

Article

Bayesian Learning of Shifted-Scaled Dirichlet Mixture Models and Its Application to Early COVID-19 Detection in Chest X-ray Images

Sami Bourouis ^{1,*} , Abdullah Alharbi ¹ and Nizar Bouguila ² 

¹ Department of Information Technology, College of Computers and Information Technology, Taif University, Taif, P.O. Box 11099, Taif 21944, Saudi Arabia; amharbi@tu.edu.sa

² The Concordia Institute for Information Systems Engineering (CIISE), Concordia University, Montreal, QC H3G 1T7, Canada; nizar.bouguila@concordia.ca

* Correspondence: s.bourouis@tu.edu.sa

Abstract: Early diagnosis and assessment of fatal diseases and acute infections on chest X-ray (CXR) imaging may have important therapeutic implications and reduce mortality. In fact, many respiratory diseases have a serious impact on the health and lives of people. However, certain types of infections may include high variations in terms of contrast, size and shape which impose a real challenge on classification process. This paper introduces a new statistical framework to discriminate patients who are either negative or positive for certain kinds of virus and pneumonia. We tackle the current problem via a fully Bayesian approach based on a flexible statistical model named shifted-scaled Dirichlet mixture models (SSDMM). This mixture model is encouraged by its effectiveness and robustness recently obtained in various image processing applications. Unlike frequentist learning methods, our developed Bayesian framework has the advantage of taking into account the uncertainty to accurately estimate the model parameters as well as the ability to solve the problem of overfitting. We investigate here a Markov Chain Monte Carlo (MCMC) estimator, which is a computer-driven sampling method, for learning the developed model. The current work shows excellent results when dealing with the challenging problem of biomedical image classification. Indeed, extensive experiments have been carried out on real datasets and the results prove the merits of our Bayesian framework.

Keywords: infection detection; COVID-19; X-ray images; image classification; bayesian inference; shifted-scaled dirichlet distribution; MCMC; gibbs sampling



Citation: Bourouis, S.; Alharbi, A.; Bouguila, N. Bayesian Learning of Shifted-Scaled Dirichlet Mixture Models and Its Application to Early COVID-19 Detection in Chest X-ray Images. *J. Imaging* **2021**, *7*, 7. <https://doi.org/10.3390/jimaging7010007>

Received: 11 November 2020

Accepted: 7 January 2021

Published: 10 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction and Related Works

Pneumonia is a severe disease issue resulting in inflammation of the lungs where a large number of people lose their lives every day. The causes of this infectious disease could be attributed to viruses or bacteria. Today, the SARS-CoV-2 virus named COVID-19 pneumonia is causing a significant outbreak around the world, having a serious impact on the health and life of several people. In particular, it causes pneumonia in humans and carries severe infections between people. Patients with COVID-19 can have acute symptoms and some may die of major organ failure. One of the critical steps in the fight against this disease is the possibility to quickly detect and track contaminated persons and place them under particular care. Early inspection of confirmed cases is of great urgency because of its infectious nature. One of the many ways of detecting the disease is by a chest radiographs of the patient. Recently, some studies have shown that studying COVID-19 from Chest X-ray images may be considered as the quickest solution to diagnose patients [1]. It is noteworthy that chest X-ray radiography is one of the interesting imaging to diagnose several related chest diseases such as pneumonia, lung cancer, emphysema and pulmonary edema [2,3]. However, sometimes this medical imaging can be subject

to error for inexperienced radiologists, while being tedious for experienced ones. Visual examination of these radiographs is generally restricted due to low infectious disease specificity. In addition, the presence of noise, the contrast which is often insufficient between the soft tissues and the overlap in appearance properties are often sources of error for an accurate diagnosis [1,4]. These inconsistencies can result in important biased decisions for clinicians.

To deal with these drawbacks and to detect infected patients, it is necessary to develop effective and automated computerized support tools able to offer radiologists desirable measures about the disease severity. These tools should also allow rapid detection and prediction of any possible infection, in particular COVID-19. Nevertheless, performing a precise analysis of big biomedical data is too difficult and time consuming because these images contain various patterns and symptoms at different stages (early, middle, advanced) [4,5]. For instance at the early stage, it is not easy at all to discover COVID-19 symptoms having acute respiratory distress syndrome in chest X-ray (CXR) scans because these symptoms can look similar to other viral infections like RSV pneumonia. Consequently, it is important to consider such assumption and to take into account robust features extraction techniques when implementing new systems.

Several promising algorithms have been implemented in the past decades to deal especially with infection detection. Some traditional machine learning-based methods are applied to support pneumonia diagnosis in children by classifying chest radiographs into normal or pneumonia cases [6]. Haar wavelet transform is also investigated as an effective feature extraction technique. Some classifiers such as FCM, DWT and WFT [2] and K-nearest neighbor (k-NN) [3] were exploited in this context to detect pneumonia infection. Nevertheless, these conventional methods fail to identify properly lung with lesions. It is true that traditional methods helped the specialists in their diagnosis, but the resulting accuracy was poor. Thus, other image processing-based systems have been proposed to address the problems of infection localization and detecting malicious lesions using, for example, SVM, Neural Networks (NN) and Deep NN (DNN) [5,7–9].

The Fully CN (FCN) method is also applied for segmenting lung in CXR [10]. Another work which is conducted using deep learning method is proposed in [11] to classify CT scan and chest X-ray into three classes: influenza-A viral pneumonia, COVID-19, and normal. The obtained accuracy is 89.3%. As a result, the accuracy is 89.3% and the training process takes a long time. After studying the related work, it is obvious that the success of supervised CNN and deep learning methods to classify CXR images and detect COVID-19 relies mainly on the size of training data. For smaller data set, these techniques are not suitable since this size is responsible for poor performances and in many cases, it becomes too difficult to generate more training data. Thus, it is important to look for other alternatives. Features extraction methods are also exploited in conjunction with some classifiers in order to extract and select relevant visual features. For instance, the ResNet50 feature extractor is used with SVM and CNN for detecting and classifying lung nodule disease in chest CT- images [12]. Other approaches such as registration and active shape models [13,14] are exploited with pixel-based statistical classification methods in order to find the boundary/region targets. For example, the lung region is determined through a non-rigid registration step between the chest radiograph of the image patient and a reference model [13].

The good results obtained from applying artificial intelligence and machine learning models to some previous epidemics are motivating researchers to provide new perspective for addressing this novel coronavirus outbreak. In particular, classifying non-Gaussian data in an unsupervised way can be of great interest for automated medical applications. Among the main existing methods to tackle this problem, statistical mixture models have recently gained considerable interest from both the theoretical and practical points of view [15–20]. This approach has led to the design of new more efficient tools. Our work is mainly based on recent research findings that have shown modeling visual data (such as images) effectively is very important for further applications such as image classification. In particular, the taking into account of the distribution of Dirichlet is very interesting to

deal with non-Gaussian data modelling [21]. Other derived models such as the scaled Dirichlet mixture (a generalization of the Dirichlet) [16] have also been shown to be effective for data grouping and classification. Further works have show that it is possible to improve these last two models by introducing an additional parameter which leads to a more flexible model. The resulting statistical mixture is called shifted-scaled Dirichlet mixture (SSDMM) and is assumed to be a generalization of the scaled model (here the Shifted term mean a perturbation in the simplex). This new model has been applied successfully for a variety of applications [22].

2. Motivations

The work developed in [22] is based on a shifted-scaled Dirichlet mixture model (SSDMM) and evaluated for data clustering and writer identification. Two important issues arise when deploying mixture models which are calculating the parameters of the mixture and determining the exact number of components that best describes the data set. These issues have been tackled recently by learning the SSDMM via deterministic Maximum Likelihood Estimator (MLE) [22]. Nevertheless, it is known that MLE has major shortcomings linked to its sensitivity at the initialization step. Therefore, a better solution especially for our case (i.e., when dealing with complex medical noisy data including COVID-19 infection) is to develop a more robust alternative based on fully Bayesian inference approach. We recall that Bayesian estimation has attracted a lot of attention for many applications [23–33]. It is also known that the Bayesian approach may be more practical due to the existance of powerful simulation techniques like MCMC [29]. Moreover, the model complexity can be easily solved using for example the marginal likelihood-based technique. Thus, our focus in this paper is to implement an effective Bayesian learning method for SSDMM in order to take into account the complexity of medical data and to overcome the drawbacks of frequentist (deterministic) approaches [34,35]. To the best of our knowledge, such an approach has never been tackled before, especially for the problem of chest x-ray images classification.

The rest of this paper is organized as follows. In next section, the finite shifted-scaled Dirichlet mixture model and the Bayesian approach are exposed. Experimental results and the merits of our approach are introduced in Section 4. Finally, we end this work and provide some possible extensions to be treated in the future.

3. Bayesian Framework for the Shifted-Scaled Dirichlet Mixture Model

We start this section by revising both the Dirichlet and scaled Dirichlet distributions, and then introduce a new generalization of these distributions named shifted-scaled Dirichlet distribution (SSDD). The finite shifted-scaled Dirichlet mixture model is also presented. Then, we develop a fully Bayesian framework for learning the parameters of this finite mixture model.

3.1. Dirichlet and Scaled-Dirichlet Distributions

Definition 1 (Dirichlet distribution). *Let us consider a random vector $Y = (y_1, \dots, y_D) \in S^D$ (sample space), where $\sum_{d=1}^D y_d = 1$. We say that Y has a D -variate Dirichlet distribution with parameter $\vec{\alpha} = (\alpha_1, \dots, \alpha_D) \in \mathbb{R}_+^D$ if its density function is:*

$$Y \sim Dir^D(\vec{\alpha})$$

$$f(\vec{Y}) = p(\vec{Y}|\theta) = \frac{\Gamma(\alpha_+)}{\prod_{i=1}^D \Gamma(\alpha_i)} \prod_{i=1}^D y_i^{\alpha_i-1} \tag{1}$$

where $\vec{\alpha}$ denotes a shape parameter, $\alpha_+ = \sum_{i=1}^D \alpha_i$ and Γ indicates the Euler gamma function.

It is noted that the Dirichlet distribution with D parameters ($Y \sim Dir^D(\vec{\alpha})$) is still popular, especially when it comes to analyzing composition data, and this popularity is due to its its conjugate property with the multinomial likelihood.

Definition 2 (Scaled Dirichlet distribution). *If Y follows a scaled Dirichlet distribution, then its density function is given as:*

$$Y \sim SDir^D(\vec{\alpha}, \vec{\beta})$$

$$f(\vec{Y}) = p(\vec{Y}|\theta) = \frac{\Gamma(\alpha_+)}{\prod_{i=1}^D \Gamma(\alpha_i)} \frac{\prod_{i=1}^D \beta_i^{\alpha_i} y_i^{\alpha_i-1}}{(\sum_{i=1}^D \beta_i y_i)^{\alpha_+}} \tag{2}$$

$\vec{\alpha} = (\alpha_1, \dots, \alpha_D)$ and $\vec{\beta} = (\beta_1, \dots, \beta_D) \in \mathbb{R}_+^D$ are the parameters of this distribution. β is a scale parameter.

The scaled Dirichlet distribution has $2D$ parameters and in this case we have $Y \sim SDir^D(\vec{\alpha}, \vec{\beta})$. If the parameter β is fixed, then we obtain a Dirichlet model.

3.2. Finite Shifted-Scaled Dirichlet Mixture Model

Definition 3 (Shifted-Scaled Dirichlet distribution). *Suppose that Y follows a shifted scaled Dirichlet distribution with parameters $\vec{\alpha} = (\alpha_1, \dots, \alpha_D) \in \mathbb{R}_+^D$, $\vec{\lambda} = (\lambda_1, \dots, \lambda_D) \in \mathcal{S}^D$ and $a \in \mathbb{R}_+$. Then, the density probability of this distribution is given as:*

$$Y \sim pSDir^D(\vec{\alpha}, \vec{\lambda}, a)$$

$$f(\vec{Y}) = p(\vec{Y}|\theta) = \frac{\Gamma(\alpha_+)}{\prod_{i=1}^D \Gamma(\alpha_i)} \frac{1}{a^{D-1}} \frac{\prod_{i=1}^D \lambda_i^{-(\alpha_i/a)} y_i^{(\alpha_i/a)-1}}{(\sum_{i=1}^D (y_i/\lambda_i)^{(1/a)})^{\alpha_+}} \tag{3}$$

where $\vec{\lambda}$ denotes a location parameter.

The shifted-scaled Dirichlet distribution has $2D$ parameters and in this case we have $Y \sim pSDir^D(\vec{\alpha}, \vec{\lambda}, a)$. If the parameter $a = 1$, then we obtain a scaled Dirichlet model.

Now, suppose that we have a set of vectors $\mathcal{Y} = \{\vec{Y}_1, \vec{Y}_2, \dots, \vec{Y}_N\}$, where each vector $\vec{Y}_n = (y_{n1}, \dots, y_{nD})$ follows a mixture of SSD, then the corresponding likelihood is defined as:

$$p(\mathcal{Y}|\Theta) = \prod_{n=1}^N \sum_{k=1}^K \pi_k p(\vec{Y}_n|\theta_k) \tag{4}$$

where the model's parameters are defined by $\Theta = (\vec{\pi}, \theta)$ and $\{\pi_k\}$ are positive mixing parameters ($\sum_k \pi_k = 1$). Each vector is supposed coming from one component as $\vec{Y}_n \sim pSDir^D(\vec{\alpha}, \vec{\lambda}, a)$. The shape parameter has the role to describe the form of the shifted SDMM. The scale (a) checks how the plotting of the density is distributed and $\vec{\lambda}$ follows the location of the data densities. In the next section, we will develop our Bayesian approach based on the presented mixture of SSDD.

3.3. Fully Bayesian Learning Algorithm

In many cases, the deterministic approach (named also maximum likelihood-based technique) via the well known EM algorithm [36] is used to estimate the parameters of finite mixture models due to its simplicity. Deterministic approach assumes that $\mathcal{Z} = (\vec{Z}_1, \dots, \vec{Z}_N)$, is a missing data. Thus, if $\vec{Y}_n \in j$ then $Z_{ij} = 1$, else $Z_{nj} = 0$. Because the likelihood-technique depends on initial values and is sensitive to local minima, we propose here to overcome these limitations by developing an efficient way based on Bayesian inference to better learn the Shifted-Scaled Dirichlet mixture model. More precisely, we propose to investigate one of the effective simulation techniques called Markov Chain Monte Carlo (MCMC) via Gibbs sampler [37,38]. Thus, the complete likelihood is defined as:

$$p(\mathcal{Y}, \mathcal{Z}|\Theta) = \prod_{n=1}^N \prod_{k=1}^K (\pi_k p(\vec{Y}_n|\theta_k))^{Z_{nk}} \tag{5}$$

Using Bayes formula, the likelihood and the priors will be expressed together to define the posterior distribution like this:

$$p(\Theta|\mathcal{Y}, \mathcal{Z}) \propto p(\mathcal{Y}, \mathcal{Z}|\Theta)p(\Theta) \tag{6}$$

The proposed Bayesian algorithm for SSDMM parameters' learning is based on the following steps :

1. Initialization
2. Step t: For t = 1, ...
 - (a) Generate $\vec{Z}_i^{(t)} \sim \mathcal{M}(1; \hat{Z}_{i1}^{(t-1)}, \dots, \hat{Z}_{iM}^{(t-1)})$
 - (b) Generate $\vec{\pi}^{(t)}$ from $p(\pi|\mathcal{Z}^{(t)})$
 - (c) Generate $(\theta)^{(t)}$ from $p(\theta|\mathcal{Z}^{(t)}, \mathcal{Y})$

where $\mathcal{M}(1; \hat{Z}_{i1}^{(t-1)}, \dots, \hat{Z}_{iM}^{(t-1)})$ is a multinomial distribution of order one with parameters $(p(1|\vec{Y}_i)^{(t-1)}, \dots, p(M|\vec{Y}_i)^{(t-1)})$. Based on this algorithm, we have to evaluate $p(\pi|\mathcal{Z})$ and $p(\theta|\mathcal{Z}, \mathcal{Y})$.

3.3.1. Priors and Posteriors

The choice of priors is one of the most crucial steps in Bayesian modeling. These priors reflect our belief about the the model's parameters and are updated and enhanced according to the observed data (see for example details in [39]). In the following, the choice of the priors is addressed as well as the determining of the resulting posteriors for our fully Bayesian approach.

Estimating the posterior will lead to have our parameters $\Theta \sim p(\Theta|\mathcal{Y}, \mathcal{Z})$. In order to perform this step, we proceed with an elegant sampling technique called Gibbs sampler. This method allows the use of conditional posterior distribution in order to update each parameter.

Since no convenient conjugate prior exist for $\vec{\alpha}_k$ and a_k , we adopt a common choice for them which is the Gamma distribution $\mathcal{G}(\cdot)$:

$$p(\alpha_{kd}) = \mathcal{G}(\alpha_{kd}|u_{kd}, v_{kd}) \tag{7}$$

$$p(a_k) = \mathcal{G}(a_k|g_k, h_k) \tag{8}$$

Then, we determine the posterior distributions according to these priors and by considering the following:

$$p(\vec{\alpha}_k|\mathcal{Z}, \mathcal{Y}) \propto p(\vec{\alpha}_k) \prod_{Z_{ik}=1} p(\vec{Y}_i|\theta_k) \propto \prod_{d=1}^D p(\alpha_{kd}) \prod_{Z_{ik}=1} p(\vec{Y}_i|\theta_k) \tag{9}$$

$$p(a_k|\mathcal{Z}, \mathcal{Y}) \propto p(a_k) \prod_{Z_{ik}=1} p(\vec{Y}_i|\theta_k) \tag{10}$$

Regarding the parameter $\vec{\lambda}_k$, since it is defined in a simplex, therefore, it is a common and classic choice in Bayesian inference to choose the Dirichlet distribution as prior with parameters $\eta_k = (\eta_{k1}, \dots, \eta_{kD})$. So, it is expressed as:

$$p(\vec{\lambda}_k|\eta_k) = \frac{\Gamma(\sum_{j=1}^D \eta_{kj})}{\prod_{j=1}^D \Gamma(\eta_{kj})} \prod_{j=1}^D \lambda_{kj}^{\eta_{kj}-1} \tag{11}$$

Knowing this prior, we can estimate the posterior distribution using the following equation:

$$p(\vec{\lambda}_k|\mathcal{Z}, \mathcal{Y}) \propto p(\vec{\lambda}_k|\eta_k) \prod_{Z_{ik}=1} p(\vec{Y}_i|\theta_k) \tag{12}$$

For the prior of mixing weight $\vec{\pi}$, the common choice is the Dirichlet distribution since $\sum_{j=1}^K \pi_j = 1$. So, the mixing weight prior is expressed as:

$$p(\vec{\pi}|K, \delta) = \frac{\Gamma(\sum_{j=1}^K \delta_j)}{\prod_{j=1}^K \Gamma(\delta_j)} \prod_{j=1}^K \pi_j^{\delta_j-1} \tag{13}$$

The selected prior of \mathcal{Z} (membership variable) is defined as :

$$p(\mathcal{Z}|\vec{\pi}, K) = \prod_{j=1}^K \pi_j^{n_j} \tag{14}$$

where n_j is the total vectors in cluster j . Given the former equations Equations (13) and (14) we have

$$\begin{aligned} p(\vec{\pi}|\dots) &\propto p(\mathcal{Z}|\vec{\pi}, K)p(\vec{\pi}|K, \delta) \\ &\propto \prod_{j=1}^K \pi_j^{n_j} \frac{\Gamma(\sum_{j=1}^K \delta_j)}{\prod_{j=1}^K \Gamma(\delta_j)} \prod_{j=1}^K \pi_j^{\delta_j-1} \propto \frac{\Gamma(\sum_{j=1}^K \delta_j)}{\prod_{j=1}^K \Gamma(\delta_j)} \prod_{j=1}^K \pi_j^{n_j+\delta_j-1} \end{aligned} \tag{15}$$

This posterior is proportional to the Dirichlet distribution $(\delta_1 + n_1, \dots, \delta_K + n_K)$. In addition, the posterior of the membership \mathcal{Z} may be deduced as:

$$p(Z_i = j|\dots) \propto \pi_j p(\vec{Y}_n|\theta_j) \tag{16}$$

Finally, we choose the uniform distribution as an appropriate prior for K . This value can vary between 1 and K_{max} (K_{max} is a predefined value). We summarize the proposed model in the following graphical representation Figure 1.

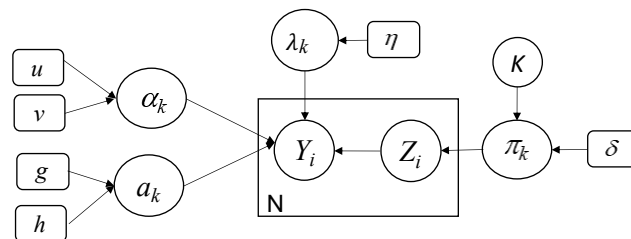


Figure 1. Graphical representation of our developed Bayesian finite shifted-scaled Dirichlet mixture model. Fixed hyperparameters are indicated by rounded boxes and random variables by circles. Y is the observed variable, Z represents the latent variable, the large box indicates repeated process, and the arcs show the dependencies between variables.

3.3.2. Complete Bayesian Estimation-Algorithm

The Gibbs sampling technique is mainly based on alternating conditional distributions for several steps. Indeed, for each iteration t , the resulted estimate Θ^t is sampled from its previous approximate Θ^{t-1} . Having all these posterior probabilities in hand, the complete MCMC-based Bayesian algorithm to learn the parameters of our finite mixture model and especially the steps of our Gibbs sampler are as follows:

1. Initialization
2. Step t: For $t = 1, \dots$
 - (a) Generate $Z_i^{(t)} \sim \mathcal{M}(1; \hat{Z}_{i1}^{(t-1)}, \dots, \hat{Z}_{iK}^{(t-1)})$
 - (b) Compute $n_k^{(t)} = \sum_{i=1}^N \mathbb{I}_{Z_{ik}^{(t)}=j}$
 - (c) Generate $\vec{\pi}^{(t)}$ from Equation (15)
 - (d) Generate $\vec{\alpha}_k^{(t)}, a_k^{(t)}$, and $\vec{\lambda}_k^{(t)}$ ($k = 1, \dots, K$) from Equations (9), (10) and (12), respectively, using random-walk Metropolis-Hastings (M-H) algorithm [40,41].

where $\mathcal{M}(1; \hat{Z}_{i1}^{(t-1)}, \dots, \hat{Z}_{iM}^{(t-1)})$ is a multinomial distribution of order one with parameters $(p(1|\vec{Y}_i)^{(t-1)}, \dots, p(M|\vec{Y}_i)^{(t-1)})$.

4. Experimental Results

The goal of this section is to evaluate and validate the developed statistical model with the different inference techniques. We have considered several real data sets of images including COVID-19 and different pneumonia types.

4.1. Data Sets

The first main COVID-19 dataset (<https://github.com/ieee8023/covid-chestxray-dataset>) for our experiments is the one developed by Cohen et al. [42]. It contains 542 Chest X-ray (CXR) images. A subset of 434 CXR images represent patients positive to COVID-19 and the rest are COVID-19 negative. The image dimension is 4248×3480 pixels. Main statistics of this dataset are given in Table 1. An illustrative sample of confirmed Coronavirus Disease 2019 (COVID-19) is given in Figure 2. This image is from a 53-year-old female who had a fever and cough for 5 days. Indeed, Multifocal patchy opacities can be seen in both lungs (arrows) [43].

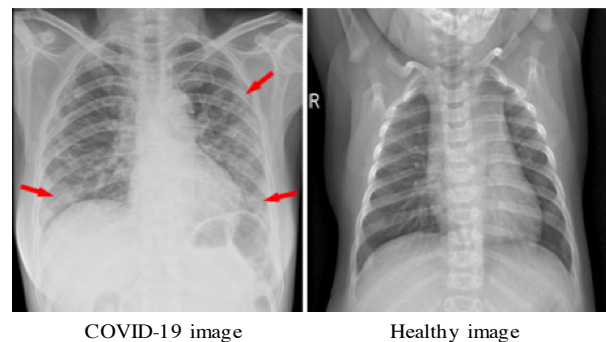


Figure 2. Illustrative sample of Chest X-Rays image with COVID-19 [43].

Table 1. Data description.

Dataset	Class	Train	Validation	Test	Total
CXR-COVID	Non-COVID-19	70	20	18	108
	COVID-19	328	80	26	434
CXR-Augmented-COVID	Non-COVID-19	512	100	300	912
	COVID-19	512	100	300	912
CXR-Pneumonia	Normal	1341	8	234	1583
	Pneumonia	3875	8	390	4273

We run also our implemented framework on another available dataset named Augmented COVID-19 Dataset (<https://data.mendeley.com/datasets/2fxz4px6d8/4>). It is collected from the previous dataset and the Kaggle one (kaggle.com/paultimothymooney/chest-xray-pneumonia). It is made up of augmented radiographics with and without COVID-19. Here, the number of images is larger than the previous dataset. Our aim is to study the performance of our model when the size of the data increases. This dataset contains 912 COVID-19 images and 912 non COVID-19 images. The augmentation process takes into account some geometric transformations and other ones such as translation, rotation, scaling, flipping, noising, blurring, etc. Some illustrative augmented images are given in Figure 3.

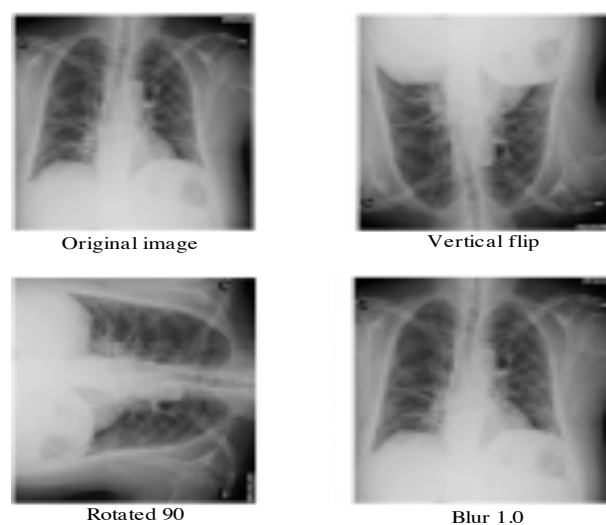


Figure 3. Illustrative examples of augmented Chest X-Rays with COVID-19 from the dataset [44].

Finally, we use the chest-xray-pneumonia to evaluate the performance. Thus, we run our algorithm on big dataset (viral, bacterial infection, and normal) Kaggle (<https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>). It contains 5856 CXR images where 1583 are normal and 4273 are infected with pneumonia. The image dimension is 1024×1024 pixels. This dataset is structured into three folders: train, test and val. Some samples are given in Figure 4. Statistics about this dataset are shown in Table 1.

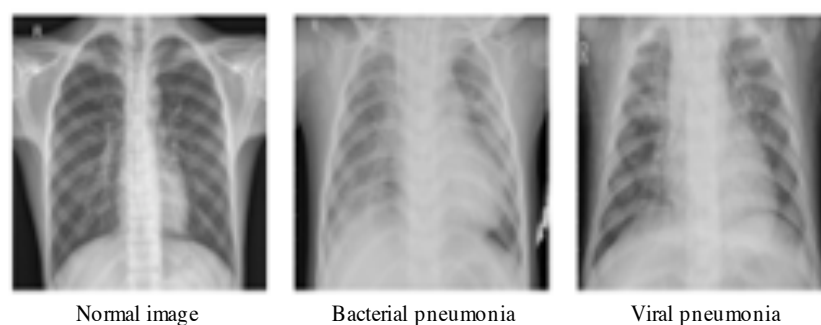


Figure 4. Illustrative samples of chest-xray-pneumonia from the dataset in [45].

4.2. Methodology

The developed model is applied to classify several images from different datasets as normal or COVID-19 affected patients using CXR images. To deal with this objective, we proceed with some preprocessing steps. After a pre-segmentation step of the lung region, we extracted some relevant features based on texture analysis. Indeed, several recently published works have shown that the lung is the basic organ which is affected by the corona COVID-19 virus. The classification is performed into two classes: normal and abnormal. Each image is modelled with a mixture of SSDDMM, then we apply the MCMC algorithm to estimate the parameters of each component. Here, the classification problem is presented in terms of assigning each image to the appropriate class using the Bayes rules. In other word, each image is affected to the class that has the greatest posterior probability. The pipeline of the proposed method is given in Figure 5.

It is noted that, in many cases, medical images such as chest x-rays are not easy to interpret; thus, it is mandatory to identify important patterns to interpret better and improve the decision. Feature extraction problem is the process of acquiring relevant information such as texture. The step of feature extraction has the role to improve the performance and accelerate the processing time. In particular, texture's structures (e.g., fine, smooth,

coarse or grained) characterize effectively visual patterns in the image. In the state of the art, many texture extraction methods have been proposed such as statistical ones which are based on different statistics order of the gray-level value. For complex images like medical ones, the use of single feature value cannot lead to satisfactory results; thus, it is important to consider more features to increase the expected performance [46]. In this work, we focus on investigating the so-called Gray Level Co-occurrence Matrix (GLCM)-based features, which has been shown to be efficient and offer interesting results in term of classification accuracy. GLCM matrix provides a co-occurrence matrix of joint probability density of the gray levels of two pixels. In this work, the second-order statistics are investigated to compute some features in order to well-discriminate lung abnormalities. In particular, the following features [47] are calculated for each image: contrast (large differences between neighboring pixels), correlation, energy, entropy, difference variance, difference entropy, inverse difference normalized, information measure of correlation, information measure of correlation. In our analysis, we focused on extracting the lungs area using image thresholding and segmentation processing which leads to identify the left and right lungs from CXR images. In order to remove noise, we applied the Gaussian filter. In Figure 6, we illustrate the obtained segmented lung using the above method. After isolating the lungs, we proceed with feature extraction step and then with classification using the proposed statistical model. The required time for feature extraction for each image is a few seconds and the model fitting taken between 20 to 30 min for the different data sets.

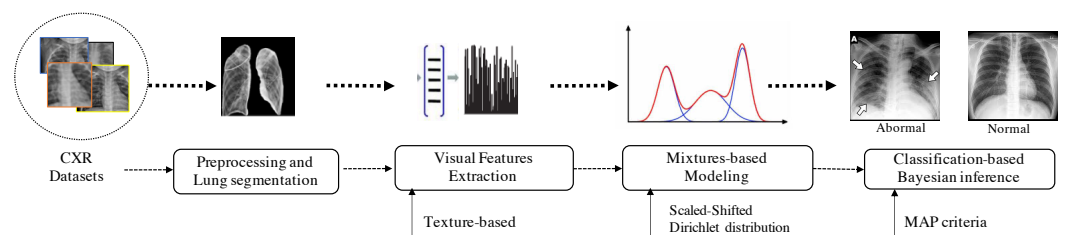


Figure 5. The pipeline of the proposed method. First, the lungs are segmented, then robust visual features are extracted. Features are modelled using the proposed mixture model (SSDDMM) and a Bayesian framework is applied to estimate the parameters of the model. Finally, images are classified on the basis of Bayes rule.

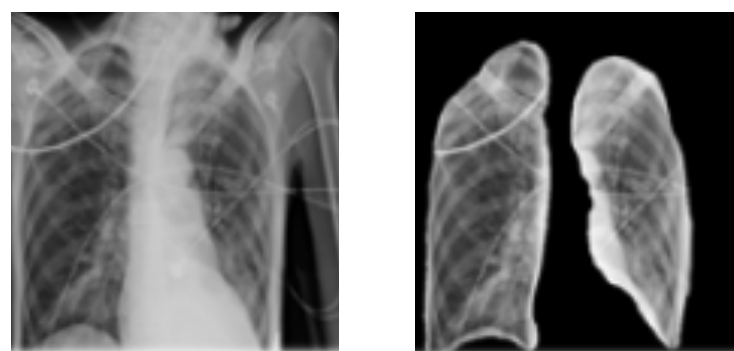


Figure 6. Process of lungs regions extraction applied on image sample from [42].

4.3. Results Analysis

In this section, we investigate our approach for COVID-19 detection. The ultimate first goal is to prove the potential of our Bayesian learning algorithm as compared to other learning method named maximum likelihood (ML) estimation. The second goal is to compare the performance of the proposed shifted-scaled Dirichlet mixture model with other methods which are Gaussian mixture-based, Gamma mixture-based, Dirichlet mixture-based and scaled Dirichlet mixture-based method. For performance investigation, we evaluate the performance of our Bayesian learning method and the rest of methods in terms of overall accuracy (ACC), detection rate (DR), and false-positive rate (FPR).

Tables 2–4, show the classification accuracies for the Test sets of each dataset when applying different generative approaches namely: Gaussian mixture model with maximum likelihood (**GMM-ML**), with Bayesian inference (**GMM-B**), Gamma mixture model with maximum likelihood (**ΓMM-ML**), Dirichlet mixture with maximum likelihood (**DMM-ML**), with Bayesian inference (**DMM-B**), scaled Dirichlet mixture with maximum likelihood (**SDMM-ML**), with Bayesian inference (**SDMM-B**), shifted scaled Dirichlet mixture with maximum likelihood (**SSDMM-ML**), and our proposed method named as shifted scaled Dirichlet mixture with Bayesian inference (**SSDMM-B**).

According to these tables, we can see clearly that, in general, all mixture models provide encouraging results taking into account the difficulty of the unsupervised learning problem. It is clear that our proposed Bayesian method for the shifted scaled Dirichlet mixture outperforms, according to the used metrics, the rest of methods. Indeed, our work has better accuracy as well as lowest false positive rate than both Dirichlet and Gaussian mixtures. We can also see that Bayesian learning provides better results than the ML approach for all models. As we can see, for CXR-COVID dataset, the SSDMM-B outperforms other models with accuracy of 89.57% compared to 88.08% for SDMM-B, 88.04% for DMM and 82.44% for GMM. Our Bayesian model is slightly better than SSDMM-MML [22]. Likewise, we came to the same conclusion for the other datasets and we reach the highest accuracy of 93.03% with our model **SSDMM-B** for the CXR-Pneumonia dataset. According to this last result, it is clear that the precision increases (and the false positive decreases) as the dataset size increases. This is can be viewed for CXR-Augmented-COVID and CXR-Pneumonia datasets which contain more images than CXR-COVID. On the basis of the overall accuracy (ACC) for three datasets (CXR-COVID, CXR-Pneumonia, and CXR-Augmented-COVID), it is obviously clear that the difference between the highest and lowest accuracy is between 5.2% and 7.46% for each dataset. The difference between some methods is about 2.26% which is also considered significant according to t-student test. The obtain results confirm the merits of the fully Bayesian formalism for shifted-scaled Dirichlet mixture which is more flexible (since it has more degrees of freedom) than the Dirichlet and the scaled Dirichlet mixtures. Its flexibility also makes it possible to easily integrate more knowledge and especially features selection mechanism into the proposed framework. On the other hand, even a small improvement is worthwhile taking into account the difficulty of the problem especially with the availability of strong machines to do the processing and simulations. Concerning the modeling uncertainty quantification, this is something that distinguishes our approach from deep learning models (black boxes). We are currently working with clinicians to be able to quantify the uncertainty and extract interpretations, as well as explanations from our models which is possible thanks to the generative nature of the deployed model.

Table 2. Overall accuracy for chest x-ray (CXR)-COVID Dataset.

Approach/Metrics	ACC(%)	DR(%)	FPR(%)
GMM-ML [48]	82.11	81.02	0.18
GMM-B [49]	83.44	82.14	0.17
ΓMM-ML [50]	85.22	83.76	0.16
DMM-ML [51]	87.99	87.88	0.14
DMM-B [52]	88.04	87.78	0.13
SDMM-ML [16]	88.08	87.84	0.13
SDMM-B [31]	88.22	88.07	0.13
SSDMM-ML [22]	89.13	88.24	0.12
SSDMM-B (our method)	89.57	88.61	0.12

Table 3. Overall accuracy for CXR-Pneumonia Dataset.

Approach/Metrics	ACC(%)	DR(%)	FPR(%)
GMM-ML [48]	87.66	85.80	0.13
GMM-B [49]	88.90	86.98	0.11
ΓMM-ML [50]	90.54	88.54	0.10
DMM-ML [51]	91.81	91.03	0.09
DMM-B [52]	92.01	91.33	0.09
SDMM-ML [16]	92.43	91.32	0.09
SDMM-B [31]	92.81	91.77	0.09
SSDMM-ML [22]	92.85	92.01	0.08
SSDMM-B (our method)	93.03	92.90	0.08

Table 4. Overall accuracy for CXR-Augmented COVID-19 Dataset.

Approach/Metrics	ACC(%)	DR(%)	FPR(%)
GMM-ML [48]	85.13	83.99	0.14
GMM-B [49]	86.77	84.08	0.13
ΓMM-ML [50]	90.24	89.14	0.10
DMM-ML [51]	88.01	87.57	0.12
DMM-B [52]	88.44	87.96	0.12
SDMM-ML [16]	89.01	88.12	0.11
SDMM-B [31]	89.88	89.12	0.10
SSDMM-ML [22]	90.10	89.01	0.09
SSDMM-B (our method)	90.33	89.12	0.09

It is also noted that the lung segmentation step is difficult particularly when it includes acute respiratory distress syndrome. This difficulty is due to the little contrast at the boundary of the lung. Moreover, when the number of images in this dataset is too small, the obtained results are lower than the case of big datasets. We can conclude that the obtained results are considered very encouraging given that we approach the classification problem in an unsupervised manner. In fact, the flexibility of the shifted-scaled mixture model and the robustness of texture-based features lead to more stable results. For COVID-19 identification through CXR images, the proposed fully Bayesian learning approach for SSDMM has confirmed that it is capable to discriminate images according to texture properties. In order to further improve these results, perhaps other descriptors are needed, especially the consideration of a robust feature selection mechanism to filter out unreliable features and keep only the most relevant ones. Please note that various studies have been proposed in the state of the art [53] which show that textures are very promising for many medical applications [54]. Here, the comparison between different feature-based techniques is beyond the scope of this article. Instead, we investigated in this work one robust texture-based descriptor to have interesting results for the classification of chest x-ray (CXR) images and corona virus covid-19 detection.

5. Conclusions

In this paper, we have addressed the problems of modeling and classification of multidimensional non-Gaussian data via a purely Bayesian learning approach based on a shifted scaled Dirichlet mixture model. We have especially tackled the problems of chest x-ray (CXR) images classification and COVID-19 detection. The flexibility and capability of the proposed statistical framework is evaluated through three public datasets related to COVID-19 and Pneumonia diseases. Unlike other statistical methods, which assume the heavy assumption that input data are Gaussian, which is not always true especially for real medical applications, the treated data in our work are modelled via non-Gaussian model and using finite mixtures of shifted scaled Dirichlet distributions that offer reasonable explanations. Our framework has provided promising results and outperforms other

methods. In particular, the Bayesian inference results are more interesting thanks to the consideration of the joint posterior distribution. In this work we have investigated an effective MCMC-based approximation technique given that exact inference in fully Bayesian methods is not easy to compute. Our implemented approach has also the advantage of being more general and extensible enough to be applied for large scale data presenting various infection's type. Future works could be devoted to extending the proposed framework via nonparametric approaches. Other promising future works include the integration of feature selection mechanism into the statistical model to improve the generalization capabilities. We hope also that many other real-world problems, including medical ones, will be addressed within the proposed framework.

Author Contributions: Conceptualization, S.B. and A.A.; methodology, N.B.; software, N.B.; validation, S.B. and A.A. and N.B.; formal analysis, S.B.; investigation, S.B. and A.A.; resources, N.B.; data curation, N.B.; writing—original draft preparation, S.B. and A.A.; writing—review and editing, N.B.; visualization, A.A.; supervision, N.B.; project administration, S.B.; funding acquisition, S.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Research Groups Program funded by Deanship of Scientific Research, Taif University, Ministry of Education, Saudi Arabia, under grant number 1-441-50.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: <https://github.com/ieee8023/covid-chestxray-dataset>; <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>; <https://data.mendeley.com/datasets/2fxz4px6d8/4>

Acknowledgments: Authors would like to thank the Deanship of Scientific Research, Taif University, Kingdom of Saudi Arabia, for their funding support under grant number 1-441-50.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Jacobi, A.; Chung, M.; Bernheim, A.; Eber, C. Portable chest X-ray in coronavirus disease-19 (COVID-19): A pictorial review. *Clin. Imaging* **2020**, *64*, 35–42. [[CrossRef](#)] [[PubMed](#)]
- Parveen, N.; Sathik, M.M. Detection of pneumonia in chest X-ray images. *J. X-ray Sci. Technol.* **2011**, *19*, 423–428. [[CrossRef](#)] [[PubMed](#)]
- Ginneken, B.V.; Stegmann, M.B.; Loog, M. Segmentation of anatomical structures in chest radiographs using supervised methods: A comparative study on a public database. *Med. Image Anal.* **2006**, *10*, 19–40. [[CrossRef](#)] [[PubMed](#)]
- Minaee, S.; Kafieh, R.; Sonka, M.; Yazdani, S.; Soufi, G.J. Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning. *Med. Image Anal.* **2020**, *65*, 101794. [[CrossRef](#)] [[PubMed](#)]
- Gordienko, Y.; Gang, P.; Hui, J.; Zeng, W.; Kochura, Y.; Alienin, O.; Rokovy, O.; Stirenko, S. Deep learning with lung segmentation and bone shadow exclusion techniques for chest x-ray analysis of lung cancer. In *International Conference on Computer Science, Engineering and Education Applications*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 638–647.
- Oliveira, L.L.G.; e Silva, S.A.; Ribeiro, L.H.V.; de Oliveira, R.M.; Coelho, C.J.; Andrade, A.L.S.S. Computer-aided diagnosis in chest radiography for detection of childhood pneumonia. *Int. J. Med. Inform.* **2008**, *77*, 555–564. [[CrossRef](#)]
- Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [[CrossRef](#)]
- Greenspan, H.; van Ginneken, B.; Summers, R.M. Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique. *IEEE Trans. Med. Imaging* **2016**, *35*, 1153–1159. [[CrossRef](#)]
- Zhao, B.; Feng, J.; Wu, X.; Yan, S. A survey on deep learning-based fine-grained object classification and semantic segmentation. *Int. J. Autom. Comput.* **2017**, *14*, 119–135. [[CrossRef](#)]
- Novikov, A.A.; Lenis, D.; Major, D.; Hladůvka, J.; Wimmer, M.; Bühler, K. Fully convolutional architectures for multiclass segmentation in chest radiographs. *IEEE Trans. Med. Imaging* **2018**, *37*, 1865–1876. [[CrossRef](#)]
- Xu, X.; Jiang, X.; Ma, C.; Du, P.; Li, X.; Lv, S.; Yu, L.; Chen, Y.; Su, J.; Lang, G.; et al. Deep Learning System to Screen Coronavirus Disease 2019 Pneumonia. *Engineering* **2020**, *6*, 1122–1129. [[CrossRef](#)]
- da Nóbrega, R.V.M.; Filho, P.P.R.; Rodrigues, M.B.; da Silva, S.P.P.; Júnior, C.M.J.M.D.; de Albuquerque, V.H.C. Lung nodule malignancy classification in chest computed tomography images using transfer learning and convolutional neural networks. *Neural Comput. Appl.* **2020**, *32*, 11065–11082. [[CrossRef](#)]

13. Candemir, S.; Jaeger, S.; Palaniappan, K.; Musco, J.P.; Singh, R.K.; Xue, Z.; Karargyris, A.; Antani, S.; Thoma, G.; McDonald, C.J. Lung Segmentation in Chest Radiographs Using Anatomical Atlases With Nonrigid Registration. *IEEE Trans. Med. Imaging* **2014**, *33*, 577–590. [[CrossRef](#)] [[PubMed](#)]
14. Xu, T.; Mandal, M.K.; Long, R.; Cheng, I.; Basu, A. An edge-region force guided active shape approach for automatic lung field detection in chest radiographs. *Comput. Med. Imaging Graph.* **2012**, *36*, 452–463. [[CrossRef](#)] [[PubMed](#)]
15. Mashrgy, M.A.; Bdiri, T.; Bouguila, N. Robust simultaneous positive data clustering and unsupervised feature selection using generalized inverted Dirichlet mixture models. *Knowl. Based Syst.* **2014**, *59*, 182–195. [[CrossRef](#)]
16. Oboh, B.S.; Bouguila, N. Unsupervised learning of finite mixtures using scaled dirichlet distribution and its application to software modules categorization. In Proceedings of the 2017 IEEE International Conference on Industrial Technology (ICIT), Toronto, ON, Canada, 22–25 March 2017; pp. 1085–1090.
17. Channoufi, I.; Bourouis, S.; Bouguila, N.; Hamrouni, K. Image and video denoising by combining unsupervised bounded generalized gaussian mixture modeling and spatial information. *Multimed. Tools Appl.* **2018**, *77*, 25591–25606. [[CrossRef](#)]
18. Fan, W.; Bouguila, N. Spherical data clustering and feature selection through nonparametric Bayesian mixture models with von Mises distributions. *Eng. Appl. Artif. Intell.* **2020**, *94*, 103781. [[CrossRef](#)]
19. Najar, F.; Bourouis, S.; Bouguila, N.; Belghith, S. Unsupervised learning of finite full covariance multivariate generalized Gaussian mixture models for human activity recognition. *Multimed. Tools Appl.* **2019**, *78*, 18669–18691. [[CrossRef](#)]
20. Najar, F.; Bourouis, S.; Zaguia, A.; Bouguila, N.; Belghith, S. Unsupervised Human Action Categorization Using a Riemannian Averaged Fixed-Point Learning of Multivariate GGMM. In Proceedings of the Image Analysis and Recognition-15th International Conference, ICIAR, Póvoa de Varzim, Portugal, 27–29 June 2018; pp. 408–415.
21. Bourouis, S.; Mashrgy, M.A.; Bouguila, N. Bayesian learning of finite generalized inverted Dirichlet mixtures: Application to object classification and forgery detection. *Expert Syst. Appl.* **2014**, *41*, 2329–2336. [[CrossRef](#)]
22. Alsuroji, R.; Zamzami, N.; Bouguila, N. Model Selection and Estimation of a Finite Shifted-Scaled Dirichlet Mixture Model. In Proceedings of the 17th IEEE International Conference on Machine Learning and Applications, ICMLA, Orlando, FL, USA, 17–20 December 2018; pp. 707–713.
23. Alroobaea, R.; Rubaiee, S.; Bourouis, S.; Bouguila, N.; Alsufyani, A. Bayesian inference framework for bounded generalized Gaussian-based mixture model and its application to biomedical images classification. *Int. J. Imaging Syst. Technol.* **2020**, *30*, 18–30. [[CrossRef](#)]
24. Kayabol, K.; Kutluk, S. Bayesian classification of hyperspectral images using spatially-varying Gaussian mixture model. *Digit. Signal Process.* **2016**, *59*, 106–114. [[CrossRef](#)]
25. Li, Z.; Xia, Y.; Ji, Z.; Zhang, Y. Brain voxel classification in magnetic resonance images using niche differential evolution based Bayesian inference of variational mixture of Gaussians. *Neurocomputing* **2017**, *269*, 47–57. [[CrossRef](#)]
26. Li, F.; Perona, P. A Bayesian Hierarchical Model for Learning Natural Scene Categories. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), San Diego, CA, USA, 20–26 June 2005; pp. 524–531.
27. Bourouis, S.; Al-Osaimi, F.R.; Bouguila, N.; Sallay, H.; Aldosari, F.M.; Mashrgy, M.A. Bayesian inference by reversible jump MCMC for clustering based on finite generalized inverted Dirichlet mixtures. *Soft Comput.* **2019**, *23*, 5799–5813. [[CrossRef](#)]
28. Robert, C. *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*; Springer Science & Business Media: New York, NY, USA, 2007.
29. Marin, J.M.; Robert, C. *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*; Springer Science & Business Media: New York, NY, USA, 2007.
30. Chen, P.; Nelson, J.D.B.; Tourneret, J. Toward a Sparse Bayesian Markov Random Field Approach to Hyperspectral Unmixing and Classification. *IEEE Trans. Image Process.* **2017**, *26*, 426–438. [[CrossRef](#)] [[PubMed](#)]
31. Bourouis, S.; Laalaoui, Y.; Bouguila, N. Bayesian frameworks for traffic scenes monitoring via view-based 3D cars models recognition. *Multimed. Tools Appl.* **2019**, *78*, 18813–18833. [[CrossRef](#)]
32. Barber, D.; Williams, C.K.I. Gaussian Processes for Bayesian Classification via Hybrid Monte Carlo. In Proceedings of the Advances in Neural Information Processing Systems 9, NIPS, Denver, CO, USA, 2–5 December 1996; Mozer, M., Jordan, M.I., Petsche, T., Eds.; MIT Press: Cambridge, MA, USA, 1996; pp. 340–346.
33. Bourouis, S.; Al-Osaimi, F.R.; Bouguila, N.; Sallay, H.; Aldosari, F.M.; Mashrgy, M.A. Video Forgery Detection Using a Bayesian RJMCMC-Based Approach. In Proceedings of the 14th IEEE/ACS International Conference on Computer Systems and Applications, AICCSA 2017, Hammamet, Tunisia, 30 October–3 November 2017; pp. 71–75.
34. Fan, W.; Bouguila, N.; Bourouis, S.; Laalaoui, Y. Entropy-based variational Bayes learning framework for data clustering. *IET Image Process.* **2018**, *12*, 1762–1772. [[CrossRef](#)]
35. Bourouis, S.; Zaguia, A.; Bouguila, N. Hybrid Statistical Framework for Diabetic Retinopathy Detection. In *Image Analysis and Recognition, Proceedings of the 15th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, 27–29 June 2018*; Lecture Notes in Computer Science; Campilho, A., Karray, F., ter Haar Romeny, B.M., Eds. Springer: Cham, Switzerland, 2018; Volume 10882, pp. 687–694.
36. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B* **1977**, *39*, 1–38.

37. Bouguila, N. Bayesian hybrid generative discriminative learning based on finite Liouville mixture models. *Pattern Recognit.* **2011**, *44*, 1183–1200. [[CrossRef](#)]
38. Gelman, A.; Carlin, J.B.; Stern, H.S.; Rubin, D.B. *Bayesian Data Analysis*, 3rd ed.; Chapman and Hall/CRC: New York, NY, USA, 2013.
39. Geiger, D.; Heckerman, D. Parameter priors for directed acyclic graphical models and the characterization of several probability distributions. In Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence, Stockholm, Sweden, 30 July–1 August 1999; pp. 216–225.
40. Congdon, P. *Applied Bayesian Modelling*; John Wiley and Sons: Hoboken, NJ, USA, 2003.
41. Chib, S.; Greenberg, E. Understanding the Metropolis-Hastings Algorithm. *Am. Stat.* **1995**, *49*, 327–335.
42. Cohen, J.P.; Morrison, P.; Dao, L.; Roth, K.; Duong, T.Q.; Ghassemi, M. COVID-19 Image Data Collection: Prospective Predictions Are the Future. *arXiv* **2020**, arXiv:2006.11988.
43. Zu, Z.Y.; Jiang, M.D.; Xu, P.P.; Chen, W.; Ni, Q.Q.; Lu, G.M.; Zhang, L.J. Coronavirus disease 2019 (COVID-19): A perspective from China. *Radiology* **2020**, *296*, 200490. [[CrossRef](#)]
44. Alqudah, A.; Qazan, S. Augmented COVID-19 X-ray images dataset. *Mendeley Data* **2020**, *4*. [[CrossRef](#)]
45. Mooney, P. Chest X-ray Images (Pneumonia). 2020. Available online: <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia> (access on November 2020).
46. Xie, J.; Jiang, Y.; Tsui, H. Segmentation of kidney from ultrasound images based on texture and shape priors. *IEEE Trans. Med. Imaging* **2005**, *24*, 45–57. [[PubMed](#)]
47. Pourghassem, H.; Ghassemian, H. Content-based medical image classification using a new hierarchical merging scheme. *Comput. Med. Imaging Graph.* **2008**, *32*, 651–661. [[CrossRef](#)] [[PubMed](#)]
48. Fernando, B.; Fromont, É.; Muselet, D.; Sebban, M. Supervised learning of Gaussian mixture models for visual vocabulary generation. *Pattern Recognit.* **2012**, *45*, 897–907. [[CrossRef](#)]
49. Figueiredo, M.A.T.; Jain, A.K. Unsupervised Learning of Finite Mixture Models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 381–396. [[CrossRef](#)]
50. Sallay, H.; Bourouis, S.; Bouguila, N. Online Learning of Finite and Infinite Gamma Mixture Models for COVID-19 Detection in Medical Images. *Computers* **2021**, *10*, 6. [[CrossRef](#)]
51. Bouguila, N.; Ziou, D. Using unsupervised learning of a finite Dirichlet mixture model to improve pattern recognition applications. *Pattern Recognit. Lett.* **2005**, *26*, 1916–1925. [[CrossRef](#)]
52. Ma, Z.; Rana, P.K.; Taghia, J.; Flierl, M.; Leijon, A. Bayesian estimation of Dirichlet mixture model with variational inference. *Pattern Recognit.* **2014**, *47*, 3143–3157. [[CrossRef](#)]
53. Smith, G.; Burns, I. Measuring texture classification algorithms. *Pattern Recognit. Lett.* **1997**, *18*, 1495–1501. [[CrossRef](#)]
54. Melendez, J.; van Ginneken, B.; Maduskar, P.; Philipsen, R.H.H.M.; Reither, K.; Breuninger, M.; Adetifa, I.M.O.; Maane, R.; Ayles, H.; Sánchez, C.I. A Novel Multiple-Instance Learning-Based Approach to Computer-Aided Detection of Tuberculosis on Chest X-Rays. *IEEE Trans. Med. Imaging* **2015**, *34*, 179–192. [[CrossRef](#)]