

# Latent Factor Model Based Classification for Detecting Abnormalities in Retinal Images

Syed Tabish Abbas and Jayanthi Sivaswamy  
CVIT, IIT-Hyderabad  
Hyderabad, India

syed.abbas@research.iit.ac.in, jsivaswamy@iit.ac.in

## Abstract

*Abnormality detection in medical images is a critical problem across image modalities and organs. Many approaches to automatic abnormality detection use discriminative methods, based on domain knowledge, to address the problem. In this paper, we investigate the effectiveness of a generative model with no assumption of domain knowledge. We propose a method for classification of tissues based on Latent Factor analysis, and demonstrate it on colour retinal images with Diabetic Retinopathy-related abnormalities. A generative model based on Gaussian latent dictionaries is used to model various structures present at a patch level in an image. The model is used to classify a given patch into one of 5 classes: namely dark and bright lesions, neo-vascularisation (NV), plain tissue background and background with vessel. Evaluation of the proposed method was done on 3 different datasets. Same and cross-dataset validation of the method yields an area under the receiver operating characteristic curve (AUC) of 0.85 for abnormality detection. The method was also modified to address a challenging problem of NV detection by posing it as a 2-class problem and the AUC for the same was found to be 0.92. This establishes the potential of LF model for abnormality detection.*

## 1. Introduction

Abnormality detection is a key problem in medical image analysis. The problem is challenging because of inter-subject variabilities, noise in data, minor structural/appearance changes with ethnicity etc. Identifying regions which deviate from normal anatomy while diagnosing a disease is a standard practise in evidence-based medicine. However, the approach taken by various CAD systems is, in contrast, significantly different as they employ elaborate multiple stages to detect various types of defects/abnormalities in an imaged tissue.

Retinal image analysis is an instance of such an approach. Diabetic Retinopathy (DR) manifests itself as various symptomatic changes in retinal image. Automatic detection systems employ separate methods for detection of vessels, optic disc, bright lesions, dark lesions etc., in the image. Existing approaches for abnormality detection typically adopt the following pipeline: (i) A pre-processing stage for multiple tasks including image enhancement; background suppression; detection and masking of vessels and the optic disc. The latter serves to suppress false positives during bright and dark lesion detections. (ii) a candidate region/patch extraction stage for which various techniques have been attempted. These include super pixel-like region extraction [10], morphological processing [5], model based matching [8] etc. (iii) a final classification stage where hand crafted features from each candidate region/patch are extracted. These features are specific to the abnormality or tissue type based on visual appearance, domain knowledge, etc. Feature extraction is done using multiple methods such as a two-stage feature selection by quadratic discriminant analysis followed by classification by k-nn classifier [10] and ensemble of classifiers [6]. Some approaches [1] do not have a separate candidate extraction stage, however, significant analysis is required in the feature selection stage.

While the above approaches belong to the class of abnormality-specific discriminative classification, a semi-parametric generative model in the form of Gaussian Mixture Model has also been reported [9]. However, this does not take spatial correlation into account and is limited to hard exudate detection. We propose the use of a *non-parametric generative model* for classification of various tissue types (both normal as well as abnormal). The proposed method does not require any preprocessing or extraction of hand crafted features and is robust to variation due to noise, illumination etc. We model various tissues at a patch level and classify them as belonging to different categories. In contrast to most existing approaches

which use only one (green) image channel, we use color information from all three channels to build our model.

There are various technical difficulties in building such a model. The sheer dimensionality of images even at patch level makes generative modelling a daunting task. Such modelling, using just the second-order statistics, involves estimating  $O(n^4)$  parameters, for a patch of size  $n \times n$ . We use latent factor analysis to make the problem tractable.

A very wide variety of approaches to latent factor analysis have been proposed in statistics literature. We use a non-parametric, latent dictionary-based factor model [2]. The motivation for using such a method is two fold. Firstly, the method handles the dimensionality problem and secondly, the method also has a clear physical interpretation in terms of images. The stochastic nature of the model incorporates noise handling natively and is inherently robust to variations in illumination, tissue pigmentation or imaging noise.

## 2. Method

A sequence of random variables  $(y_1, y_2, \dots, y_t, \dots)$  is said to be a Gaussian process iff for every finite subset of indices  $\{1, 2, \dots, t, \dots\}$  the sub-sequence is a multivariate Gaussian distribution. We model various tissues/structures in an image, at a patch level, as a Gaussian process.

Consider an  $n \times n = q$  image patch. let  $y_j$  be vectorized version of the patch, such that,  $y_j$  is a  $q \times t$  matrix where  $t$  denotes the number of channels in the image.

$y$  is assumed to be a Gaussian random process with channel-dependent covariance and channel and sample dependant mean. Formally, let  ${}^i y_j \in \mathbb{R}^q$  be the  $j^{\text{th}}$  channel of the  $i^{\text{th}}$  instance of a patch of a certain tissue type. The mean  ${}^i \mu_j$  of each channel is assumed to vary across patches ( $i$ ) to account for intra-image and inter-image illumination variation, whereas, the covariance  $\Sigma_j$  is assumed to be constant for different patch instances. Constant covariance models the uniformity of structure and spatial relation irrespective of bias/illumination variation or other noise. We then have,

$${}^i y_j \sim N({}^i \mu_j, \Sigma_j)$$

where  ${}^i \mu_j$  is a class and instance dependent  $q \times 1$  vector.  $\Sigma_j$  is a  $q \times q$  class-dependent matrix, which has in general  $q(q+1)/2$  parameters.

Such a model will hence require estimating  $O(q^2)$  parameters, making it impractical even for moderate patch size  $q$ . To make the model practically plausible we reduce the dimensionality of the problem using latent factor analysis[2]. We assume the following latent factor model for  ${}^i y_j$ .

$${}^i y_j = \Lambda \eta_j + \epsilon_j, \quad \epsilon \sim N(0, \Sigma_0)$$

where  $\eta_j$ , is a  $k$ -dimensional latent factor vector, and  $\Lambda$  is a  $q \times k$  loading matrix.  $\epsilon_j \sim N(0, \Sigma_0)$  is the modelling

noise with  $\Sigma_0 = \text{diag}(\sigma_1^2, \dots, \sigma_q^2)$ . The above expansion of  $y_j$  leads to the following decomposition of  $\Sigma$ .

$$\Sigma_j = \Lambda \Lambda^T + \Sigma_0$$

where  $T$  denotes transpose operation.  $\Lambda$  is a  $q \times k$  matrix with  $k \ll q$ . The decomposition factors the covariance matrix into a sum of a low rank and a diagonal matrix. Note that,  $\Lambda$  has  $O(qk)$  parameters, while  $(\Sigma_0)$  has  $O(q)$  parameters. Such modelling, therefore, reduces the number of independent parameters of  $\Sigma_j$  from  $q(q+1)/2$  to  $q(k+1)$ . However, if  $q$  is large, estimation of  $q(k+1)$  parameters may still poses a significant challenge. Further reduction in dimensionality is achieved by letting  $\Lambda$  to be a linear combination of *latent covariance dictionary functions*  $\zeta_{lk}$  such that,

$$\Lambda = \Theta \zeta = \sum_{l=1}^L \theta_{rl} \zeta_{ls}$$

where  $\Theta \in \mathbb{R}^{q \times L}$  is a non-stochastic class-dependant loading matrix which weighs the *covariance dictionary functions*  $\zeta \in \mathbb{R}^{L \times k}$ . We specify a prior on the covariance dictionary function  $\zeta_{lk}(\cdot) \sim N(0, d)$ , where  $d$  is a squared exponential covariance function. Note that both  $k, L \ll q$ . Using the above equation,  $y_j$  may expressed as,

$${}^i y_j = \sum_{m=1}^k {}^i \eta_m \sum_{l=1}^L \theta_{jl} \zeta_{lm} + {}^i \epsilon_j$$

In order to allow the mean  ${}^i \mu$  to evolve non-parametrically and encode dependencies across channels,  $\eta_j$  is modelled in terms of *dictionary functions*  $\psi_i$  as follows

$$\eta_i = \psi_i + \nu_i, \quad \nu_i \sim N(0, I_k)$$

where  $\psi_i = [\psi_i^1, \dots, \psi_i^k]^T$ . Each of the  $k$  *latent dictionary functions* of  $\psi_i$  is a Gaussian random function, with squared exponential correlation.

After marginalizing the noise terms  $\epsilon$  and  $\nu$ , we obtain the mean and covariance structure of  ${}^i y_j$  as follows.

$${}^i \mu_j = \Theta \zeta_j \psi_i$$

$$\Sigma_j = \Theta \zeta_j \zeta_j^T \Theta^T + \Sigma_0$$

where  $\Theta, \zeta$ , and  $\psi$  are learned from labelled training data [2].

Next, we demonstrate how the above latent factor model (LFM) can be used to detect multiple abnormalities related to DR, from retinal images.

## 3. Experiments And Results

Diabetic retinopathy affected retinal images contain multiple lesions present at different locations depending

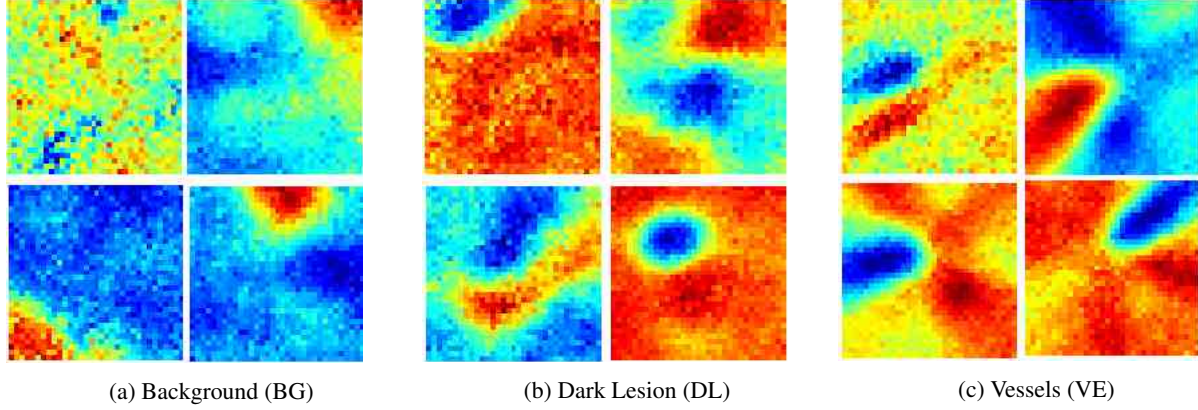


Figure 1: Sample loading factors ( $\theta$ ) for background, dark lesions and vessels.

on the stage of the disease. Given an image patch, we aim to classify a given patch, into one of 5 different classes: dark lesions (DL) or hemorrhages, bright lesions (BL) or hard/soft exudates, neo-vascularisation (NV), plain tissue background (BG) and background with vessels (VE).

In the training stage, labelled patches are used to learn the structure of  $\Sigma$ , and  $\mu$  in the form of  $\Theta$ ,  $\zeta$ , and  $\psi$  for each class [2]. Figure 1 shows a visualisation (as a heat map) of the columns of  $\Theta$  for 3 different classes of a retinal image patch. The visualisation is achieved by reshaping the columns into patches of size  $n \times n (= q)$ . It can be noted that the  $\theta$  columns for the background class have a uniform texture with some variations, which is typical for retinal background tissue; those of the dark lesions class contain irregular shaped spots akin to their structure while those of the vessel class have linear structures. Figure 1 illustrates the physical interpretation of the model where each tissue patch is composed of a weighted combination of these latent factors.

In the testing stage, given an unlabelled patch, a posterior probability score is computed and used to classify the patch. The posterior probability computation involves approximating the following integral.

$$p(y_j | \{y_j^{train}\}_{i=1}^5) = \int p(y_j | \Theta, \zeta, \Sigma_0, \psi) d\Omega$$

where,  $\Omega = \{\Theta, \zeta, \Sigma, \psi_i\}$  and  $j = R, G, B$  colour channels. The integral involves matrices of size  $(q \times j)^2$  making it computationally infeasible. Using Monte Carlo technique, the integral is approximated by the following sum

$$p(y_j | \{y_j^{train}\}_{i=1}^5) = \frac{1}{M} \sum_{m=1}^M p(y_j | \Omega_m)$$

In our experiments, 150 samples ( $M = 150$ ) were drawn to approximate the posterior. The size of each patch was

chosen to be  $33 \times 33 = q$  as it fits the typical size and structure of abnormalities we aim to classify. Each patch was chosen to be represented by the raw R,G,B ( $t = 3$ ) colour vector, rather than higher order features, in order to assess the effectiveness of LFM for multi-class classification. Image patches from 3 different datasets were used in our experiments: DriDB [7] with 50 images, DiaretDB1 [3] with 89 images and a private dataset (KHPDR) shared by the authors of [11] with 76 images. The last dataset was used only for drawing NV patches.

Training was done on a set of 750 patches ( $=150 \times 5$  classes). For all classes, except NV, patches were drawn from the DRiDB while NV patches were drawn from KHPDR. Testing was done on 2 sets: Set A, made of 200 patches (ensuring no overlap with training set) drawn from DriDB and KHPDR; Set B, made of 1732 patches drawn from DiaretDB1, for cross-validation.

Two types of experiments were performed. The first aimed at assessing the potential of the LFM-based approach for multi-class classification, while the second experiment aimed at assessing it for a 2-class classification problems. For the first experiment, the Mahalanobis distance ( $D_M$ ) between the maximum likelihood estimate of a class mean  $\mu$  to other classes distributions was measured. This metric has also been used in [1] for detecting DR using AM-FM features. For the second experiment, the area under the curve (AUC) of the receiver operating characteristic curves (ROC) was computed. The computed  $D_M$  between the 5 classes is presented in Table 1. The discriminatory power is proportional to the distance. The results indicate that BG and NV are most difficult to distinguish while VE and BL are the easiest. While distinguishing DL from VE or BG is difficult because of their similar color, confusion between DL and BG is compounded by their structural similarity. This is confirmed by the smaller  $D_M$  values between DL and BG. All these observations are consistent with the appearance of sample patches in Fig 2.

A gross problem of interest in DR detection is assessing if a target abnormality is present/absent in a given image patch. This can be posed as a 2-class problem and addressed using LFM-based approach. An SVM (with 3<sup>rd</sup> degree polynomial kernel) was trained on the log posterior probability scores generated by the LFM for each image patch. This was done for the following combinations: *Lesions* (BL+DL) versus others (NV, BG, VE); NV versus others (BL,DL,BG,VE) and *Normal* (BG,VE) versus *Abnormal* (BL,DL and NV). Figures 4, 3 show the ROC plots with a 5-fold validation on the test sets A and B for these combinations. The plots in Fig.4 for Set A were derived using a 15-dimensional (=5 classes × 3 colour channels) vector of probability scores while those in Fig.3 for Set B were derived using a 12-dimensional (= 4 × 3) vector as this set does not have NV samples.

The AUC values obtained on Set A (0.8533) and Set B (0.8500) for Abnormal/Normal classification are very close, indicating the robustness of the model across datasets. In terms of assessing presence/absence of abnormalities, the performance is best for NV (AUC=0.92) and worst (0.79) for Lesions class. The reason for this trend can be explained based on the  $D_M$  values in Table 1 which indicates that it is difficult to discriminate between NV and BG patches and the latter is now a part of ‘others’ class. Likewise, as per the  $D_M$  values, the DL class (a sub-class of Lesion class) is difficult to distinguish from both BG and VE classes, both of which are in the ‘others’. This causes the AUC to be the lowest for Lesion vs ‘others’. From these points it also follows that the performance of Abnormal/Normal classification (AUC=0.85) will be between that of NV/others and Lesions/others. It is noteworthy that NV detection, in general, is very challenging and has received scant attention in literature. Hence, a high AUC of 0.92 for NV/others, where ‘others’ include lesions and VE classes, is significant and underscores the strength of the LFM approach.

DR detection in [1], also aimed at discriminating between multiple classes at a patch level, reports assessment results (using  $D_M$ ) on a public dataset. However, the annotations publicly available are only at image-level and not region/lesion-level. Hence, it is not possible to do a proper comparative evaluation. Furthermore, the numerical values of  $D_M$  cannot be compared directly, as the dimensionality of the underlying feature spaces are different. While only hard exudates are considered in [1], our experiments included both hard as well as soft exudates in the BL class. Nevertheless, some broad common trends can be observed: it is difficult to distinguish DL class from VE and BG classes;  $D_M$  values exhibit largest variability in the case of NV discrimination against other classes (3rd row of Table 1). There is also a notable difference in that BL and VE classes are reported in [1] to be least separable, whereas they are maximally separable in our case.

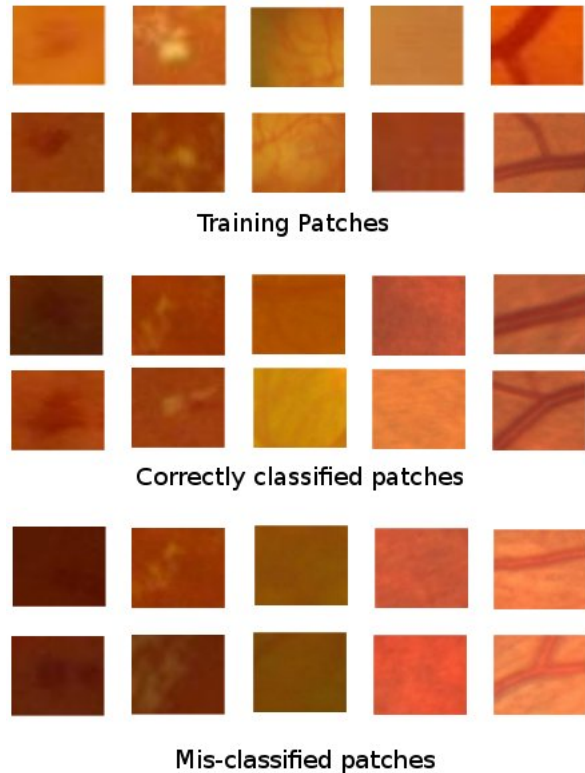


Figure 2: Sample patches (columns 1 to 5) with dark lesion, bright lesion, neo-vascularisation (NV), background and vessel. Training and testing samples are from DridB and DiaretdB, respectively. All NV samples are from KHPDR.

Table 1: Average Mahalanobis distance between classes.

	DL	BL	NV	BG	VE
DL	0	2.5387	2.0602	1.3211	1.4910
BL	2.5387	0	3.4514	2.0770	6.3084
NV	2.0602	3.4514	0	0.5214	3.6098
BG	1.3211	2.0770	0.5214	0	2.4756
VE	1.4910	6.3084	3.6098	2.4756	0

#### 4. Concluding remarks

In this paper, we investigated the application of a Gaussian Process-based LFM for classifying DR-related abnormalities in retinal images. While extensive work has been reported on discriminative models for the problem, there is a lacuna on the use of generative models. Generative models have been proven [4] to outperform discriminative approaches when the training set size is small, a scenario that is typified by the medical domain, where generating annotations is an arduous task.

Our method is general and applicable to different abnor-

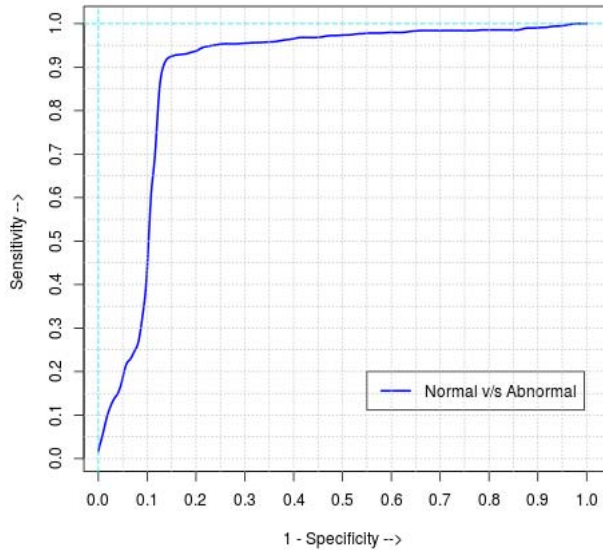


Figure 3: ROC plot for normal vs abnormal classification on DiaretDB.

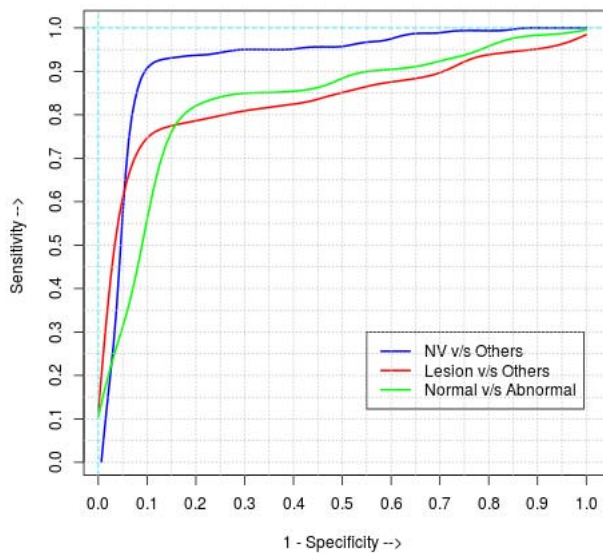


Figure 4: ROC plots for DriDB. Normal vs abnormal classification (green), NV vs others (blue) and Lesions vs others (red).

malities, contrary to most methods reported in literature [5], [6], [8], [12] which are specific to particular abnormalities. It does not assume any domain knowledge, thus obviating the need for extraction of hand crafted features. Methods like [1] which handle multiple abnormalities entail an elaborate procedure for extraction and selection of features. The nature of modeling causes the method to be immune to noise and variations due to illumination etc. The robustness of the method is illustrated by consistent performance

across datasets. The bottle neck of the system is in the computation of the posterior probability, which needs to be explored in future work.

## References

- [1] C. Agurto et al. Multiscale am-fm methods for diabetic retinopathy lesion detection. *Medical Imaging, IEEE Transactions on*, 2010. 1, 3, 4, 5
- [2] E. Fox and D. Dunson. Bayesian nonparametric covariance regression. *arXiv preprint arXiv:1101.2017*, 2011. 2, 3
- [3] T. Kauppi et al. The diaretdb1 diabetic retinopathy database and evaluation protocol. In *BMVC*, 2007. 3
- [4] A. Y. Ng and M. I. Jordan. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. In *Advances in Neural Information Processing Systems 14*. MIT Press, 2002. 4
- [5] M. Niemeijer et al. Automatic detection of red lesions in digital color fundus photographs. *Medical Imaging, IEEE Transactions on*, 2005. 1, 5
- [6] M. Niemeijer et al. On combining computer-aided detection systems. *Medical Imaging, IEEE Transactions on*, 2011. 1, 5
- [7] P. Prentas et al. Diabetic retinopathy image database (driDB): a new database for diabetic retinopathy screening programs research. In *Image and Signal Processing and Analysis (ISPA)*, 2013. 3
- [8] G. Quellec et al. Optimal filter framework for automated, instantaneous detection of lesions in retinal images. *Medical Imaging, IEEE Transactions on*, 2011. 1, 5
- [9] C. I. Sánchez et al. Retinal image analysis based on mixture models to detect hard exudates. *Medical Image Analysis*, 2009. 1
- [10] L. Tang et al. Splat feature classification with application to retinal hemorrhage detection in fundus images. *Medical Imaging, IEEE Transactions on*, 2013. 1
- [11] M. Vatanparast and A. Harati. A feasibility study on detection of neovascularization in retinal color images using texture. In *ICCKE*, 2012. 3
- [12] D. Zhang et al. Robust hemorrhage detection in diabetic retinopathy image. In *Pattern Recognition (ACPR), 2011 First Asian Conference on*, pages 209–213. IEEE, 2011. 5