16}$$

V. IMAGE CLASSIFICATION BASED ON SIMILARITY MEASUREMENT BETWEEN gΓMM

The similarity between images (or mixtures) is measured by evaluating the inconsistencies or similarities between the mixtures. In this work, we apply the developed statistical model gΓMM to represent the histogram of textures and to compare between images. In other words, similarity quantity between mixtures gΓMM is one of the main questions for the image classification tools. Given the model gΓMM which is characterized by Θ (our model parameter). In this case gΓMM is sensitive to selected values of its parameter that is a little modification in the parameter can cause a marked difference in its distribution). Therefore, it is clear that we coul not directly measure the distance between parameters. To overcome this problem and to fully exploit the advantages of the model gΓMM (i.e. its parameters), we will instead consider some effective similarity measures such as Fisher, Symmetrized Kullback-Leibler Distance (SKLD), and Bhattacharyya distances for similarity measurement between different mixtures.

A. FISHER DISTANCE (FD)

Fisher distance [38] is one of the efficient techniques to measure the similarity between two distributions or two mixtures. In our case, we address the general case (i.e. mixtures), and thus similar mixtures have same log-likelihood gradients in the mixture space. To use Fisher distance, we need first to derive the log-likelihood gradient of each generalized Gamma mixture model (gΓMM). Then, similarity between two mixtures (Y and Y') is determined as:

$$\mathcal{FD}(Y, Y') = U_Y^{tr}(\Theta) I^{-1}(\Theta) U_{Y'}(\Theta') \tag{17}$$

The derivatives of the log-likelihood (i.e the gradient of $\log(p(\mathcal{Y}|\Theta))$) with respect to $\Theta = (\pi, \alpha, \beta, \lambda)$ are expressed after simplification as:

$$U_Y(\Theta) = \nabla \log(p(\mathcal{Y}|\Theta)) = \frac{\partial \log(p(\mathcal{Y}|\Theta))}{\partial \Theta} \tag{18}$$

$$\frac{\partial \log(p(\mathcal{Y}|\Theta))}{\partial \pi_j} = \sum_{i=1}^{N} \left[\frac{\hat{Z}_{ij}}{\pi_j} - \frac{\hat{Z}_{i1}}{\pi_1} \right] \tag{19}$$

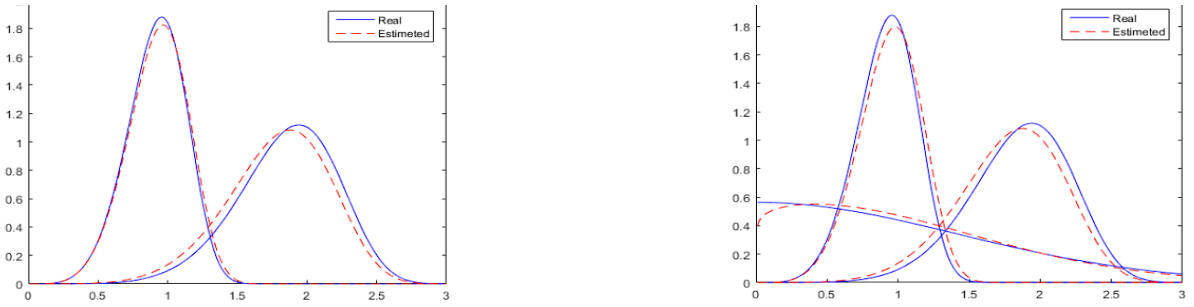
$$\frac{\partial \log(p(\mathcal{Y}|\Theta))}{\partial \alpha_j} = \sum_{i=1}^{N} \hat{Z}_{ij} \left[-\frac{\beta_j}{\alpha_j} + \frac{\lambda_j Y_i^{\lambda_j}}{\alpha_j^{\lambda_j+1}} \right] \tag{20}$$

$$\frac{\partial \log(p(\mathcal{Y}|\Theta))}{\partial \beta_j} = \sum_{i=1}^{N} \hat{Z}_{ij} \left(log(Y_i) - log(\alpha_j) - \frac{\psi(\frac{\beta_j}{\lambda_j})}{\lambda_j} \right) \tag{21}$$

$$\frac{\partial \log(p(\mathcal{Y}|\Theta))}{\partial \lambda_j} = \sum_{i=1}^{N} \hat{Z}_{ij} \left[\frac{1}{\lambda_j} - \left(\left(\frac{Y_i}{\alpha_j} \right)^{\lambda_j} log \left(\frac{Y_i}{\alpha_j} \right) \right) + \frac{\psi(\frac{\beta_j}{\lambda_j})\beta_j}{\lambda_j^2} \right] \tag{22}$$

B. SYMMETRIZED KULLBACK-LEIBLER DISTANCE (SKLD)

The second efficient solution is called Symmetrized Kullback-Leibler distance (SKLD) [39]. This distance is derived from the KL divergence and has the advantage of offering a more balanced measurement than the asymmetrized KLD, which makes it more appropriate for classification task. In order to derive the SKLD distance, we first



(a) Two generalized Gamma mixtures (b) Three generalized Gamma mixtures

FIGURE 3. Fitting real data (blue plot) with generalized Gamma mixture model (red plot).

need to calculate the asymmetrized-KLD between two mixtures p_1 and p_2 using the following equation:

$$\begin{aligned}&\mathcal{SKLD}(p_1(\vec{Y}|\Theta_1), p_2(\vec{Y}'|\Theta_2))\\&\quad = \mathcal{KLD}(p_1(\vec{Y}|\Theta_1), p_2(\vec{Y}'|\Theta_2))\\&\qquad +\mathcal{KLD}(p_2(\vec{Y}'|\Theta_2), p_1(\vec{Y}|\Theta_1)) \quad (23)\\&\mathcal{KLD}(p_1(\vec{Y}|\Theta_1), p_2(\vec{Y}'|\Theta_2))\\&\quad = e^{-BF(p_1(\vec{Y}|\Theta_1), p_2(\vec{Y}'|\Theta_2))} \quad (24)\end{aligned}$$

where B is a real positive factor used for computational stability purpose, and

$$\begin{aligned}&F(p_1(\vec{Y}|\Theta_1), p_2(\vec{Y}'|\Theta_2))\\&= \int_{\omega} p_1(\vec{Y}|\Theta_1) \log \frac{p_1(\vec{Y}|\Theta_1)}{p_2(\vec{Y}'|\Theta_2)} + p_2(\vec{Y}'|\Theta_2) \log \frac{p_2(\vec{Y}'|\Theta_2)}{p_1(\vec{Y}|\Theta_1)}\end{aligned}$$

Noted here that is not impossible to find a closed expression for the KL divergence, therefore we choose to apply an effective approximate based on the so-called Monte Carlo method [40], [41].

$$\mathcal{KLD}(p_1(\vec{Y}|\Theta_1), p_2(\vec{Y}'|\Theta_2)) \approx \frac{1}{L} \sum_{i=1}^{L} \log \frac{p_1(\vec{Y}_i|\Theta_1)}{p_2(\vec{Y}'_i|\Theta_2)} \quad (25)$$

C. BHATTACHARYYA DISTANCE (BD)

The third distance used to compare between mixtures is called "Bhattacharyya distance" [42]. This distance is defined as:

$$\mathcal{BD}_{\frac{1}{2}}(\vec{Y}_1, \vec{Y}_2) = \int_0^{\infty} p(\vec{Y}|\Theta_1)^{1/2} q(\vec{Y}|\Theta_2)^{1/2} d\vec{Y} \quad (26)$$

Here also, an analytical expression for the Bhattacharyya distance is intractable given that we deal with mixtures of generalized gamma distributions. Therefore, we proceed like the case of symmetrized Kullback-Leibler (KL) distance by approximating a Bhattacharyya distance using an efficient approximate based on the Monte Carlo simulation technique [40], [41]. The distance is then expressed as:

$$\begin{aligned}\mathcal{BD}_{\frac{1}{2}}(\vec{Y}_1, \vec{Y}_2) \approx &\frac{\beta}{N_1} \sum_{i=1}^{N_1} \frac{p^{1/2}(\vec{Y}_i|\Theta_1)}{Z_1} p^{1/2}(\vec{Y}_i|\Theta_1)\\&+\frac{1-\beta}{N_2} \sum_{i=1}^{N_2} \frac{q^{1/2}(\vec{Y}_i|\Theta_2)}{Z_2} q^{1/2}(\vec{Y}_i|\Theta_2) \quad (27)\end{aligned}$$

where $\beta \in [0, 1]$ and the normalized factors Z_1, Z_2 are used for the densities p and q.

VI. EXPERIMENTAL RESULTS

In order to show the effectiveness of the statistical mixture model in terms of goodness-of-fit with respect to data histogram, we provide some plots in Fig 3. These results are conducted on one-dimensional (D = 1) data sets generated from both two and three generalized Gamma densities (M = 2, N = 100). These plots show clearly that the proposed model fits well the data.

For real data sets, the implemented approach includes first a visual features extraction step through LBP detector. Indeed, each image is encoded with and LBP features vector and then modelled through the mixture gΓMM. Then, we measure the similarity between images (i.e. mixtures) using the different implemented distances. Indeed, the main idea is to investigate these distances effectively via the model parameters in order to make our proposed scheme particularly appropriate for image classification (or covid-19 detection) systems with various image databases. The statistical models are built to distinguish the normal and COVID affected patients using Chest X-ray and CT scan images. We performed different image processing steps. After processing the data, we extracted a list of statistical parameters. The majority of studies have shown that the primary organs that are affected by this disease are the lungs. In our analysis, we focused on extracting the lungs area using image thresholding and segmentation processing, we identified and isolated left and right lungs from Chest X-ray images using the method in [43]. In order to remove noise from the image, we applied the Gaussian filter. In our experiments, we run the Monte-Carlo approximation algorithm within 10.000 samples.

A. X-RAY IMAGES CLASSIFICATION FOR COVID-19 DETECTION

As mentioned before, we experiment our model on the challenging Covid-19 infection detection context via processing Chest X-ray images. The example in Fig 4 depicts an example of Chest X-ray image that contains covid-19. The main source of COVID-19 images is the available dataset[1] made

[1] https:https://github.com/ieee8023/covid-chestxray-dataset

(a) Normal case (b) COVID-19 case

FIGURE 4. Chest X-ray images of (a) normal and (b) COVID-19 patients.

TABLE 1. CXR data statistics.

Data/Class	Train	Validation	Test	Total
Non-COVID-19	70	20	18	108
COVID-19	328	80	26	434
Total	398	100	44	542

TABLE 2. Overall accuracy for CXR-COVID Dataset.

Approach/Metrics	ACC(%)	DR(%)	FPR(%)
GMM	82.11	81.02	0.18
ΓMM	85.22	83.76	0.16
GMM-FD	83.43	82.29	0.17
GMM-KD	83.27	82.20	0.17
GMM-BD	83.25	82.18	0.17
ΓMM-FD	86.01	84.11	0.16
ΓMM-KD	85.99	84.08	0.16
ΓMM-BD	85.94	84.03	0.16
gΓMM-FD	89.01	87.90	0.12
gΓMM-KD	87.71	87.01	0.12
gΓMM-BD	87.67	86.96	0.12

by Cohen *et al.* [44], [45]. This dataset was developed by gathering several Chest X-ray (CXR) images. At the time of writing, this dataset comprises in total 542 images with different type of pneumonias. Currently, a subset of 434 CXR images of patients potentially positive to COVID-19 and the rest are COVID-19 negative. Metadata are provided for every image, including several information such as the patient ID, the location and other annotations. Details related to the number of images are given in table 1.

For performance investigation, we run the three learning approaches for finite generalized Gamma mixture model and we evaluated their performance in terms of overall accuracy (Acc), detection rate (DR), and false-positive rate (FPR). For this experiment, the selected value of B for the Symmetrized Kullback-Leibler distance is equal to 2. Table 2 shows the obtained results for the CXR-COVID data sets when applying different learning approaches namely Gamma mixture (**ΓMM**), Gaussian mixture(**GMM**), Gaussian mixture(**GMM**) with different distances (Fisher, symmetrized Kullback-Leibler, Bhattacharyya), Gamma mixture with different distances, and with generalized Gamma mixture with Fisher distance (**gΓMM-FD**), with symmetrized Kullback-Leibler distance (**gΓMM-KD**), and with Bhattacharyya distance (**gΓMM-BD**). Please note that our implementation tackles the problem of grouping images in an unsupervised way which is more difficult than the supervised one. According to the results, we can see clearly that the proposed generalized Gamma mixture model provides better results using the three distances-based measures, as compared to Gaussian-based and Gamma-based models. In particular, the Fisher distance allows to provide the best result with overall accuracy (Acc = 89.01%), however with the **GMM** and **ΓMM** it is about 82.11% and 85.22%, respectively. Finally, we can explain the fact that these results are not really that high (for this particular dataset) since we are not dealing with a large data set.

TABLE 3. Augmented data statistics.

Data/Class	Train	Validation	Test	Total
Non-COVID-19	512	100	300	912
COVID-19	512	100	300	912

B. AUGMENTED X-RAY IMAGES CLASSIFICATION FOR COVID-19 DETECTION

We conducted also our experiments on several other augmented chest x-ray images from the publicly dataset "Augmented COVID-19 Dataset".[2] This dataset contains augmented X-ray images for COVID-19 disease detection and the number of images is more important than the previous dataset. Indeed, the size of the new covid-19 images is increased from 48 to 912 images which make the dataset appropriate for learning. The augmentation process is considered very important since the number of covid-19 samples is low, and this process includes some image transformations such as rotation, flipping, translation, and scaling. The augmented dataset is collected from two online available datasets Kaggle[3] and.[4] Table 3 describes this dataset. For this application, the value of the parameter B for the SKLD distance is selected experimentally and it is equal to 8. Table 4 depicts the classification accuracies for this repository when applying different generative approaches. Accordingly, we can see clearly that our proposed approach with the different measures outperforms both the Gaussian and Gamma models. Indeed, our work has better accuracy as well as lowest false positive rate than these models. We can also notice that the Fisher distance is slightly better (Acc = 94.08%) than the KL-based (Acc = 93.90%) and Bhattacharyya-based distances (Acc = 93.88%), respectively. These results encourage the choice for flexible generalized Gamma mixtures with different distances which are able to improve the

[2]https://data.mendeley.com/datasets/2fxz4px6d8/4
[3]https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
[4]https:https://github.com/ieee8023/covid-chestxray-dataset

TABLE 4. Overall accuracy for CXR-Augmented-COVID Dataset.

Approach/Metrics	ACC(%)	DR(%)	FPR(%)
GMM	85.13	83.99	0.14
ΓMM	90.24	89.14	0.10
GMM-FD	85.98	84.38	0.14
GMM-KD	85.91	84.41	0.14
GMM-BD	85.88	84.35	0.14
ΓMM-FD	91.13	89.86	0.10
ΓMM-KD	91.09	89.80	0.10
ΓMM-BD	91.03	89.77	0.10
gΓMM-FD	94.08	93.17	0.09
gΓMM-KD	93.90	93.09	0.09
gΓMM-BD	93.88	93.08	0.09

TABLE 5. Pneumonia data statistics.

Data/Class	Train	Validation	Test	Total
Normal	1341	8	234	1583
Pneumonia	3875	8	390	4273
Total	5216	16	624	5856

classification accuracy. Moreover, the size of this dataset which is increased compared to the previous one, the flexibility of the generalized Gamma mixture, and the effectiveness of texture-based features lead to more stable results here.

C. MULTI-LABEL PNEUMONIA CLASSIFICATION

In this work we also addressed the issue of detecting pneumonia. For this purpose, we conducted several experiments on the publicly repository Kaggle. This dataset contains two categories (Pneumonia/Normal) for lung and it is structured into 3 folders (train, test, val). It is noted that pneumonia is a lung inflammation caused by a viral or bacterial infection. The current dataset includes 1583 normal CXR cases and 4273 pneumonia images, for a total of 5856 images. The number of the images used in the experiment from this dataset is depicted in table 5. Table 6 shows the obtained results for all methods. We note here that the value of the parameter B for the SKLD distance is equal to 8. It is clear here that all mixture models are able to provide encouraging good results. Likewise, we came to the same conclusion with this dataset and we obtain the highest accuracy (ACC about 92%) with our approach. We note also that the precision increases (and the false positive decreases) as the dataset size increases given that the current dataset (CXR-Pneumonia) contains more images than the previous ones (i.e. CXR-COVID and CXR-Augmented-COVID). As a result, it is obviously clear to confirm the merits of our proposed framework thanks to its flexibility (since it has more degrees of freedom) with respect to the Gaussian and Gamma mixtures. Moreover, thanks to the texture characteristics which are one of the most important descriptors, in particular for medical applications, the modeling of these characteristics with the proposed statistical model leads to good results of differentiation between the images. [46].

TABLE 6. Overall accuracy for CXR-Pneumonia Dataset.

Approach/Metrics	ACC(%)	DR(%)	FPR(%)
GMM	87.66	85.80	0.13
ΓMM	90.54	88.54	0.10
GMM-FD	88.25	86.90	0.12
GMM-KD	88.22	86.83	0.12
GMM-BD	88.18	86.79	0.12
ΓMM-FD	90.88	88.60	0.10
ΓMM-KD	90.85	88.53	0.10
ΓMM-BD	90. 84	88,51	0.10
gΓMM-FD	91.98	91.11	0.09
gΓMM-KD	91.77	91.05	0.09
gΓMM-BD	91.75	91.02	0.09

TABLE 7. CT-scan data statistics.

Data/Class	Train	Validation	Test	Total
Non-COVID	146	15	34	195
COVID-19	183	57	35	275
Total	329	72	69	470

TABLE 8. Overall accuracy for CT-COVID-19 Dataset.

Approach/Metrics	ACC(%)	DR(%)	FPR(%)
GMM	78.45	76.35	0.22
ΓMM	81.05	80.71	0.20
GMM-FD	78.19	77.08	0.21
GMM-KD	78.14	77.03	0.21
GMM-BD	78.09	76.99	0.21
ΓMM-FD	81.98	81.12	0.19
ΓMM-KD	81.92	81.09	0.19
ΓMM-BD	81.73	81.03	0.19
gΓMM-FD	84.23	83.16	0.17
gΓMM-KD	83.93	83.01	0.17
gΓMM-BD	83.83	82.99	0.17

D. CT-SCAN IMAGES CLASSIFICATION FOR COVID-19 DETECTION

In this work we also addressed the detection of COVID-19 patients in CT scans. It is noted that CT scans are promising in giving precise analysis and screening of COVID-19. Our experiments are based on the public dataset [47] containing in total 470 CT scans where 275 are positive cases for COVID-19 and 195 are negative cases and are labeled as Non-COVID-19. It is noted that this dataset is storing more images over time, till April 28, 2020. Moreover, it is verified by senior radiologists who have performed several diagnoses on many COVID-19 patients. Table 7 describes this dataset. In this case, the value of B for the Symmetrized Kullback-Leibler distance is equal to 2. Based on the results in Table 8, we notice that our developed model still provides better results with the CT-scan dataset, despite the complexity of these images. Indeed, for this kind of images, it should be noted that the steps of modelling and characterization are

difficult especially when these images include acute respiratory distress syndrome. This difficulty is due to the little contrast at the boundary of the lung. In addition, the number of images in this dataset is too small. For all these reasons, the obtained results are lower than the previous datasets. Our goal is to analyze these CTs images in order to detect and predict whether a patient is affected or not with COVID-19.

VII. CONCLUSION

In this paper, we have presented a three-parameter generalized gamma mixture model for modeling chest x-ray (CXR) images via texture image classification. The advantage of this mixture model over the existing Gaussian model is that it provides more flexibility to control the shape of model which is critical for classification systems. In order to make our proposed framework particularly appropriate for image classification problem, we derived some efficient measures based on Fisher, Kullback-Leibler and Bhattacharyya distances for the mixtures of generalized Gamma distributions. Thus, the proposed approach involves the extraction of robust texture descriptors, the learning of the developed mixture model, and the classification of CXR and CT scans using the derived mixtures-based distances. Through challenging applications involving pneumonia/covid-19 detection and texture classification, we have shown how our flexible model in conjunction with effective distances measures between mixtures can be used and provide excellent modeling and classification capabilities. One of the advantages of our developed framework is it can be easily extended to incorporate the selection of relevant features which can further improve the expected classification results. We hope that other real-world applications will be addressed as part of the approach we are proposing. Extensive experiments were conducted on standard medical datasets and the results demonstrate that the proposed approach outperforms many other generative approaches. In our future work we will study if the proposed framework can be extended via nonparametric Bayesian principle in order to increase the classification accuracy. Future works could be also devoted to avoiding the limitations of Monte-Carlo approximations by considering for example the Cauchy-Schwarz divergence as previously done successfully for different mixture models [23], [29], [30]. We plan to investigate this work for other tasks such as image segmentation by classifying the smaller regions and evaluate it for other related image classification tasks such as object recognition.

REFERENCES

[1] A. Jacobi, M. Chung, A. Bernheim, and C. Eber, "Portable chest X-ray in coronavirus disease-19 (COVID-19): A pictorial review," *Clin. Imag.*, vol. 64, pp. 35–42, Aug. 2020.

[2] B. Vellingiri, K. Jayaramayya, M. Iyer, A. Narayanasamy, V. Govindasamy, B. Giridharan, S. Ganesan, A. Venugopal, D. Venkatesan, H. Ganesan, K. Rajagopalan, P. K. S. M. Rahman, S.-G. Cho, N. S. Kumar, and M. D. Subramaniam, "COVID-19: A promising cure for the global panic," *Sci. Total Environ.*, vol. 725, Jul. 2020, Art. no. 138277.

[3] M. F. Hashmi, S. Katiyar, A. G. Keskar, N. D. Bokde, and Z. W. Geem, "Efficient pneumonia detection in chest xray images using deep transfer learning," *Diagnostics*, vol. 10, no. 6, p. 417, Jun. 2020.

[4] Y. Lai, Y. Ping, W. He, B. Wang, J. Wang, and X. Zhang, "Variational Bayesian inference for finite inverted Dirichlet mixture model and its application to object detection," *Chin. J. Electron.*, vol. 27, no. 3, pp. 603–610, May 2018.

[5] G. J. McLachlan and D. Peel, *Finite Mixture Models*. Hoboken, NJ, USA: Wiley, 2004.

[6] F. Najar, S. Bourouis, A. Zaguia, N. Bouguila, and S. Belghith, "Unsupervised human action categorization using a Riemannian averaged fixed-point learning of multivariate GGMM," in *Proc. 15th Int. Conf. Image Anal. Recognit. (ICIAR)*, Povoa de Varzim, Portugal, Jun. 2018, pp. 408–415.

[7] A. M. Khan, H. El-Daly, and N. M. Rajpoot, "A gamma-Gaussian mixture model for detection of mitotic cells in breast cancer histopathology images," in *Proc. 21st Int. Conf. Pattern Recognit. (ICPR)*, Tsukuba, Japan, Nov. 2012, pp. 149–152.

[8] F. Najar, S. Bourouis, N. Bouguila, and S. Belghith, "A new hybrid discriminative/generative model using the full-covariance multivariate generalized Gaussian mixture models," *Soft Comput.*, vol. 24, no. 14, pp. 10611–10628, Jul. 2020.

[9] S. Bourouis, F. R. Al-Osaimi, N. Bouguila, H. Sallay, F. Aldosari, and M. Al Mashrgy, "Bayesian inference by reversible jump MCMC for clustering based on finite generalized inverted Dirichlet mixtures," *Soft Comput.*, vol. 23, pp. 5799–5813, May 2018.

[10] M. A. T. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3, pp. 381–396, Mar. 2002.

[11] N. Bouguila, D. Ziou, and J. Vaillancourt, "Unsupervised learning of a finite mixture model based on the Dirichlet distribution and its application," *IEEE Trans. Image Process.*, vol. 13, no. 11, pp. 1533–1543, Nov. 2004.

[12] T. Elguebaly and N. Bouguila, "A hierarchical nonparametric Bayesian approach for medical images and gene expressions classification," *Soft Comput.*, vol. 19, no. 1, pp. 189–204, Jan. 2015.

[13] R. Alroobaea, S. Rubaiee, S. Bourouis, N. Bouguila, and A. Alsufyani, "Bayesian inference framework for bounded generalized Gaussian-based mixture model and its application to biomedical images classification," *Int. J. Imag. Syst. Technol.*, vol. 30, no. 1, pp. 18–30, 2020.

[14] F. Najar, S. Bourouis, N. Bouguila, and S. Belghith, "Unsupervised learning of finite full covariance multivariate generalized Gaussian mixture models for human activity recognition," *Multimedia Tools Appl.*, vol. 78, no. 13, pp. 18669–18691, Jul. 2019.

[15] Y. Li, G. Wang, M. Li, J. Li, L. Shi, and J. Li, "Application of CT images in the diagnosis of lung cancer based on finite mixed model," *Saudi J. Biol. Sci.*, vol. 27, no. 4, pp. 1073–1079, Apr. 2020.

[16] S. Bourouis, A. Zaguia, N. Bouguila, and R. Alroobaea, "Deriving probabilistic SVM kernels from flexible statistical mixture models and its application to retinal images classification," *IEEE Access*, vol. 7, pp. 1107–1117, 2019.

[17] I. Channoufi, S. Bourouis, N. Bouguila, and K. Hamrouni, "Image and video denoising by combining unsupervised bounded generalized Gaussian mixture modeling and spatial information," *Multimedia Tools Appl.*, vol. 77, no. 19, pp. 25591–25606, Oct. 2018.

[18] M. Azam and N. Bouguila, "Multivariate bounded support laplace mixture model," *Soft Comput.*, vol. 24, no. 17, pp. 13239–13268, Sep. 2020.

[19] S. Jorjandi, H. Rabbani, R. Kafieh, and Z. Amini, "Statistical modeling of optical coherence tomography images by asymmetric normal laplace mixture model," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 4399–4402.

[20] X. Cong-Hua, X. Wen-Bin, and C. Jin-Yi, "Medical image denoising by generalised Gaussian mixture modelling with edge information," *IET Image Process.*, vol. 8, no. 8, pp. 464–476, Aug. 2014.

[21] H. Rabbani, R. Nezafat, and S. Gazor, "Wavelet-domain medical image denoising using bivariate Laplacian mixture model," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 12, pp. 2826–2837, Dec. 2009.

[22] I. Channoufi, S. Bourouis, N. Bouguila, and K. Hamrouni, "Spatially constrained mixture model with feature selection for image and video segmentation," in *Proc. 8th Int. Conf. Image Signal Process.*, Cherbourg, France, Jul. 2018, pp. 36–44.

[23] K. Kampa, E. Hasanbelliu, and J. C. Principe, "Closed-form cauchy-schwarz PDF divergence for mixture of gaussians," in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2011, pp. 2578–2585.

[24] C. Beckmann, M. Woolrich, and S. Smith, "Gaussian/gamma mixture modelling of ICA/GLM spatial maps," in *Proc. 9th Int. Conf. Funct. Mapping Human Brain*, New York, NY, USA, vol. 19, Jun. 2003, p. S985.

[25] N. Bouguila, K. Almakadmeh, and S. Boutemedjet, "A finite mixture model for simultaneous high-dimensional clustering, localized feature selection and outlier rejection," *Expert Syst. Appl.*, vol. 39, no. 7, pp. 6641–6656, Jun. 2012.
[26] D. Ziou, N. Bouguila, M. S. Allili, and A. El-Zaart, "Finite gamma mixture modelling using minimum message length inference: Application to SAR image analysis," *Int. J. Remote Sens.*, vol. 30, no. 3, pp. 771–792, Feb. 2009.
[27] H.-C. Li, V. A. Krylov, P.-Z. Fan, J. Zerubia, and W. J. Emery, "Unsupervised learning of generalized gamma mixture model with application in statistical modeling of high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2153–2170, Apr. 2016.
[28] G. Vegas-Sanchez-Ferrero, S. Aja-Fernandez, C. Palencia, and M. Martin-Fernandez, "A generalized gamma mixture model for ultrasonic tissue characterization," *Comput. Math. Methods Med.*, vol. 2012, Dec. 2012, Art. no. 481923.
[29] H. Rami, A. Drissi El Maliani, and M. El Hassouni, "A finite mixture of weibull-based statistical model for texture retrieval in the complex wavelet domain," *IEEE Access*, vol. 7, pp. 130144–130155, 2019.
[30] H. Rami, L. Belmerhnia, A. D. E. Maliani, and M. E. Hassouni, "Texture retrieval using mixtures of generalized Gaussian distribution and Cauchy–Schwarz divergence in wavelet domain," *Signal Process.: Image Commun.*, vol. 42, pp. 45–58, Mar. 2016.
[31] J. Xie, Y. Jiang, and H.-t. Tsui, "Segmentation of kidney from ultrasound images based on texture and shape priors," *IEEE Trans. Med. Imag.*, vol. 24, no. 1, pp. 45–57, Jan. 2005.
[32] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1657–1663, Jun. 2010.
[33] L. Liu, P. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Local binary features for texture classification: Taxonomy and experimental study," *Pattern Recognit.*, vol. 62, pp. 135–160, Feb. 2017.
[34] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the *EM* algorithm," *J. Roy. Stat. Soc., B (Methodol.)*, vol. 39, no. 1, pp. 1–22, 1977.
[35] N. Bouguila and D. Ziou, "High-dimensional unsupervised selection and estimation of a finite generalized Dirichlet mixture model based on minimum message length," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 10, pp. 1716–1731, Oct. 2007.
[36] R. A. Baxter and J. J. Oliver, "Finding overlapping components with MML," *Statist. Comput.*, vol. 10, no. 1, pp. 5–16, Jan. 2000.
[37] B. Alghabashi and N. Bouguila, "Finite multi-dimensional generalized gamma mixture model learning based on MML," in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, M. A. Wani, M. M. Kantardzic, M. S. Mouchaweh, J. Gama, and E. Lughofer, Eds. Orlando, FL, USA: IEEE Press, Dec. 2018, pp. 1131–1138.
[38] T. S. Jaakkola and D. Haussler, "Exploiting generative models in discriminative classifiers," in *Proc. Adv. Neural Inf. Process. Syst.*, Denver, CO, USA, Nov./Dec. 1998, pp. 487–493.
[39] P. J. Moreno, P. Ho, and N. Vasconcelos, "A Kullback-Leibler divergence based kernel for SVM classification in multimedia applications," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, S. Thrun, L. K. Saul, and B. Schölkopf, Eds. Vancouver, BC, Canada: MIT Press, Dec. 2003, pp. 1385–1392.
[40] N. Bouguila, "Deriving kernels from generalized Dirichlet mixture models and applications," *Inf. Process. Manage.*, vol. 49, no. 1, pp. 123–137, Jan. 2013.
[41] A. B. Chan, N. Vasconcelos, and P. J. Moreno, "A family of probabilistic kernels based on information divergence," Univ. California, San Diego, CA, USA, Tech. Rep. SVCL-TR-2004-1, 2004.
[42] T. Jebara and R. Kondor, "Bhattacharyya expected likelihood kernels," in *Proc. 16th Annu. Conf. Comput. Learn. Theory, 7th Kernel Workshop, COLT/Kernel Comput. Learn. Theory Kernel Mach.*, in Lecture Notes in Computer Science, vol. 2777, B. Schölkopf and M. K. Warmuth, Eds. Washington, DC, USA: Springer, Aug. 2003, pp. 57–71.
[43] A. Boulmerka, M. Saïd Allili, and S. Ait-Aoudia, "A generalized multiclass histogram thresholding approach based on mixture modelling," *Pattern Recognit.*, vol. 47, no. 3, pp. 1330–1348, Mar. 2014.
[44] J. Paul Cohen, P. Morrison, and L. Dao, "COVID-19 image data collection," 2020, *arXiv:2003.11597*. [Online]. Available: http://arxiv.org/abs/2003.11597
[45] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "COVID-19 image data collection: Prospective predictions are the future," 2020, *arXiv:2006.11988*. [Online]. Available: https://arxiv.org/abs/2006.11988
[46] J. Melendez, B. van Ginneken, P. Maduskar, R. H. H. M. Philipsen, K. Reither, M. Breuninger, I. M. O. Adetifa, R. Maane, H. Ayles, and C. I. Sanchez, "A novel multiple-instance learning-based approach to computer-aided detection of tuberculosis on chest X-Rays," *IEEE Trans. Med. Imag.*, vol. 34, no. 1, pp. 179–192, Jan. 2015.
[47] X. Yang, X. He, J. Zhao, Y. Zhang, S. Zhang, and P. Xie, "COVID-CT-dataset: A CT scan dataset about COVID-19," 2020, *arXiv:2003.13865*. [Online]. Available: http://arxiv.org/abs/2003.13865