

Classifying and grouping mammography images into communities using Fisher information networks to assist the diagnosis of breast cancer

Meenal Srivastava¹, Ivan Olier¹, Patrick Riley¹, Paulo Lisboa¹,
Sandra Ortega-Martorell¹

¹ Department of Applied Mathematics, Liverpool John Moores University, Byrom Street,
Liverpool, UK.

meenal.srivastava@gmail.com, I.A.OlierCaparoso@ljmu.ac.uk,
P.J.Riley@2014.ljmu.ac.uk, P.J.Lisboa@ljmu.ac.uk, S.OrtegaMartorell@ljmu.ac.uk

Abstract. The aim of this paper is to build a computer based clinical decision support tool using a semi-supervised framework, the Fisher Information Network (FIN), for visualization of a set of mammographic images. The FIN organizes the images into a similarity network from which, for any new image, reference images that are closely related can be identified. This enables clinicians to review not just the reference images but also ancillary information e.g. about response to therapy. The Fisher information metric defines a Riemannian space where distances reflect similarity with respect to a given probability distribution. This metric is informed about generative properties of data, and hence assesses the importance of directions in space of parameters. It automatically performs feature relevance detection. This approach focusses on the interpretability of the model from the standpoint of the clinical user. Model predictions were validated using the prevalence of classes in each of the clusters identified by the FIN.

Keywords: Medical image analysis, Community detection in medical images, Fisher Distance, Fisher Information Metric, Information Geometry, Riemannian Space

1 Introduction

Breast cancer is the most frequent cancer among women, impacting 2.1 million women each year, and also causes the greatest number of cancer-related deaths among women (1). This type of cancer usually takes time to develop and symptoms become evident very late. Currently, there is no effective way to cure later stage breast cancer, therefore early and accurate detection of tumor plays a vital role in improving the prognosis, as it allows for better treatment planning.

Besides physical examination, commonly used modalities for breast screening are mammography, ultrasonography, Magnetic Resonance Imaging (MRI) and core-needle biopsy. Among these techniques, mammography is considered the best, cheapest way to detect the tumor (2). However sensitivity of mammography can vary considerably due to factors like radiologist's experience, human error and image quality (3). Visual clues during early stages are subtle and varied in appearance making diagnosis difficult.

Many times, abnormalities are hidden by breast tissue structure. According to statistical reports, patients with dense breasts have high chances of receiving false negatives for lesions (4). Appearance of normal tissue is also highly variable and complex in mammograms, making tumor identification more difficult. Micro-calcifications are frequently missed. In the current practice, pathological confirmation of malignancy and tumor grade characterization is done with a biopsy, which is an invasive, painful procedure, that also carries risk of tumor cell migration (5).

Most of existing work done in breast cancer classification is categorizing into normal and tumor class. However, tumor can be either benign or malignant, with the latter being the one requiring treatment. Hence, the two classes need to be differentiated as well. In this paper we attempt to classify mammography images into these three classes: malignant, benign or normal. Our aim is to propose a methodology that can assist clinicians in managing their breast cancer patients, by visualizing mammographic images in a different, novel way, using Fisher Information Network (FIN) (6).

FIN provides a global view of data due to the use of Fisher metric and displays a meaningful structure that implicitly informs about underlying class probabilities. The FIN framework (7) can be used both for visualization of data and for constructing interpretable retrieval-based classifier since connection weights contain accurate information about similarity between data points. This approach focusses on both accurate prediction as well as interpretability of output.

In our proposed methodology we construct a FIN using probability density estimates calculated for three classes- normal, benign, malignant. The aim is to detect underlying patterns and structure in breast cancer images. We test our approach using an existing, publicly available database (8). By producing FIN's visualization of similarity networks we expect to elucidate the underlying data structure, community membership and class prediction. In addition, by dividing the network into communities and stratifying the data according to classification labels, we expect to be able to provide instances similar to new query image, which can then be analyzed in order to better understand new instances.

The paper is organized as follows: Section 2 describes the dataset used for constructing and testing the FIN. Section 3 describes the methodology to construct the network. In Section 4, we present the results and discuss its significance and implications. Finally, Section 5 presents our conclusions.

2 Material

Images from the Mammographic Image Analysis Society (Mini MIAS) database (8), [available at <http://peipa.essex.ac.uk/info/mias.html>] are used in this paper. The dataset is arranged in pairs of films, where each pair represents the left (even filename numbers) and right (odd filename numbers) mammogram of a single patient. The dataset is composed of 326 mammograms of right and left breast, from 161 patients, where 51 observations were labeled as malignant, 66 as benign and 209 as normal/healthy. In this work we randomly under-sampled the normal observations to balance the number of cases per class in the dataset. Hence, the dataset used in this paper contains 66 benign, 51 malignant and 62 normal observations, for a total of 179.

The dataset includes radiologist's "truth" markings for locations of abnormalities present in images. For these images, coordinates for center as well as approximate radius of circle enclosing the abnormality, is available. In case of calcifications, center locations and radii are for clusters rather than individual calcifications. The dataset also contains information related to background tissue of patients, and margin/shape of abnormalities. Table 1 shows the details of the dataset used in this paper regarding tissue, abnormalities and class.

Table 1 Mini MIAS Dataset used in Project

Tumour shape	Tumour type		Breast type		
	Benign	Malignant	Dense	Fatty	Glandular
Architectural Distortions	9	10	7	6	6
Asymmetric	6	9	7	4	4
Calcification	13	12	14	5	6
Circumscribed	20	4	3	13	8
Miscellaneous	7	8	2	8	5
Spiculated	11	8	7	5	7

Classes	Breast type			Total per class
	Dense	Fatty	Glandular	
Benign	23	23	20	66
Malignant	17	18	16	51
Normal	11	31	20	62

Figure 1 is an image from the dataset illustrating the challenges faced in automatic evaluation/classification of these medical images. Artefacts like patient's names, labels, etc. are present in the images. Pectoral muscles having same intensity as tumors are visible in mediolateral oblique (MLO) view and needs to be removed. The intensity of normal region in dense breast is similar to that of tumor region in fatty and glandular breast (see Figure 2 b, c and f). Additionally, tumors can have varied shape, size and characteristics making classification difficult. Figure 2 shows examples of the different tissue types that can be found in this dataset: (a) fatty breast with normal tissue (b) fatty breast with mass tissue (c) dense breast with normal tissue (d) dense breast with mass tissue (e) glandular breast with normal tissue (f) glandular breast with tumour tissue.

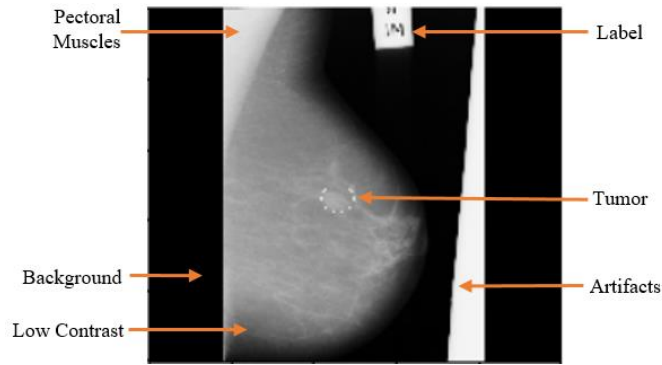


Figure 1: Problems in Mammogram Image in mini-MIAS dataset.

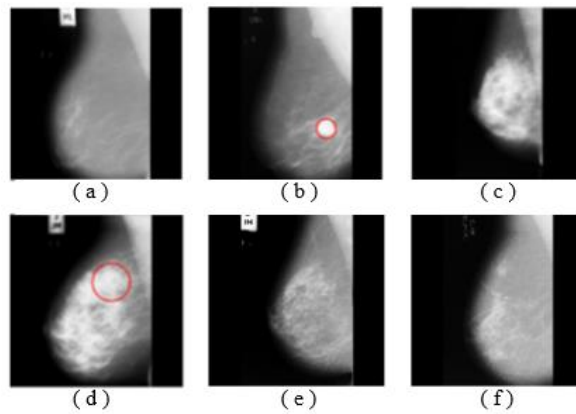


Figure 2: Sample images from Mini MIAS dataset showing different tissue types.

3 Methodology

As discussed in the introduction section, the aim of the paper is to detect communities based on similarities detected in mammography images of cancer patients. For this, we proposed to follow the steps below (see Figure 3):

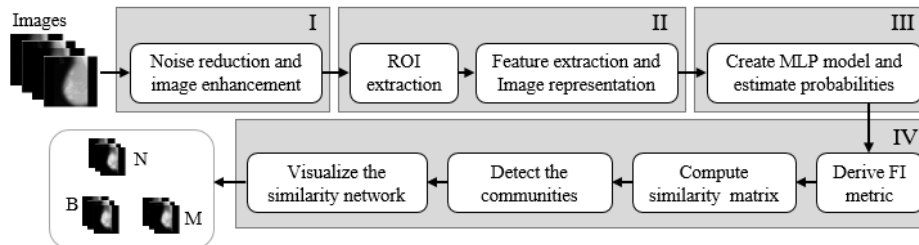


Figure 3: Pipeline followed for creating the Fisher Information Network

I) Image Pre-processing: This is the first step where noise in images is removed using 3X3 size median filter followed by the removal of image artefacts using image binarization and intensity thresholding. The breast profile is extracted by removing pectoral muscles using slant line approximation. Images are then morphologically enhanced using top hat filter (9). Structuring element used was “box” of “radius” 5.

II) Image Representation: Before representing images with features, ROI of size 50X50, containing normal, benign tumor or malignant tumor regions, is extracted from enhanced images. For images with tumor, the center of abnormalities defined in dataset was taken as the ROI center. For the normal samples, coordinates of ROI center were manually defined. ROI extraction reduced the training time significantly. It also helped to provide an initial representative area for benign, malignant, and normal instances for our model to learn differences from. The image representation involves two stages, feature extraction and feature selection, explained below:

- a. Feature Extraction: In this paper, texture features are extracted to represent images. Texture contains information about structural arrangement of surfaces. It specifies the roughness or coarseness of an object surface and can be described as a pattern with some element of regularity. Texture features have been proven useful in differentiating mass and normal breasts tissues, and according to (10) they can outperform intensity features. The reason for this can be that tumor area exhibits low texture compared to normal parenchyma. Texture features can be extracted by many methods, for example statistical methods such as gray level co-occurrence matrix (GLCM) in (11), model based methods such as Markov random field (MRF) model (12), and transformation based methods such as wavelet decomposition (13). However, most studies use GLCM to extract texture features. In this paper, statistical first order texture features, and statistical second order texture features from GLCM, together with gray level run length matrix (GLRLM) are extracted.
- b. Feature Selection: This step is performed to remove noisy, redundant and irrelevant features and retain the optimal set of effective and discriminating features. Specifically, features were ranked using Classification and Regression Tree (CART) (14). As interpretability was the focus of the project, CART was preferred over other methods. CART also takes into account discriminative power of the variable with respect to target variable.

III) Calculate probabilities: Classification for the three classes is performed using Multi-Layer Perceptron (MLP) with 10 fold cross-validation on 70% of the dataset used for training the algorithm and then testing on the remaining 30% of dataset to evaluate its performance. The hidden layer is set up with 3 hidden units and weight decay regularization of 0.1. Once the algorithm is tuned, probability densities of the three classes are calculated, which will then be used to calculate pairwise distances, producing the Fisher distance matrix.

IV) Creating the FIN: Next step is to detect the structure in the dataset by constructing the FIN. Details for this can be found below (also illustrated in Figure 3):

- a. Fisher Information metric (6): This metric is derived from the Fisher distance matrix, and defines a Riemannian space where distances reflect similarity with respect to a given probability distribution. This metric is informed about generative

properties of data, and hence can assess the importance of directions in space of parameters. FI combines multiple variables together to assess patterns and evaluate stability in system. It is obtained by differentiating the logarithm of the conditional probability $p(x|\theta)$ with respect to x and summing over all possible classifications:

$$FI(x) = E_{p(c|x)}\{(\nabla_x \log p(c|x))^T (\nabla_x \log p(c|x))\} = \quad (1)$$

$$-E_{p(c|x)}\{\nabla_x^2 \log p(c|x)\} .$$

Where $E_{p(c|x)}$ denotes the expectation over the density function $p(c|x)$ and ∇_x is the gradient with respect to x .

Given the FI metric, the infinitesimal distance between a pairs of neighboring points in the data space is given by the quadratic differential form:

$$d(x, x + \Delta x)^2 = \Delta x^T FI(x) \Delta x . \quad (2)$$

An important property of this metric is that it automatically scales each dimension of the data space according to its degree of relevance with respect to class membership, expanding directions along which $p(c|x)$ changes rapidly and compressing those where the variation is little. The result is a Riemannian space where the posterior class membership probability changes evenly in all directions.

- b. Similarity matrix and community detection: After calculating the distance matrix with a Gaussian radial kernel, the similarity matrix is then computed from them, resulting in the adjacency matrix which defines network structure. The communities are then detected by maximizing modularity using Newman's spectral optimization (15), resulting in clusters that best represent the graph structure. In medical database, patients do not interact as in social networks. Hence in medical networks, the presence of an edge indicates similarity between patients (observations), and the weights determine the strength of links. Central nodes in each community can also be found and used as representatives of the clusters giving a set of characteristic points to associate with each of the communities.
- c. Visualization of similarity network: This is done using classical Multi-Dimensional Scaling (MDS) (16). MDS uses a distance matrix to produce representation of points in lower dimensional Euclidean space such that the distance between them approximate as closely as possible the dissimilarities between corresponding instances in the original matrix. Applying MDS is also key as the mapping of the matrix onto the Euclidean space will allow many commonly used methods from signal processing to be applied. All this, while also retaining the distance structure generated by the FI matrix.

4 Results and Discussion

Preprocessing images with median filter removed the noise and morphological enhancement increased the contrast between bright and dark areas so that the classifier performance was optimized. The balanced (taking into account the number of cases per class) accuracy of the initial MLP model was 78.98% on the separated test set, with a standard deviation of 6.89. This was considered acceptable as compared to the

literature. Uppal and Naseem achieved an accuracy of 96.97% on MIAS dataset using fusion of cosine transform for classification into tumor and non-tumor(17) . MLP classifier in this paper distinguished fully between normal and tumor class. This initial MLP was created to generate the probability densities (i.e. probability of class membership) to generate the Fisher Information (FI) metric.

Figure 4 shows the three-dimensional representation of the FIN for the three classes involved, i.e. normal, malignant and benign. As expected, the normal class is well separated from both tumor classes, which are in turn also reasonably well-separated from each other, with some misclassifications.

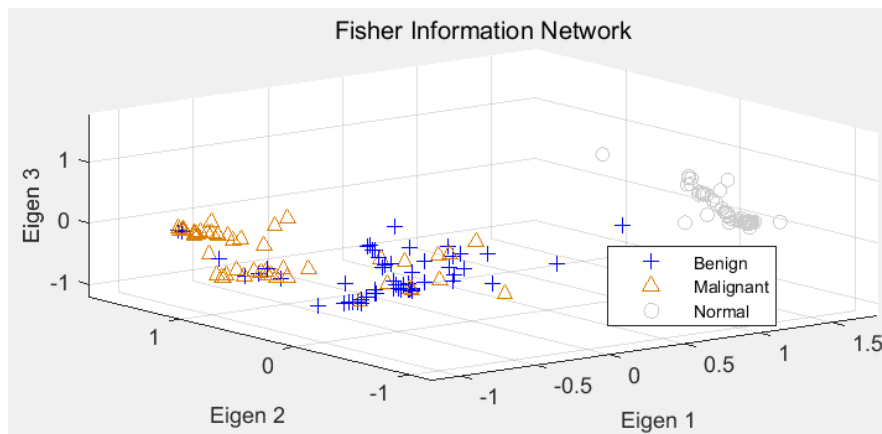


Figure 4: FIN Representation of Mini MIAS Dataset.

Figure 5 contains two- and three-dimensional representations of the dataset with communities identified by the network for the three classes. In this representation, edges are displayed only between members of same community to highlight cluster membership. The Fisher network is able to separate the data in three clearly distinguishable groups of communities, with normal cases being fully separated from the tumor cases.

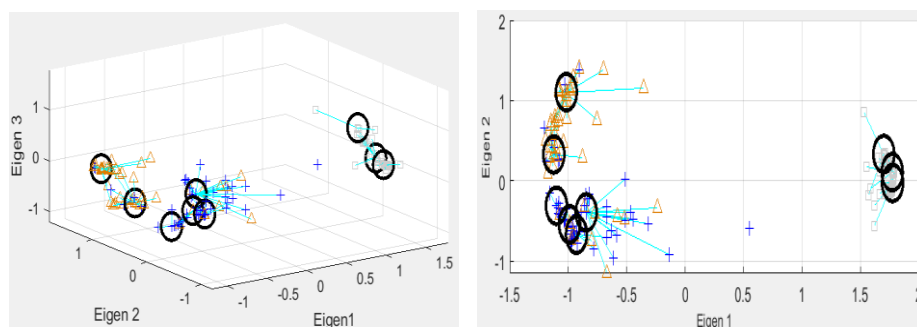


Figure 5: Three- and two-dimensional representations of the FIN with Communities identified by each network marked.

Figure 6 further analyses the purity of the communities detected by the FIN. It separates the normal cases entirely in three communities (1, 5 and 8) of 35, 23 and 4 subjects, respectively. These three communities not only do not contain observations from tumor samples, but also are represented very far from the rest of them representing the tumors. Communities 2, 4, 7 and 9 are mainly representing the benign cases, while the malignant tumors are mainly represented by communities 3 and 6. There is a benign case which is not assigned any of these nine communities and exists as a singleton in the center of the network (referred to as community 10 in the paper). See Table 2 for more details.

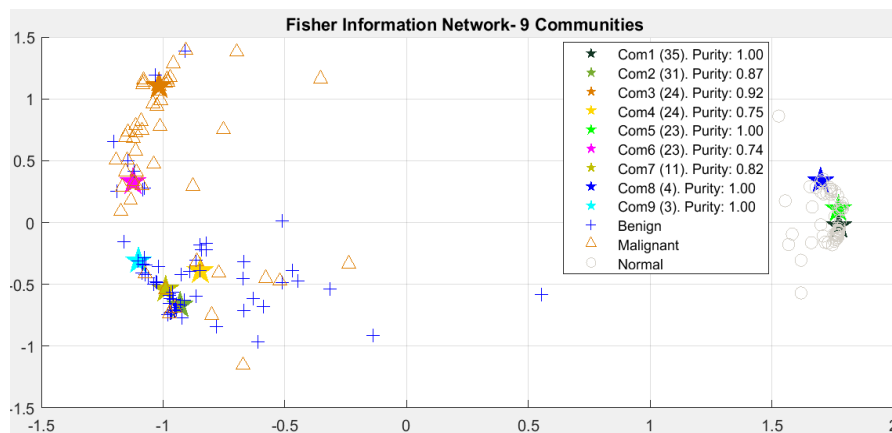


Figure 6: FIN Representation with size and purity for each community marked

Table 2 Number of cases per community of the main represented class.

Communities	Number of cases per community	Main class represented
1, 5 and 8	35, 23 and 4 normal cases, resp.	Normal: 62 (out of 62) – 100%
2, 4, 7, 9 and 10	27, 18, 9, 3 and 1 benign cases, resp.	Benign: 58 (out of 66) – 87.9%
3 and 6	22 and 17 malignant cases, resp.	Malignant: (39 out of 51) – 76.5%

The accuracies for the tumor classes were 76.5% and 87.9% for the malignant and benign classes, respectively. A total of 8 benign cases were falsely identified as malignant, and 12 malignant cases were falsely predicted as benign. None of the tumors were classified as normal, and none of the normal were classified as tumors, either benign or malignant, for a striking 100% accuracy in the separation of tumor from healthy tissue.

Another interesting aspect of the FIN representation in Figure 6 is the shape formed by all the scattered tumor cases projected, going from community 2 (mainly benign) to community 3 (mainly malignant). Possibly, all the area in between is showing a representation of the different gradations of the tumor. This is a very useful representation as it provides a level of confidence in the prediction, which is key when analyzing new observations. Furthermore, when projecting a new case in the map, it would be possible to look at neighboring cases and learn their characteristics, treatments, outcomes and prognosis, and with that information improve the diagnosis and prognosis of the analyzed patient.

The advantages of having used the FIN in this work can be summarized as follows:

- 1) The FIN automatically filters relevant variables based on their contribution with respect to the classification problem, measured by their influence on the posterior class probabilities. This allowed us to create a model that amplifies distances along the direction of the classes of interest, i.e. normal, benign and malignant.
- 2) It preserves the topology of the input space, producing affinity measures that reflect the data structure, which is later infused into the similarity network. It helps in understanding the hidden mechanisms that generated the data. This makes distances accurate measures of dissimilarity even when the number of covariates is large, as is the case in this study.
- 3) The framework can be used as an interpretable retrieval-based classifier, and, even more importantly, the results obtained are interpretable. This was key in this study as it allowed us not only to produce highly accurate classifiers, but also to learn for each case the level of confidence in the predicted outcome.

5 Conclusion

It can be seen that the FIN is able to distinguish between tumor and normal regions with 100% accuracy and separates benign and malignant tumor classes reasonably with some misclassifications. Nodes at the ends of the network are clear in terms of membership while membership of center nodes is relatively unclear and needs to be investigated further. Fisher information metric, used to construct the network, is informed about the generative properties of the data, and thus assesses the importance of directions in the space of the parameters. As the metric is Riemannian, expected variation of probability density caused by a distortion in parameters is different depending on the location of the space in which it is measured. Thus, Fisher information metric provides an elegant, clearly defined and statistically rigorous solution which is visualized as communities in a network, besides being classified into benign, malignant and normal classes.

This paper enhances image morphologically, with median filter for noise removal. Statistical second order texture features were extracted with GLCM and GLRLM along with first order texture features. Other methods for image enhancement, noise removal and extracting features can also be tested in the future.

References

1. World Health Organization. Breast cancer. WHO website. 2018.
2. Homer MJ. Mammographic interpretation : a practical approach. 2nd ed. New York: McGraw-Hill, Health Professions Division; 1997. 376 p.
3. Rangayyan RM, Ayres FJ, Leo Desautels JE. A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. *J Franklin Inst.* 2007 May 1;344(3–4):312–48.
4. Oelze ML. Quantitative ultrasound techniques and improvements to diagnostic

- ultrasonic imaging. In: IEEE International Ultrasonics Symposium, IUS. 2012.
5. Tadayyon H, Sadeghi-Naini A, Wirtzfeld L, Wright FC, Czarnota G. Quantitative ultrasound characterization of locally advanced breast cancer by estimation of its scatterer properties. *Med Phys*. 2014 Jan 2;41(1):012903.
 6. Ruiz H, Ortega-Martorell S, Jarman IH, Martín JD, Lisboa PJG. Constructing similarity networks using the Fisher information metric. In: European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN). Bruges, Belgium; 2012. p. 191–6.
 7. Ruiz H, Jarman IH, Martín JD, Lisboa PJG. The role of Fisher information in primary data space for neighbourhood mapping. In: European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN). Bruges, Belgium; 2011. p. 381–6.
 8. Suckling J, Parker J, Dance D. The mammographic image analysis society digital mammogram database. In: *Exerpta Medica International Congress Series*. 1994.
 9. Zhao D, Shridhar M, Daut DG. Morphology on detection of calcifications in mammograms. In: [Proceedings] ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing. IEEE; 1992. p. 129–32 vol.3.
 10. Yao J, Chen J, Chow C. Breast Tumor Analysis in Dynamic Contrast Enhanced MRI Using Texture Features and Wavelet Transform. *IEEE J Sel Top Signal Process*. 2009 Feb;3(1):94–100.
 11. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. *IEEE Trans Syst Man Cybern*. 1973 Nov;SMC-3(6):610–21.
 12. Cross GR, Jain AK. Markov random field texture models. *IEEE Trans Pattern Anal Mach Intell*. 1983 Jan;5(1):25–39.
 13. Laine A, Fan J. Texture classification by wavelet packet signatures. *IEEE Trans Pattern Anal Mach Intell*. 1993;15(11):1186–91.
 14. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees*. 1984.
 15. Newman MEJ. Detecting community structure in networks. In: *European Physical Journal B*. 2004. p. 321–330.
 16. Young G, Householder AS. Discussion of a set of points in terms of their mutual distances. *Psychometrika*. 1938 Mar;3(1):19–22.
 17. Uppal MTN. Classification of mammograms for breast cancer detection using fusion of discrete cosine transform and discrete wavelet transform features. *Biomed Res*. 27(2).