Applied GIS

ISSN: 1832-5505

Vol-11 Issue-03 July 2023

Analyzing Breast Cancer Images Histopathologically using Deep Learning

T SAI SRAVANI JYOTHIRMY/ V.SHANTHI

Assistant Professor[1,2]
DEPARTMENT OF ECE
Dr. K.V. Subba Reddy Institute of Technology

Abstract

Among all cancer types, breast cancer has the greatest morbidity rates and is thus a significant public health concern. If caught early, many life-threatening diseases are treatable, and the patient's prognosis improves dramatically. This, however, is a difficult and time-consuming procedure that calls for the expertise of pathologists. Automatic breast cancer identification using analysis of histological images has important clinical and prognostic implications for patients. However, conventional feature extraction techniques can only glean a few surface-level characteristics from pictures, and therefore need expert expertise to choose relevant characteristics. High-level abstract characteristics may be automatically extracted from photos using deep learning algorithms. As a result, we implement it to use supervised and unsupervised deep convolutional neural networks for analysing breast cancer histopathology pictures. As a first step, we used transfer learning to modify the Inception V3 and Inception ResNet V2 architectures for use with the binary and multi-class problems of breast cancer histopathology image classification. Subclasses were then rebalanced against Ductal Carcinoma as the baseline by flipping pictures top to bottom, left to right, and anticlockwise by 90 and 180 degrees to remove the bias introduced by the histological images' uneven distribution. When compared to previous approaches and our own experimental findings, the Inception V3 and Inception ResNet V2 based histopathological image categorization of breast cancer is clearly the best option currently available.

Keywords:

pictures of histopathology for breast cancer diagnosis; deep convolutional neural networks; autoencoder; transfer learning; classification; clustering.

INTRODUCTION

One of the most pressing problems in public health today is cancer. Cancer diagnoses rose by 28% between 2006 and 2016, and an estimated 2.7 million additional cancer cases would appear in 2030, according to data compiled by the International Agency for Research on Cancer (IARC) of the World Health Organization and the Global Burden of Disease Cancer Collaboration (Boyle and Levin, 2008; Moraga-Serrano, 2018). For women, breast cancer ranks high in both frequency and mortality (1.7 million new cases, 535,000 deaths, and 14.9 million DALYs) (Moraga-Serrano, 2018). As a result, detecting breast cancer early is crucial. Although X-ray, MRI (Magnetic Resonance Imaging), ultrasound, etc. have all been used to identify breast malignancies for over 40 years (Stenchiest et al., 1978), biopsy procedures are still the primary methods relied on to diagnose breast cancer accurately. Surgical biopsy, vacuum-assisted biopsy, and fine-needle aspiration are the most common biopsy methods. Samples of cells or tissues are collected, fixed, and stained before being examined under a microscope (Vita et al., 2014). Pathologists then make a diagnosis based on the histopathological pictures

(Spanhol et al., 2016a). Histopathological image analysis is a complex and time-consuming process that calls for expert understanding. Additionally, the amount of expertise of the pathologists participating may impact the results of the study. Therefore, the diagnosis and prognosis of breast cancer depend critically on computer-aided (Aswathy and Jagannath, 2017) interpretation of histological images. Nonetheless, the following obstacles slow the progress of building tools for doing this analysis. To begin, breast cancer histopathology photos are detailed, high-resolution pictures full of interesting shapes and patterns. When dealing with several classes, classification may be very challenging due of the variety within a class and the consistency across classes. The second difficulty is that current feature extraction techniques for breast cancer histopathology pictures have their limits. Existing techniques for feature extraction, such as the scale-invariant feature transform (SIFT) (Lowe, 1999) and the grey level co-occurrence matrix (GLCM) (Hara lick et al., 1973), depend on supervised data. However, the feature extraction effectiveness is quite poor, and the computational burden is very large, since it requires previous knowledge of the data to identify meaningful features. The final retrieved features are rather unimportant and low-level characteristics of histopathological pictures. As a consequence, the resulting model may be inadequate for its intended categorization task.

CONNECTED TEXTS

ISSN: 1832-5505

Vol-11 Issue-03 July 2023

Several significant advances in the field of image-based breast cancer detection have been made over the course of 40 years of study. This research may be broken down into two groups, one using conventional machine learning techniques, and the other using deep learning techniques. The former group relies on time-consuming and rather inefficient, abstract characteristics and is mostly used to limited datasets of breast cancer pictures. The latter group is capable of handling huge data and automatically extracting much more abstract characteristics from data. For the categorization of microscopic biopsy images, for instance, Zhang et al. (2013) introduced a novel cascade random subspace ensemble approach with rejection options in 2012. In this classification scheme, we use a pair of ensembles of random subspace classifiers. From the initial K-class classification issue (K = 3), we generate a collection of K binary classification problems, and the corresponding set of support vector machines makes up the first ensemble. Rejected samples from the first ensemble are the focus of the second ensemble, a Multi-Layer Perceptron ensemble. 361 photos were used to evaluate the system, with 119 representing normal tissue, 102 representing cancer in situ, and 140 representing lobular carcinoma or invasive ductal carcinoma. Twenty percent of photographs from each class were randomly selected for testing, while the other images were utilised for training. The overall accuracy of its classifications was 99.25%, its dependability was 97.50%, and its rejection rate was 1.94%. 500 pictures from 50 breast cancer patients were nucleus segmented by Kowal et al. (2013) using four different clustering techniques. Then, three distinct classification strategies were used to the pictures, separating benign from malignant tumours. There were 50 healthy cases and 50 cases with cancer represented by 10 photos each, for a total of 500 photographs.

Using 50-fold cross-validation, they were able to improve classification accuracy to between 96% and 100%. Using examination of cytological pictures of tiny needle biopsies, Flick et al. (2013) demonstrated a technique for diagnosing breast cancer. In order to create classifiers for the biopsies, 25 features of the nuclei were used, and four classic machine learning techniques were employed: KNN (K nearest neighbour with K = 5), NB (naive Bayes classifier with kernel density estimate), DT (decision tree), and SVM (support vector machine with Gaussian radial basis function kernel and scaling factor = 0.9). A total of 737 microscopic pictures of small needle biopsies taken from 67 individuals were used to evaluate these classifiers; this data set included 25 benign (275 images) and 42 malignant (462 images) cases. The highest recorded efficiency is 98.51 %. Using nuclear segmentation from cytological pictures, George et al. (2014) suggested a technique for diagnosing breast cancer. MLP (multilayer perceptron using the backpropagation technique), PNN, LVQ (learning vector quantization), and SVM (support vector machine) were the four models used for classification. George et al. provide a table (Table 5) detailing the model parameters (2014). With only 92 photos, 45 of benign tumours and 47 of malignant tumours are enough to achieve a classification accuracy of 76-94% using 10-fold cross validation. Asri et al. (2016) evaluated the efficiency of four machine learning algorithms, SVM, DT, NB, and KNN, on the 699-sample Wisconsin Breast Cancer dataset (including 458 benign and 241 malignant cases). The maximum level of accuracy (97.13%) was found to be reached using SVM with 10-fold cross-validation, as shown by the results of the experiments.

METHODS And Data Collections

Spanhol et al. (2016a) released the BreaKHis dataset utilised in this work. There are 7,909 histological pictures included, representing 82 breast cancer cases seen in the clinic. In order to see the database, go to http:// web.inf.ufpr.br/vri/breast-cancer-database. Each picture was obtained by a pathologist from a surgical sample of a patient's breast tissue to preserve the tissue's native structural and molecular makeup. Haematoxylin and eosin staining was then used to get the resulting pictures. In the end, pathologists used microscope observations to assign each picture its correct class designation. All breast cancer histopathology pictures are 700 460 three-channel RGB micrographs. These histological pictures of breast cancer were collected using objective lenses of varying multiples; hence the whole dataset was divided into four categories: 40X, 100X, 200X, and 400X.

These supplementary datasets are categorised by whether or not the tumours in question are benign or malignant. Consequently, there are four distinct types of tumours, two of which are benign and two of which are malignant. Adenosis (A), Fibroadenoma (F), Phyllodes Tumour (PT), and Tubular Adenoma (T) are all examples of benign tumours (TA). Ductal carcinoma, lobular carcinoma, mucinous carcinoma, and papillary carcinoma are all types of malignant tumours (PC). The Break His dataset's sample descriptions are included in Table 1. Each of the histopathological pictures of breast cancer must be turned into a 299 299 image to match the needed input size of the network structure, which is 299 299 for both the Inception V3 and Inception ResNet V2 networks utilised in this article.

Applied GIS

ISSN: 1832-5505

Vol-11 Issue-03 July 2023

TensorFlow's picture preparation tools, such as cropping the border box, resizing, and altering the saturation, were employed in the transformation process. By doing so, a three-channel picture with pixel values normalised to the range [1, 1] was produced that matched the model's input size. The datasets for the four magnification factors were randomly divided into training and testing subsets in the ratio of 7:3 to guarantee generalizability of the experimental findings in the classification task.

Classification System

In this part, we'll talk about our experience utilising the Inception V3 (Szeged et al., 2016) and Inception ResNet V2 (Szeged et al., 2017) deep learning models to categorise histopathological pictures of breast cancer, as well as our analysis of those findings.

Connectedness and Classification Networks

Our studies make use of the Inception V3 (Szeged et al., 2016) and Inception ResNet V2 (Szeged et al., 2017) networks, both of which were suggested by the same authors in 2016 and 2017, respectively. In the ILSVRC competition, it was shown that the Inception ResNet V2 network, when trained on large datasets, could outperform the Inception V3 network. The presence of residual connections in the Inception ResNet V2 network distinguishes it from the Inception V3 network. In this study, we use these two networks to classify histopathological pictures of breast cancer to see whether the experimental findings from Inception ResNet V2 are better than those from Inception V3 on small datasets. Figure 1 depicts these network diagrams. Figure 1 demonstrates the structural similarity between the two networks. The first few layers implement a signature transformation using classic convolutional and pooling layers, while the core is made out of a stack of Inception modules. At last, the SoftMax function is used to output the results via the fully-connected layer. The Inception modules of the Inception V3 and Inception ResNet V2 networks are fundamentally different from one another. Each Inception module of the Inception V3 network is made up of filters of varying sizes, such as 1 1, 1 3, and 3 1, to improve the network's flexibility to various convolution kernels. By adding a residual unit to each Inception module, the Inception ResNet V2 network is able to prevent the degradation of the network gradient that is often seen as layer counts grow. In addition to a variety of filter.

TABLE 1 | Image distribution of different subclasses in different magnification factors

Magnification	Benign				Malignant				Total
	A	F	PT	TA	DC	LC	MC	PC	
40X	114	253	109	149	864	156	205	145	1,995
100X	113	260	121	150	903	170	222	142	2,081
200X	111	264	108	140	896	163	196	135	2,013
400X	106	237	115	130	788	137	169	138	1,820
Total	444	1,014	453	569	3,451	626	792	560	7,909
#Patients	4	10	3	7	38	5	9	6	82

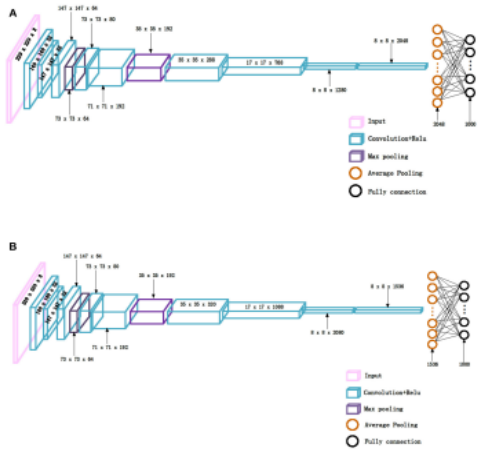


FIGURE 1 The network structures, (A) Inception_V3, (B) Inception_ResNet_V2

At larger network sizes, layer-jumping is possible thanks to residual connections, mitigating the decline brought on by more nodes. Comparing the 8x8 Inception modules of Inception V3 with Inception ResNet V2, as seen in Figure 2. For more information, please see the cited works (Szeged et al., 2016, 2017).

Skill Transfer

An important application of deep learning is transfer learning (Pan and Yang, 2010). It's common knowledge that you can't train a sophisticated deep network from start with a little dataset. In addition, there are no predetermined guidelines for developing a network architecture to accomplish a certain goal. Instead of recreating the wheel, we may utilise the information gathered as pre-training for our unique research aim by adopting the model and the parameters achieved by other researchers via time-consuming and computationally costly training on the extremely large picture dataset of ImageNet. After that, we may retrain the model's final specified fully-connected layer using a modest quantity of data and still receive satisfactory results on our goal job.

ISSN: 1832-5505

Vol-11 Issue-03 July 2023

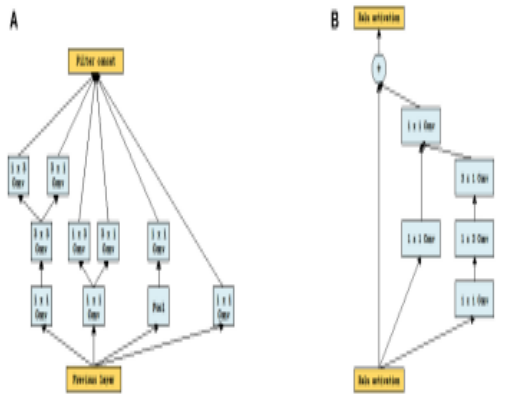


FIGURE 2 | The inception module of size 8 × 8 in two networks, (A) Inception_V3, (B) Inception_ResNet_V2

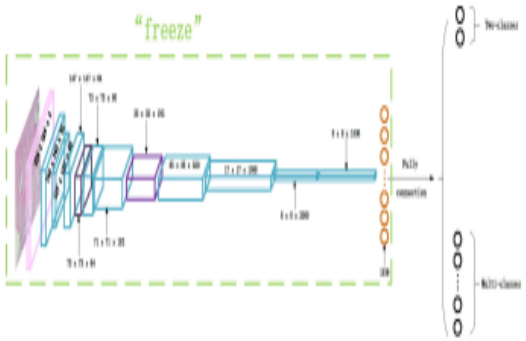


FIGURE 3 | The Inception_ResNet_V2 network structure for transfer learning.

Transfer learning is adopted in this paper to classify the histopathological images of breast cancer using Inception_V3 and Inception_ResNet_V2 networks. We first downloaded the models and parameters of Inception_V3 and Inception_ResNet_V2 networks trained on the ImageNet dataset. The dataset is composed of about 1.2 million training images, 50,000 validation images, and 100,000 testing images. This comprises a total of 1,000 different categories. Then, we froze all of the parameters before the last layer of the networks. We modified the number of neurons of the last fully-connected layer as 2 for binary classification and 8 for multi-class classification. After that, the parameters of the fully-connected layer are trained on the histopathological images of breast cancer. The modified network structure of the Inception_ResNet_V2 network is shown in Figure 3. The modified Inception_V3 network structure is similar, so it is omitted. Our classification process was developed based on the TensorFlow deep learning framework. The Adam (adaptive moment estimation) (Kingman and Ba, 2014) algorithm was used in the training process to perform optimization by iterating through 70 epochs using the histopathological image dataset of breast cancer. The batch size is set to 32 in the experiments, and the initial learning rate is 0.0002 (Bergstrom and Bagnio, 2012). Then, the exponential decay method is adopted to reduce the learning rate and ensure that the model moves through iterations quickly at the initial training stage. This also helps to provide more stability at the later stage and makes it easier to obtain the optimal solution. The decay coefficient is set as 0.7 (Bergstrom and Bagnio, 2012), and the decay speed is set so that the decay occurs every two epochs. The specific decay process is shown in (1), where decayed_learning_rate is the current learning rate, learning rate is the initial learning rate, decahydrate is the decay coefficient, global step is the current iteration step, and decay steps is the decay speed.

$$decayed\_learning\_rate = learning\_rate \times decay\_rate^{(global\_step/decay\_steps)} \quad (1)$$

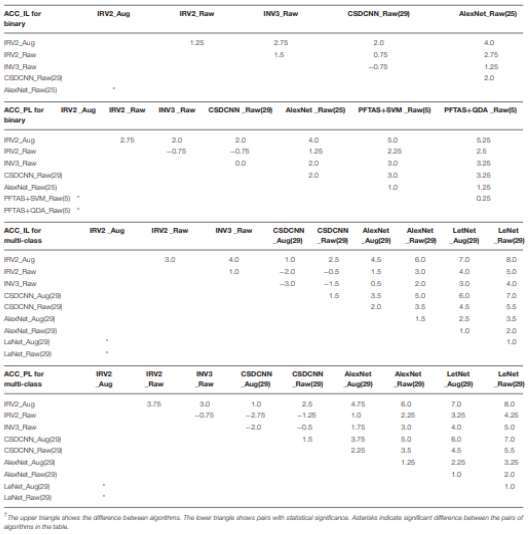
Climatic Grouping Outcomes

In this part, we'll talk about how the Inception ResNet V2 network can automatically extract important features from breast cancer histopathology pictures, which is a huge time-saver. Histopathological pictures of breast cancer are processed using Inception ResNet V2 to extract features along 1,536 dimensions, and then clustered using the K-means technique. Moreover, a new AE (Autoencoder) network with the dimensions [1536, 500, 2] is built to apply a non-linear modification to the 1,536-dimensional feature vectors produced by Inception ResNet V2. This allows us to get, in low dimensional space, the 2-dimensional characteristics of histological pictures of breast cancer, which may then be used by K-means. Clustering results from K-means using features extracted by Inception ResNet V2 are represented by IRV2+Kmeans, while clustering results from K-means using features altered by our proposed AE utilising features retrieved by Inception ResNet V2 are represented by IRV2+AE+Kmeans.

TABLE 3 Paired rank comparison of algorithms in ACC_IL and AII_PL for binary and multi-class classification

ISSN: 1832-5505

Vol-11 Issue-03 July 2023



ACC_IL for binary	IRV2_Aug	IRV2_Raw	INV3_Raw	CSDCNN_Raw(29)	AlexNet_Raw(25)
IRV2_Aug		1.25	2.75	2.0	4.0
IRV2_Raw			1.5	0.75	2.75
INV3_Raw				−0.75	1.25
CSDCNN_Raw(29)					2.0
AlexNet_Raw(25)	*				

ACC_PL for binary	IRV2 _Aug	IRV2 _Raw	INV3 _Raw	CSDCNN _Raw(29)	AlexNet _Raw(25)	PFTAS+SVM _Raw(5)	PFTAS+QDA _Raw(5)
IRV2_Aug		2.75	2.0	2.0	4.0	5.0	5.25
IRV2_Raw			−0.75	−0.75	1.25	2.25	2.5
INV3_Raw				0.0	2.0	3.0	3.25
CSDCNN_Raw(29)					2.0	3.0	3.25
AlexNet_Raw(25)						1.0	1.25
PFTAS+SVM_Raw(5)	*						0.25
PFTAS+QDA_Raw(5)	*						

ACC_IL for multi-class	IRV2 _Aug	IRV2 _Raw	INV3 _Raw	CSDCNN _Aug(29)	CSDCNN _Raw(29)	AlexNet _Aug(29)	AlexNet _Raw(29)	LetNet _Aug(29)	LeNet _Raw(29)
IRV2_Aug		3.0	4.0	1.0	2.5	4.5	6.0	7.0	8.0
IRV2_Raw			1.0	−2.0	−0.5	1.5	3.0	4.0	5.0
INV3_Raw				−3.0	−1.5	0.5	2.0	3.0	4.0
CSDCNN_Aug(29)					1.5	3.5	5.0	6.0	7.0
CSDCNN_Raw(29)						2.0	3.5	4.5	5.5
AlexNet_Aug(29)							1.5	2.5	3.5
AlexNet_Raw(29)								1.0	2.0
LeNet_Aug(29)	*								1.0
LeNet_Raw(29)	*								

ACC_PL for multi-class	IRV2 _Aug	IRV2 _Raw	INV3 _Raw	CSDCNN _Aug(29)	CSDCNN _Raw(29)	AlexNet _Aug(29)	AlexNet _Raw(29)	LetNet _Aug(29)	LeNet _Raw(29)
IRV2_Aug		3.75	3.0	1.0	2.5	4.75	6.0	7.0	8.0
IRV2_Raw			−0.75	−2.75	−1.25	1.0	2.25	3.25	4.25
INV3_Raw				−2.0	−0.5	1.75	3.0	4.0	5.0
CSDCNN_Aug(29)					1.5	3.75	5.0	6.0	7.0
CSDCNN_Raw(29)						2.25	3.5	4.5	5.5
AlexNet_Aug(29)							1.25	2.25	3.25
AlexNet_Raw(29)								1.0	2.0
LeNet_Aug(29)	*								1.0
LeNet_Raw(29)	*								

*The upper triangle shows the difference between algorithms. The lower triangle shows pairs with statistical significance. Asterisks indicate significant difference between the pairs of algorithms in the table.

degree of separation and condensation; applicable even when no label data is available. SSE has a range of [1, 1]. More tightly clustered samples and more widely spaced samples from various groups are indicated by larger SSE values. Closer SSE values near 1 suggests tighter grouping.

Assessment of Outcomes

In this part, we will evaluate the clustering outcomes of IRV2+AE+Kmeans and IRV2+Kmeans in terms of external criteria, such as ACC, ARI, AMI, and the internal measure SSE. Clustering findings according to the aforementioned four criteria are shown for datasets of varying magnifications in Figure 7. Figure 7 displays experimental data that demonstrates the following. (1) On all datasets and scaling factors, IRV2+AE+Kmeans produces superior clustering results than IRV2+Kmeans in terms of ARI, AMI, SSE, and ACC. By encoding the features retrieved by the Inception ResNet V2 network, our proposed AE network is able to provide considerably more abstract and expressive features. Without applying any alteration to the features generated by Inception ResNet V2, (2) the values of ARI, AMI, SSE, and ACC for the same clustering are increasing. Thirdly, using features generated by the Inception ResNet V2 network, the highest clustering accuracy (ACC) is 59.3% on the 40X dataset, while utilising features converted by the proposed AE network using extracted features from the Inception ResNet V2 network, the best ACC is 76.4% on the 200X dataset. In conclusion, IRV2+AE+Kmeans has a best ACC of 76.4%, whereas IRV2+Kmeans has a top ACC of 59.3%.

Conclusion

In this study, we presented our approaches for analysing breast cancer histopathology pictures using deep convolutional neural networks such as Inception V3 and Inception ResNet V2 that have been trained using transfer learning techniques. These two networks have already been trained using ImageNet, a massive picture database. Then the settings and structure they've learnt are permanently set. The fully-connected layer's parameters are retrained, and its number of neurons is adjusted such that it is optimal for our job. This allows the model to be used to breast cancer histopathology pictures for either binary or multi-class categorization. By comparing our experimental findings to those of previous research, we show that the Inception ResNet V2 network is better than the Inception V3 network when it comes to analysing histopathology pictures of breast cancer.

We also find that our experimental findings improve greatly on the original datasets when we use the enriched datasets. This is particularly true when using our histopathology pictures of breast cancer for multi-class categorization. By comparing the experimental findings, we find that the Inception ResNet V2 network is able to extract much more informative features than the other networks we used as references. K-means, a common clustering technique, was used to analyse histopathology pictures of breast cancer, and the results showed that the optimal K value for K-means could be determined by using the internal criteria of SSE. For breast cancer histopathology pictures, the suggested AE network may identify far more informative, low dimensional characteristics.

REFERENCES

[1]Araújo, T., Arista, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., et al. (2017). Classification of breast cancer histology images using convolutional neural networks. PLoS ONE 12:e0177544. doi: 10.1371/journal.pone.0177544

[2] Asri, H., Mousannif, H., Al Moatassime, H., and Noel, T. (2016). Using machine learning algorithms for breast cancer risk prediction and diagnosis. Procedia Comput. Sci. 83, 1064–1069. doi: 10.1016/j.procs.2016.04.224

[3] Aswathy, M., and Jagannath, M. (2017). Detection of breast cancer on digital histopathology images: present status and future possibilities. Inform. Med. Unlocked 8, 74–79. doi: 10.1016/j.imu.2016.11.001

[4] Bayramoglu, N., Kannala, J., and Heikkilä, J. (eds) (2016). "Deep learning for magnification independent breast cancer histopathology image classification," in 23rd International Conference on Pattern Recognition (ICPR), 2016. (Cancun: IEEE).

[5] Bengio, Y., Courville, A., and Vincent, P. (2013). "Representation learning: A review and new perspectives," in IEEE Transactions on Pattern Analysis and Machine Intelligence 35, 1798–1828. doi: 10.1109/TPAMI.2013.50

[6] Bergstra, J., and Bengio, Y. (2012). Random search for hyper-parameter optimization. J. Machine Learn. Res. 13, 281–305.

ISSN: 1832-5505

Vol-11 Issue-03 July 2023

[7] Borg, A., Lavesson, N., and Boeva, V. (eds) (2013). Comparison of Clustering Approaches for Gene Expression Data. Aalborg: SCAI.

[8] Boyle, P., and Levin, B. (2008). World Cancer report 2008: IARC Press. International Agency for Research on Cancer.

[9] Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recog. 30, 1145–1159.

[10] doi: 10.1016/S0031-3203(96)00142-2 Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of p-values. R. Soc. Open Sci. 1:140216. doi: 10.1098/rsos.140216

[11] Ellis, P. D. (2010). The Essential Guide to Effect Sizes: Statistical Power, MetaAnalysis, and the Interpretation of Research Results. New York, NY: Cambridge University Press.

[12] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115–118. doi: 10.1038/nature21056

[13] Filipczuk, P., Fevens, T., Krzyzak, A., and Monczak, R. (2013). Computeraided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies. IEEE Trans. Med. Imaging 32, 2169–2178. doi: 10.1109/TMI.2013.2275151

[14] George, Y. M., Zayed, H. H., Roushdy, M. I., and Elbagoury, B. M. (2014). Remote computer-aided breast cancer detection and diagnosis system based on cytological images. IEEE Syst. J. 8, 949–964. doi: 10.1109/JSYST.2013. 2279415

[15] Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316, 2402–2410. doi: 10.1001/jama.2016.17216.