



HscoreNet: A Deep network for estrogen and progesterone scoring using breast IHC images

Monjoy Saha^{a,b,*}, Indu Arun^c, Rosina Ahmed^c, Sanjoy Chatterjee^c, Chandan Chakraborty^a

^aSchool of Medical Science and Technology, Indian Institute of Technology Kharagpur, West Bengal, India

^bDepartment of Biomedical Informatics, Emory University-School of Medicine, Atlanta, Georgia, United States

^cTata Medical Center, Kolkata, West Bengal, India

ARTICLE INFO

Article history:

Received 17 February 2019

Revised 19 November 2019

Accepted 9 January 2020

Available online 10 January 2020

Keywords:

Breast

Estrogen

Progesterone

Encoder

Decoder

ABSTRACT

Estrogen and progesterone receptors serve as an important predictive and prognostic biomarkers for breast cancer immunohistological analysis. For breast cancer prognosis, pathologists manually compute the score based on the visual expression and the number of immunopositive and immunonegative nuclei. This manual scoring technique is time-consuming, cumbersome, expensive, error-prone, and susceptible to intra- and interobserver ambiguities. To solve these issues, we proposed a deep neural network (i.e., *HscoreNet*), which consists of three parts, i.e., encoder, decoder, and scoring layer. A total of 600 (300 ER and 300 PR) regions of interest at $40\times$ magnification from 100 histologically confirmed slides were used in this study. The size of each region of interest was 2048×1536 pixels (width \times height). The encoder layer has been used to transform input pixels into a lower-dimensional representation, whereas the decoder reconstructs the output of the encoder through minimization of a cost function. The decoder generates an image that only contains immunopositive and immunonegative nuclei. The output of the decoder is fed to the input of the scoring layer. This layer computes the Histo-score or H-score based on the staining intensity, the color expression, and the number of immunopositive and immunonegative nuclei. Pathologists compute this score to subcategorize the cancer grades and to decide proper treatment procedures. Our proposed approach is affordable, accurate, and fast. We achieved excellent performance, with 95.87% precision and 94.53% classification accuracy. Our proposed approach streamlines the human error-prone and time-consuming process. This methodology can also be used for other types of histology and immunohistology image segmentation and scoring.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Breast cancer (BC) is a devastating carcinoma that mostly affects females (age groups between 40 to 70 years) rather than men worldwide [1,2]. This cancer accounts for almost 30% of all new carcinomas among females [3]. In 2017, approximately 255,180 new cases of BC (252,710 female and 2,470 male) and 41,070 deaths (40,610 female and 460 male) were reported in the United States [4]. It is assumed that in 2019 in the United States, nearly 271,270 new cases (268,600 female and 2,670 male) and 42,260 deaths (41,760 female and 500 male) will occur [3]. In the year 2018, the Global Cancer Statistics (GLOBOCAN) reported that in the European Union, BC was in the third position among all leading diseases [5]. As per the global BC statistics, in 2012, almost 522,000 deaths and

1,671,000 new cases were registered [6]. From the past and present cancer statistics, we can say that BC is alarming and needs to be controlled properly at an early stage. For breast cancer prognosis, histological tissue analysis is considered the gold standard procedure. In this regard, a pathologist first collects tissues from the suspected regions (mostly from the tumor regions) of the breast and then processes those tissues by fixing, cutting, and staining. They mostly use hematoxylin and eosin (H&E) and immunohistological (IHC) stains to stain the tissues [7,8]. H&E stain is used to categorize BC based on grades (e.g., grade I, grade II, grade III, and grade IV), whereas IHC analysis is performed to confirm the severity of cancer based on the expression of different hormonal proteins. The main hormonal proteins are estrogen (ER), progesterone (PR), Ki-67 and HER-2 [9–12].

As per the report, 70% of BC cases express ER and PR expression. Fig. 1 shows ER and PR stained images at $40\times$ magnification. The cancerous cells that exhibit ER-positive expression have a better prognosis than the ER-negative cells [13]. The ER-positive status

* Corresponding author at: Department of Biomedical Informatics, Emory University-School of Medicine, Atlanta, Georgia, United States
E-mail address: monjoy.saha@emory.edu (M. Saha).

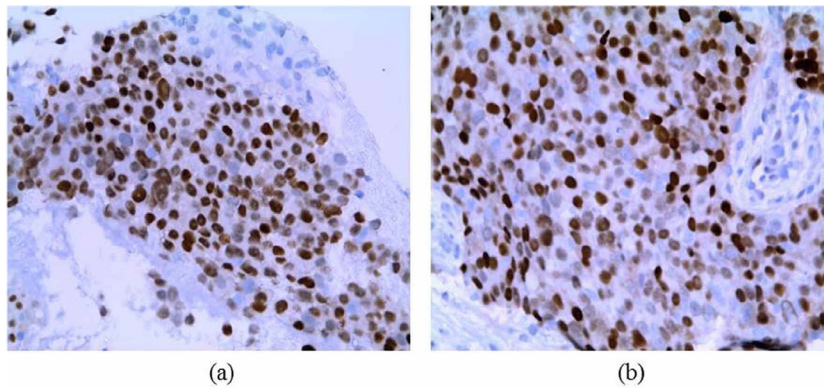


Fig. 1. Breast cancer ER and PR stained images at $40 \times$ magnification: (a) estrogen receptor; (b) progesterone receptor. The brown colored nuclei are called immunopositive, and the blue colored nuclei are called immunonegative.

signifies the necessity of postsurgical hormonal therapy to the patient. PR expression is observed due to ER-positive breast epithelium [14]. The coexpression of PR and ER in BC is very significant. Hence, PR is called a postreceptor biomarker of ER. The PR and ER scores help the clinicians decide the treatment procedure and appropriate drugs for hormonal therapy [15]. ER-positive patients have a survival rate seven to eight times greater than ER-negative patients [16].

In most cases, pathologists determine the ER and PR scores manually. This process has a chance of getting erroneous results. Moreover, the manual technique is time-consuming, costly, and prone to inter- and intra-observer variabilities. Furthermore, in the remote areas where advanced medical instruments and experienced pathologists are unavailable, manual tissue analysis creates an unnecessary delay in starting the treatment [17]. Hence, there is an urgent need to develop an automated technique to compute the ER and PR scores.

Many researchers around the globe have proposed different machine learning approaches for H&E stained tissues analysis [18,19], but very few works in which IHC images were used have been reported. The complexity of IHC images is much higher than that of H&E stained images. Our group already proposed different deep learning-based approaches for Ki-67 and HER-2 stained tissue analysis [7,20]. Hence, we are interested in the development of a deep network for ER and PR stained images.

In this manuscript, we propose a deep neural network composed of an encoder, a decoder, and a scoring layer for ER and PR scoring using IHC stained images. Our proposed deep network utilizes the most recent advances in deep layer design and training techniques. This manuscript is organized as follows: introduction, related works, materials and methods, results and discussion, and conclusion and future work.

2. Related works

The use of machine learning approaches for pathology data analysis is growing day by day. One of the main reasons for using machine learning is its capability of solving any complex task very efficiently. Moreover, machine learning approaches are fast, reliable, and affordable. There are many articles available in which machine learning algorithms were proposed for object (e.g., nuclei, mitosis, etc.) segmentation, detection, classification, regression, and many more purposes. The most popular image analysis algorithm includes watershed [18], Otsu thresholding [21], regression [22], and gradient energy tensor [23]. In addition to conventional techniques, various deep learning (DL) models have been proposed by various researchers. A few of them are sparse auto-encoder [24] and U-net [25]. We have carried out an extensive lit-

erature survey on the application of machine learning in pathology, especially H&E and IHC, and radiology image analysis. We observed that there is no reliable tool available for automated scoring of ER and PR using IHC images. Here, we list some of the research articles that are relevant to our objectives.

A Fuzzy C-means based approach for nuclei segmentation and classification using ER-stained IHC images was proposed in [26]. The authors reported 93.1% accuracy, 93.2% specificity, and 95.7% sensitivity. One of the very popular tools in IHC image analysis, Immunoratio, was proposed for ER and PR scoring [27]. Immunoratio is a web-based tool that uses ImageJ in the backend. The spatial dimension of ER and PR signaling using the XD2 software package has also been used [28]. The authors of [29] utilized Adobe Photoshop for ER status evaluation at $40 \times$ image magnification. They used red-scale absorption characteristics for calculating the staining intensity. To find the ER and PR level, researchers have proposed random forest classifiers in some articles [30]. In [31], ER status was predicted using dynamic breast contrast-enhanced magnetic resonance images. They used Otsu thresholding, morphological dilation, and a connected component algorithm.

The authors of [32] reported color normalization, morphological operations, and a thresholding-based algorithm for nuclei detection. They reported 97.6% accuracy. In [33], the authors worked on H&E stained ER-positive images. They extracted Haralick texture features from the region of interests and then classified based on the grades. The authors of [34] used thresholding, concave point detection, and watershed algorithms for nucleus segmentation using immuno-fluorescence images. They reported 97.47% segmentation accuracy (SA).

Deep convolutional networks also have been used for image (medical or natural) segmentation. Some of the popular deep segmentation architectures are U-Net [25], SegNet [35], and Bayesian SegNet [36]. The U-Net architecture is being widely used for brain tumor segmentation using magnetic resonance images [37,38], sparse-view CT image reconstruction [39], pancreas segmentation using CT images [40], ventricular segmentation [41], breast mass segmentation using digital mammogram [42], and Retinal Vessel Segmentation [43], for example. Deep learning models are also implemented for nuclei detection and classification [44–46]. In [47], the authors proposed sparse auto-encoder based unsupervised technique for nucleus detection using H&E-stained digital pathology images. Encoder-decoder based auto-encoders were also reported for image clustering and binarization [48,49]. The authors of [50] proposed a convolutional network for cell classification using ER-stained whole-slide images. They tested in the region of 1200 cells and reported 88.8% overall accuracy. Table 1 summarizes different methodologies proposed for ER and PR image analysis.

Table 1

Summarization of ER and PR scoring approaches as per the existing state-of-art.

Categories	Year	Cancer type	Hormone name	Methodology used	Remarks	References
Conventional techniques	2016	Breast cancer	ER	Fuzzy-C-means	They achieved 93.1% accuracy, 95.7% sensitivity and 93.2% specificity	[26]
	2016		ER, HER-2 and TN	Region growing, logistic regression	They achieved 73.53 %, accuracy for ER	[51]
	2014		ER	Color deconvolution, thresholding, morphological operations	97.6% scoring accuracy reported	[32]
	2013		ER	Semi-quantitative	ER positivity is associated with reduced risk of BC	[52]
	2013		ER, PR, Ki-67 and CK5	Semi-quantitative	ER positivity correlated with PR positivity	[53]
	2012		ER	Semi-quantitative	57.9% Correlation coefficient achieved	[54]
	2008		ER and PR	Color space conversion, contrast adjustment, mean shift and watershed segmentation, random forest	They achieved a potential outcome	[30]
Deep Learning		Till date attempt has not been made for automated scoring of ER and PR using deep learning				

The major contributions of our manuscript are as follows:

- The originality of *HscoreNet* lies in the manner in which the encoder and decoder maintain the balance of image resolution during downsampling and upsampling.
- An additional scoring layer for H-score calculation is implemented.
- It is a valuable addition in terms of main quantification in the already existing established techniques for ER and PR scoring.

3. Materials and methods

3.1. Materials

3.1.1. Slide preparation and image acquisition

To carry out this research, ethical approval has been granted by the Tata Medical Center (TMC), Kolkata (Ref. no. EC/GOVT/07/14; dated August 11, 2014), and the Indian Institute of Technology, Kharagpur (Ref. no. IIT/SRIC/SAO/2015; dated July 23, 2015). The patient and their close relatives had signed the patient consent form. The IHC slides were prepared and maintained by the experts. IHC staining of ER and PR tissue sections was done as per the institutional guidelines and regulations. Paraffin-embedded tissue sections were fixed using formalin. The slides were stained using the respective monoclonal antibodies. A total of 600 (300 ER and 300 PR) regions of interest at $40 \times$ magnification from 100 histologically confirmed slides were obtained using a Zeiss whole slide imager (Zeiss Axio Imager M2 microscope with Axiocam ICc5camera) at a constant brightness and contrast. The size of each region of interest was 2048×1536 pixels (width \times height). All the images contained almost 359,884 (231,851 immunopositive and 128,033 immunonegative) nuclei.

The texture, morphology, and color distribution profiles of ER and PR-positive cases were visually similar, and vice-versa. Therefore, in our proposed *HscoreNet*, all immunopositive nuclei have been considered one category, and similarly, all immunonegative nuclei have been considered a different category. Due to the significant variation in the dataset, the proposed *HscoreNet* model can accept any brown colored nuclei as immunopositive and blue colored nuclei as immunonegative irrespective of different antigens.

3.2. Methods

3.2.1. Proposed HscoreNet model and learning schemes

Our proposed DL based framework has been developed predominantly inspired by SegNet [35] and U-Net [25]. The work

flow diagram of the proposed *HscoreNet* is shown in Fig. 2. The *HscoreNet* architecture mainly consists of the encoder, decoder and scoring layer. The encoder transforms input pixels into a lower-dimensional representation, whereas the decoder is used to reconstruct the output of the encoder through the minimization of cost function [55,56]. The output of the decoder is the segmented image, which contains only immunopositive and immunonegative nuclei. The scoring layer is used to classify and score the nuclei based on the staining intensity, color expression and the number of immunopositive and immunonegative nuclei. The *HscoreNet* is trained in a supervised fashion, which enables extracting useful features from the unannotated dataset. The overall training and testing diagram of *HscoreNet* is shown in Fig. 3. The proposed ER and PR scoring architecture is presented in Table 2 [here, MP denotes the max-pooling operation].

The encoder-decoder part of *HscoreNet* directly enumerates the conditional probability when the input sequence is applied. Mathematically, the decoder can be represented as $P(d_1, d_2, \dots, d_O | x_1, x_2, \dots, x_T)$ where d_0 , and x_T represent the output and hidden layers, respectively. The input is denoted by T , whereas O represents the output length. For *HscoreNet*, the input is usually a sequence of input image patches, while the output is a sequence of images that only contain immunopositive and immunonegative nuclei. The decoder output can be written as [57]

$$P(d_1, d_2, \dots, d_O | x_1, x_2, \dots, x_T) = \prod_{o=1}^O P(d_o | d_1, d_2, \dots, d_{o-1}, c_0) \quad (1)$$

where c_0 represents a context vector and represented as $c_0 = \sum_s \beta_s x_T$.

Here, $\beta_s \in [0,1]$ and $\sum_s \beta_s = 1$; $x_T = (\vec{x}_T, \overleftarrow{x}_T)$.

The weight β_s is calculated as below:

$$\beta_s = \frac{\exp(e_s)}{\sum_{s'} \exp(e_{s'})} \quad (2)$$

where $e_s = a(T_{O-1}, g_s)$ and a feed-forward neural network is assigned to $a(\cdot)$.

The posterior probability of d_o is computed as

$$P(d_o | d_1, d_2, \dots, d_{o-1}, c_0) = g(d_{o-1}, f(d_{o-1}, c_0), c_0) \quad (3)$$

Here, $f(d_{o-1}, c_0)$ represents the recurrent hidden layer, and $g(d_{o-1}, f(d_{o-1}, c_0), c_0)$ denotes the softmax function.

The encoder layer consists of three-convolution-layer block-1, two-convolution-layer block-2, and four max-pooling layers. The convolution layer block-1 consists of two repeated convolution layers, weight normalization, ReLU, and a dropout layer. The convolution layer block-2 consists of three repeated convolution layers,

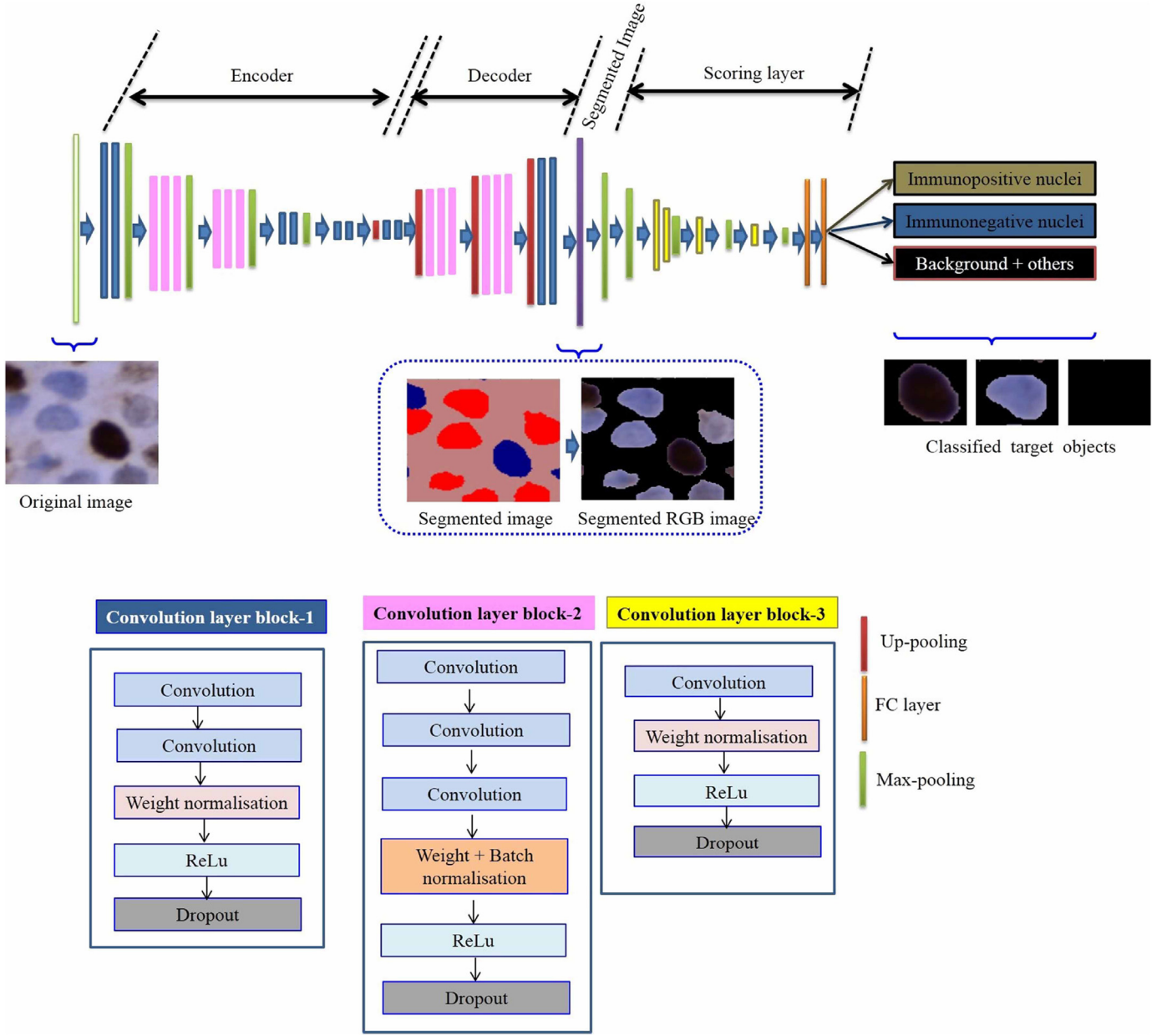


Fig. 2. The work flow diagram of the proposed HscoreNet.

weight normalization, batch normalization, ReLu, and a dropout layer. The decoder part includes two-convolution-layer block-1 and two-convolution-layer block-2.

The decoder layer helps in up-sampling the input feature maps using the memorized max-pooling indices and produces sparse feature maps. Each layer is described below in detail.

3.2.2. Convolutional layer

The convolution layer is used to extract low- and high-level hierarchical features from input image pixels. In the proposed model, multiple layers have been stacked in a manner such that the $(n-1)^{th}$ layer's output is the input of the $(n)^{th}$ layer. A convolution layer is used to learn convolutional filter and calculate feature maps. Mathematically, feature maps can be represented as below [58]:

$$F_m^n = f \left\{ \sum_i (F_i^{n-1} \times G_{im}^n) + \sigma_m^n \right\} \quad (4)$$

where $i \in [0, \sum_j F_{in}^{n-1} - 1]$ and $m \in [0, \sum_j F_{out}^n - 1]$ represent the input and output feature maps, respectively.

3.2.3. Pooling layer

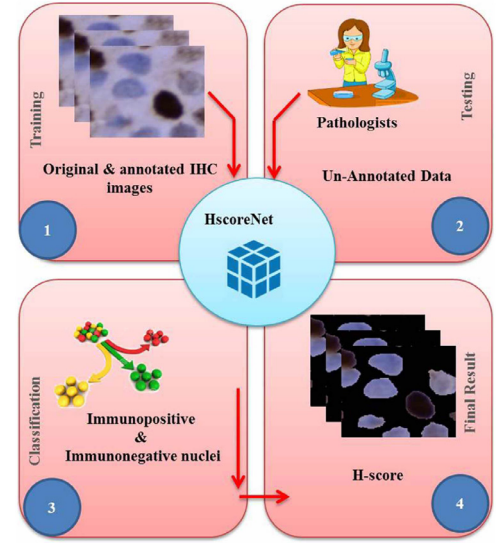
The pooling layer, especially max-pooling, is used to decrease the spatial dimension of an image. The pooling layers reduce over-fitting and computational complexity. The max-pooling and up-pooling layers are employed to create translational invariance over single spatial shifts over input images. To store and capture the boundary information of a feature map, max-pooling indices are memorized for the feature map. In the encoder and decoder parts, max-pooling and up-pooling operations, respectively, are employed.

3.2.4. Rectified linear unit (ReLU)

A rectified linear unit, i.e., ReLu, works as a gradient descent vector and an activation function. This layer boosts the performance of the network by increasing the nonlinearity of the deci-

Table 2
Proposed HscoreNet architecture.

Layer name	Sl. number	Type	Maps	Number of Neurons	Filter Size	Layer Name	Sl. number	Type	Maps	Number of Neurons	Filter Size	Layer's Name	Sl. Number	Type	Maps	Number of Neurons	Filter Size
Encoder	0	Input Image	3	100 × 100	-	Decoder	17	Up-pooling-1	64	4 × 4	2 × 2	Scoring Layer	32	MP-5	64	50 × 50	2 × 2
	1	Conv-1	64	99 × 99	2 × 2		18	Conv-13	64	8 × 8	2 × 2		33	MP-6	64	25 × 25	2 × 2
	2	Conv-2	64	98 × 98	2 × 2		19	Conv-14	64	9 × 9	2 × 2		34	Conv-23	64	23 × 23	2 × 2
	3	MP-1	64	49 × 49	2 × 2		20	Up-pooling-2	64	10 × 10	2 × 2		35	Conv-24	64	22 × 22	2 × 2
	4	Conv-3	64	48 × 48	2 × 2		21	Conv-15	64	20 × 20	2 × 2		36	MP-7	64	11 × 11	2 × 2
	5	Conv-4	64	47 × 47	2 × 2		22	Conv-16	64	21 × 21	2 × 2		37	Conv-25	64	10 × 10	2 × 2
	6	Conv-5	64	46 × 46	2 × 2		23	Conv-17	64	22 × 22	2 × 2		38	MP-8	64	5 × 5	2 × 2
	7	MP-2	64	23 × 23	2 × 2		24	Up-pooling-3	64	23 × 23	2 × 2		39	Conv-26	64	4 × 4	2 × 2
	8	Conv-6	64	22 × 22	2 × 2		25	Conv-18	64	46 × 46	2 × 2		40	MP-9	64	2 × 2	2 × 2
	9	Conv-7	64	21 × 21	2 × 2		26	Conv-19	64	47 × 47	2 × 2		41	FC-1	-	100	1 × 1
	10	Conv-8	64	20 × 20	2 × 2		27	Conv-20	64	48 × 48	2 × 2		42	FC-2	-	3	1 × 1
	11	MP-3	64	10 × 10	2 × 2		28	Up-pooling-4	64	49 × 49	2 × 2						
12	Conv-9	64	9 × 9	2 × 2	29	Conv-21	64	98 × 98	2 × 2								
13	Conv-10	64	8 × 8	2 × 2	30	Conv-22	64	99 × 99	2 × 2								
14	MP-4	64	4 × 4	2 × 2	31	Output Image	3	100 × 100	2 × 2								
15	Conv-11	64	3 × 3	2 × 2													
16	Conv-12	64	2 × 2	2 × 2													

**Fig. 3.** The overall training and testing process of HscoreNet

sion function. Mathematically, ReLu can be defined by Eq. 5 [20]:

$$k(z) = \max(0, z) \quad (5)$$

where k represents models output function.

3.2.5. Fully Connected (FC) layer

The FC layer is often used as a final layer or classification layer in a classification problem. In our DL architecture, the FC layer's height and width of each blob has been assigned to 1.

3.2.6. Dropout Layer

Dropout is a regularization technique which increases the accuracy of a model by reducing overfitting and preventing complex-adaptation on training data [59].

3.2.7. Scoring layer and its significance

The scoring layer has been developed to compute the Histo-score, or H-score, automatically. Pathologists manually calculate the H-score using ER- and PR-stained IHC images to sub-categorize the cancer grade and to decide proper treatment procedures. As per the standard scoring system, the H-score is calculated based on the staining intensity, color expression, and numbers of immunopositive and immunonegative nuclei [14]. The manual process of computing the H-score is time-consuming, error-prone, and dependent on the experience of a pathologist. Hence, we have proposed the scoring layer, which will automatically calculate the H-score.

Our proposed scoring layer consists of four convolution layer block-3, five max-pooling layers, and two fully connected layers (FC). The FC layers are used to obtain independent class probabilities.

3.3. Parameter initialization

In our proposed methodology, each patch (RGB image) of size $\omega \times \omega = 100 \times 100 = 10000$ pixels is used. The numbers of training and validation samples are considered 65% and 35%, respectively, out of 600 images [size of each image is 2048×1536 pixels (width \times height)]. The training and testing datasets were completely two different sets, and there was no overlap exists. To make our model more generalized, we have performed five-fold cross-validation across all the datasets. The training and validation

batch size is set to 28. The testing interval and maximum iteration are assigned to 4500 and 500,000 respectively. The other important parameters include the learning rate ($= 0.01$), weight decay ($=0.045$) and momentum ($=0.9$).

3.4. Training the HscoreNet

We compute the weight or batch normalization layer across the training dataset and use it during validation. As per our observations, these normalizations provide a better result. We train the whole model using the end-to-end stochastic gradient descent method.

3.5. Evaluation measures

Object-level (nuclei detection) and pixel-level (nuclei shape and size) evaluation measures, i.e., precision (Pr), recall (Re), F-score, negative predictive value (NPV) and accuracy, are calculated as follows [24]:

$$\text{Pr (\%)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100 \quad (6)$$

$$\text{Re (\%)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100 \quad (7)$$

$$\text{F score (\%)} = 2 \times \left(\frac{\text{Pr} \times \text{Re}}{\text{Pr} + \text{Re}} \right) \times 100 \quad (8)$$

$$\text{NPV (\%)} = \frac{\text{TN}}{\text{TN} + \text{FN}} \times 100 \quad (9)$$

$$\text{Accuracy (\%)} = \frac{\text{TP} + \text{TN}}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \times 100 \quad (10)$$

In the above (Eqs. 7–11), TP, TN, FP and FN indicate true positive, true negative, false positive and false negative, respectively.

4. Experimental results and discussion

For the training of our model, we sampled 100×100 pixels of high-power fields (HPFs) from the regions of interests of size 2048×1536 pixels. The sampling process was random and had multiple HPFs from each patient. The HPFs of batch size 28 were first converted into lightning memory-mapped database (LMDB) format, supported by the Caffe deep learning framework, and then fed into our proposed model. The training took almost 72 hours. For the testing, the HPFs were sampled from the testing datasets and then converted into LMDB format. Finally, the LMDB-formatted datasets were next fed into the trained model for evaluation.

4.1. Qualitative results

The qualitative results of the *HscoreNet* are shown in Fig. 4. In Fig. 4(b), immunopositive, immunonegative and background are shown in blue, red and pink color, respectively. The final segmented images are shown in 4(c). These results suggest that the segmentation outcomes from *HscoreNet* are visually similar and produce natural boundaries.

4.2. Quantitative results

The *HscoreNet*-based methodology yields an accurate result in term of Pr, Re, F-score, NPV, and accuracy. The high precision value indicates the strength of the proposed model. Fig. 5(a) shows the Receiver Operating Characteristic curve (ROC), where the area under the curve (AUC) is 96%. From this ROC, we can conclude that our proposed *HscoreNet* is an emerging architecture that is

marginally better for ER and PR scoring than similar types of networks. The calculated confidence intervals of precision and recall in a 95% confidence interval using the bootstrapping method are shown in Fig. 6. The *HscoreNet* yields the highest precision and recall values.

4.3. Sensitivity analysis

Fig. 5 (b) illustrates the sensitivity analysis graph. The X-axis represents the window size, and the Y-axis represents the AUC of ROC curves for the *HscoreNet* model. From this figure, it is clearly shown that *HscoreNet* yields a better AUC value when the window size is 100×100 . Hence, we have chosen an image patch size of 100×100 for all of our experiments.

4.4. H-score calculation

As per the scoring system, the H-score is calculated using the formula mentioned in Eq. 11 [14,60–62], where WNS denotes the percentage of weak nuclear staining, MNS is the percentage of moderate nuclear staining, and SNS represents the percentage of strong nuclear staining. The staining intensity of each nucleus has been scored as 0= nuclear stain absent, 1= weak nuclear stain present, 2= moderate nuclear stain present, or 3= strong nuclear stain present, as indicated in the literature.

$$\text{H-score} = (1 \times \text{WNS}) + (2 \times \text{MNS}) + (3 \times \text{SNS}) \quad (11)$$

The H-score value ranges from 0 to 300. The positivity of both ER and PR has been considered when the H-score is equal or greater than 50. The reference ranges and computerized results are reported in Table 3. In this table, the reference range shows the standard category and their ranges, which are already widely used in pathology. We compared the H-score of both the pathologists with the proposed methodology. The receptor negative category error rate is 0.39%, the weakly receptor positive type error rate is 0.46%, and the receptor positive category error rate is 0.63%.

4.5. Comparison with other methods

The performance of the proposed *HscoreNet* was compared with that of the most recent and popular methods, including Immunoratio [63], Leica image analysis software and the most recently published deep learning-based semantic segmentation technique, i.e., SegNet [35].

The Immunoratio and Leica image analysis software have been developed using conventional image analytic techniques. The work flow of Immunoratio software consists of a set of algorithms related to background correction, stain separation, thresholding, particle segmentation, filtering and calculation of the total nuclear area. In contrast, the SegNet technique has been developed using deep learning. The SegNet framework consists of encoder and decoder networks. The encoder layer of SegNet is identical to the VGG16 network. The decoder network has been used for the mapping from low-resolution feature maps to full input resolution feature maps. Fig. 7 shows the results of the comparison of the proposed methodology with the state-of-the-art. From the comparison results, we conclude that our proposed methodology performs better at segmentation, detection, classification and scoring of ER and PR images. Table 4 illustrates the quantitative evaluation of segmentation and classification results on our image dataset with different models. The results mentioned in this table are the average outcomes of five-fold cross-validation. Fig. 8(a) shows comparison graphs of various methodologies and their performances. Fig. 9 shows false positive and false negative cases, where the red-colored arrow in 9(a) shows false positive case and 9(c) signifies false negative case.

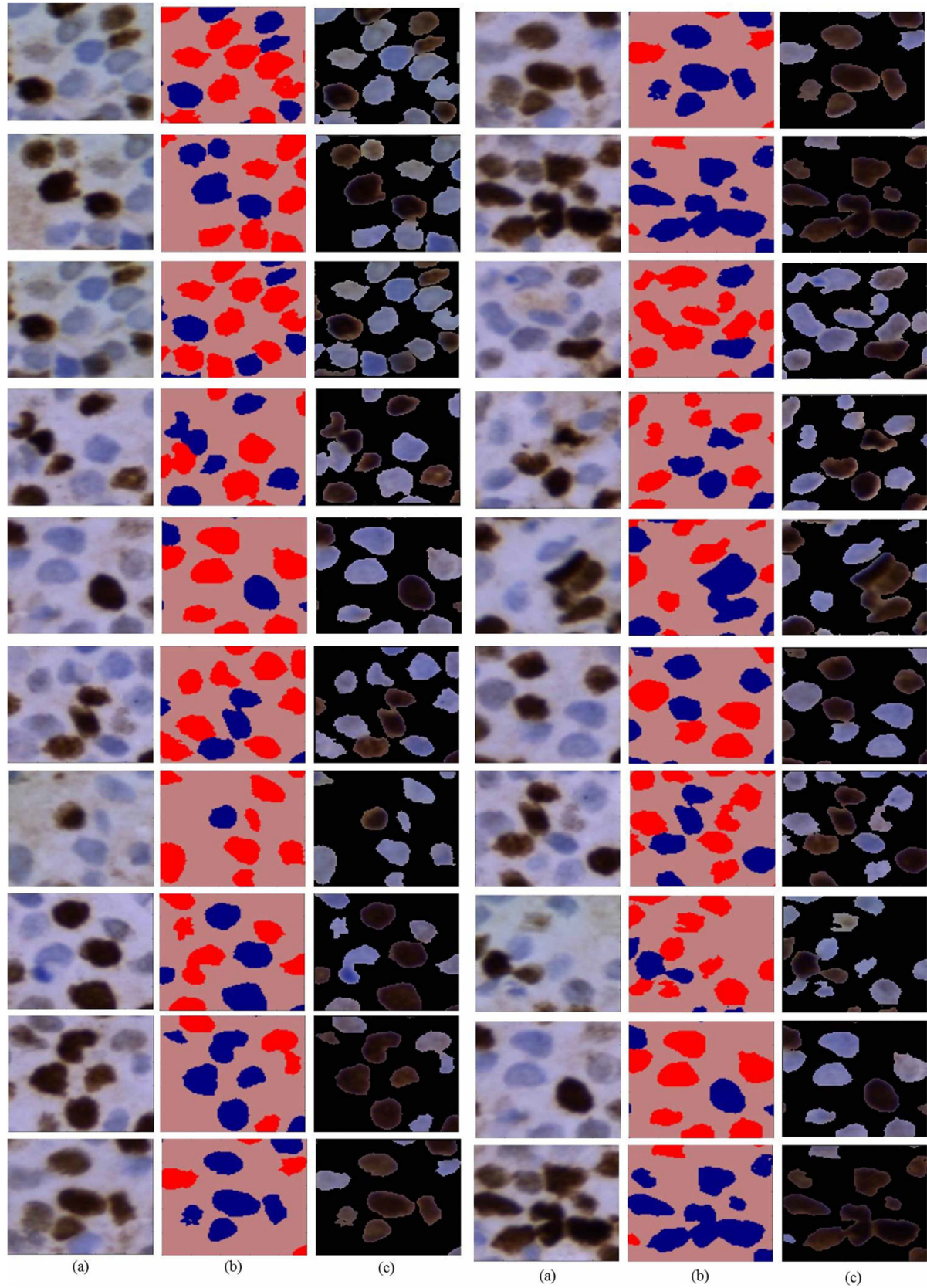


Fig. 4. Segmentation and classification results of ER- and PR-stained BC images at $40 \times$ magnification: (a) original image patch; (b) segmented and classified nuclei corresponding to blue (immunopositive), red (immunonegative) and pink (background) colors; (c) final segmented RGB image, which contains only immunopositive and immunonegative nuclei.

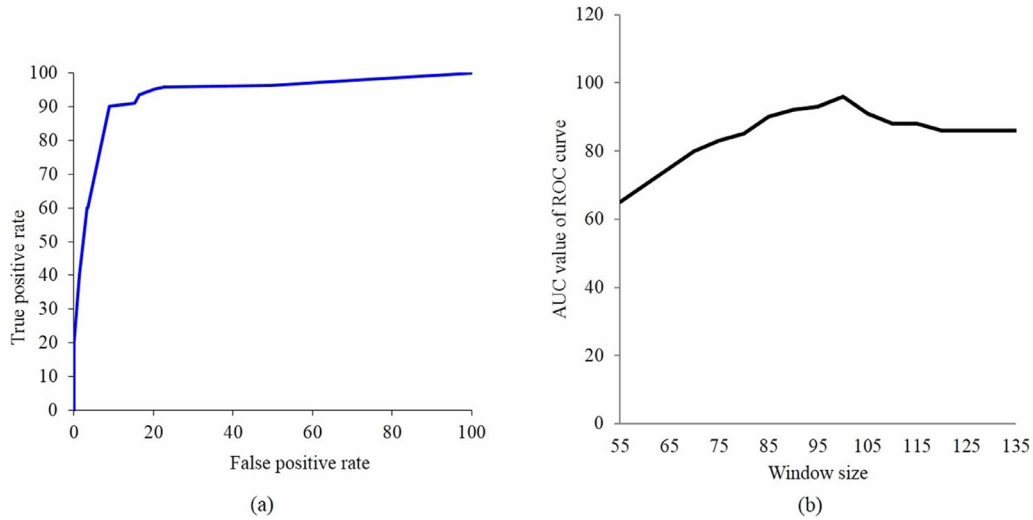


Fig. 5. (a) The ROC curve; (b) plot of window size versus AUC for the HscoreNet model.

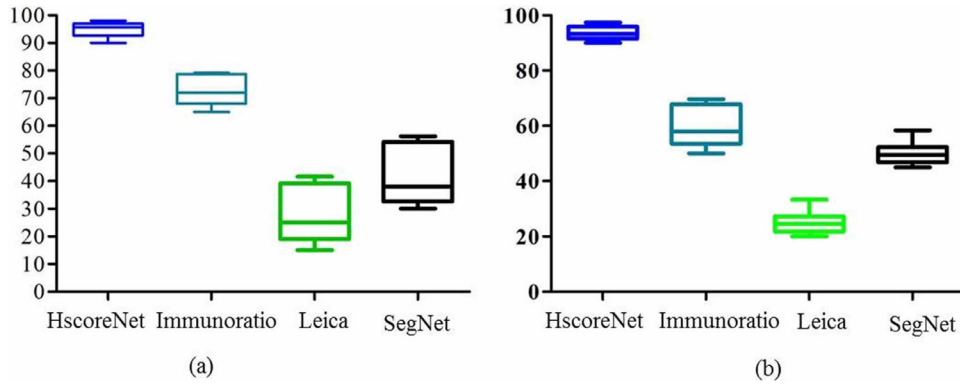


Fig. 6. The estimated confidence interval: (a) precision (b) recall.

Table 3

The results show that the proposed method results are corroborated by the two expert pathologists.

Hormone receptor	Details	Reference range [14]	Expert	Manual score	HscoreNet score
ER and PR	Receptor negative	0-49	Expert-1	27.65	28.23
			Expert-2	28.01	28.23
			Average	27.83	28.23
	Weakly receptor positive	50-99	Expert-1	76.00	76.56
			Expert-2	77.02	76.56
			Average	76.51	76.56
	Receptor positive	≥ 100	Expert-1	118.75	120.00
			Expert-2	120.00	120.00
			Average	119.37	120.00

Table 4

Performance evaluation of the proposed method w.r.t. standard techniques available in the public domain.

Models	Dataset	Pr (%)	Re (%)	F-score (%)	NPV (%)	Accuracy (%)
Immunoratio	Our BC IHC image dataset	78.41	68.65	73.20	53.68	67.63
Leica Image Analysis		37.64	33.33	35.35	0	21.47
SegNet		52.2	58.41	55.12	3.96	38.73
Proposed Methodology		95.87	95.64	96.49	92.13	94.53

4.6. Computation time

Computation time is one of the most vital factors in machine learning. For this reason, we always tried to keep the patch size as small as possible. A small patch size decreases the computation time and increases the detection efficiency. The algorithm was developed using Python and MATLAB libraries on a machine with two

NVIDIA Titan X Pascal GPUs and an AMD Opteron processor with 128 GB RAM. The model took almost 3 days ($24 \times 3 = 72$ hours) for training and, on average, takes 1.67 seconds to detect all the immunopositive and immunonegative nuclei. Fig. 8(b) shows the loss and accuracy curves with respect to the number of iterations. There are a few limitations of our method, which are discussed below:

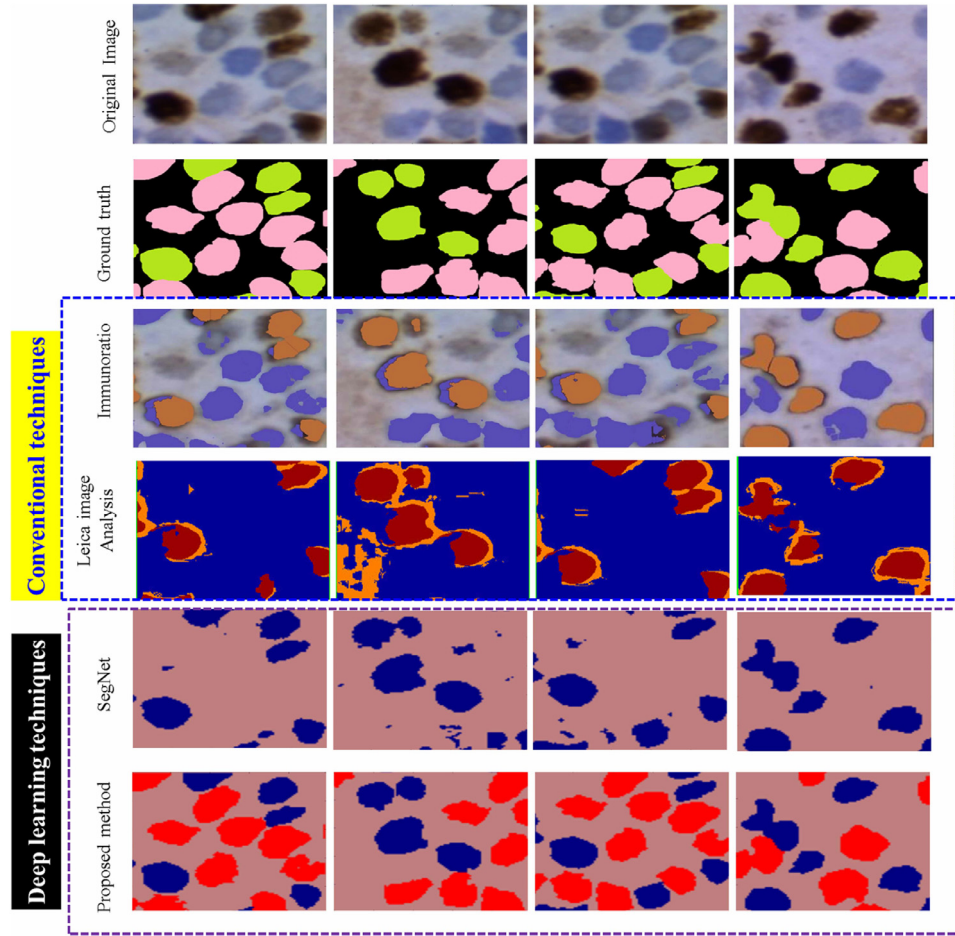


Fig. 7. Outcome-based comparison between the proposed method and the state-of-the-art methods, where it can be observed that the proposed method provides better visualization of immunopositive and immunonegative nuclei.

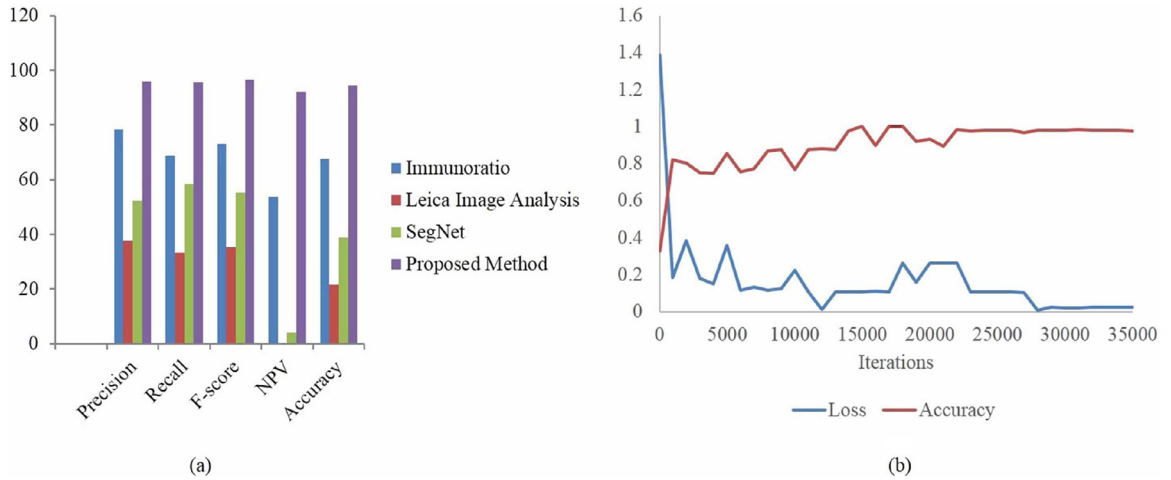


Fig. 8. (a) Comparison graphs of various methodologies and their performances; (b) loss and accuracy curves with respect to iterations.

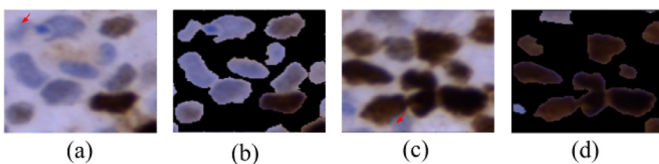


Fig. 9. False positive case: (a) original image; (b) segmented image. False negative case: (c) original image; (d) segmented image.

4.7. Limitations

- The color of immunonegative nuclei sometime creates ambiguities between the background and nucleus. Therefore, this point should be remembered during patch selection.
- Benchmarks and a large number of datasets are missing. This is one of the main obstacles in this research.

5. Conclusion and future work

In this manuscript, we present a novel deep learning network for the scoring of ER and PR using breast IHC images, i.e., *HscoreNet*. In addition to scoring, the network can also be used for semantic segmentation, classification, and image reconstruction. Our proposed network contains mainly three compartments, i.e., encoder, decoder, and scoring layer. The encoder transforms input image pixels into a lower-dimensional representation, whereas the decoder is used to reconstruct the output of the encoder through minimization of a cost function. The output of the decoder is the reconstructed image, which contains only immunopositive and immunonegative nuclei. The final layer is a scoring layer, which is responsible for the calculation of the H-score. The concept of a scoring layer is unique and has been used for the first time for this type of study.

The overall experiments have demonstrated the outstanding performance of *HscoreNet* in comparison with the current state-of-the-art methods. Moreover, it streamlines the human error-prone and time-consuming process and facilitates screening and diagnosis in the early detection of BC. The network has been tested only on the IHC datasets. Hence, the performance might vary if different types of datasets are used, or there might be need to adjust in the model based on the input data.

In the future, it would be interesting to explore the performance of *HscoreNet* on other disease types, e.g., lung cancer, glioblastoma, and prostate cancer. The performance of the model can also be tested on H&E stained images.

Acknowledgments

M. Saha would like to acknowledge Department of Science and Technology (DST), Govt. of India, for providing the INSPIRE fellowship (IVR Number: 201400105113) and Indo French Centre for Promotion of Advanced Research (CEFIPRA) for Raman-Charpak fellowship 2015 (RCF-IN-0071). Rest of the co-authors acknowledge Ministry of Human Resource Development (MHRD), Govt. of India for financial support to carry out this work (Grant no. 4-23/2014 T.S.J. date: 14-02-2014) under SSLs track, a mega-initiative by IIT Kharagpur. Thanks to Banuja Kumari Acharya for helping in reviewing the manuscript.

References

- [1] M. Saha, I. Arun, B. Basak, S. Agarwal, R. Ahmed, S. Chatterjee, R. Bhargava, C. Chakraborty, Quantitative microscopic evaluation of mucin areas and its percentage in mucinous carcinoma of the breast using tissue histological images, *Tissue Cell* 48 (3) (2016) 265–273.
- [2] B. Gecer, S. Aksoy, E. Mercan, L.G. Shapiro, D.L. Weaver, J.G. Elmore, Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks, *Pattern Recognit.* 84 (2018) 345–356.
- [3] R.L. Siegel, K.D. Miller, A. Jemal, *Cancer statistics, 2019*, CA: A Cancer Journal for Clinicians 69 (1) (2019) 7–34.
- [4] A.C. Society, *Breast cancer facts & figures 2017–2018*, 2017.
- [5] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, *Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries*, CA: A Cancer Journal for Clinicians 68 (6) (2018) 394–424.
- [6] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D.M. Parkin, D. Forman, F. Bray, *Cancer incidence and mortality worldwide: sources, methods and major patterns in globocan 2012*, *Int. J. Cancer* 136 (5) (2015) E359–E386.
- [7] M. Saha, C. Chakraborty, *Her2Net: a deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation*, *IEEE Trans. Image Process.* 27 (5) (2018) 2189–2200.
- [8] M. Saha, I. Arun, S. Agarwal, R. Ahmed, S. Chatterjee, C. Chakraborty, *Imprint cytology-based breast malignancy screening: an efficient nuclei segmentation technique*, *J. Microsc.* (2017).
- [9] H. Masmoudi, S.M. Hewitt, N. Petrick, K.J. Myers, M.A. Gavrielides, *Automated quantitative assessment of HER-2/neu immunohistochemical expression in breast cancer*, *IEEE Trans. Med. Imag.* 28 (6) (2009) 916–925.
- [10] H. Irshad, A. Veillard, L. Roux, D. Racocanu, *Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential*, *IEEE Rev. Biomed. Eng.* 7 (2014) 97–114.
- [11] E. Ficarra, S. Di Cataldo, A. Acquaviva, E. Macii, *Automated segmentation of cells with IHC membrane staining*, *IEEE Trans. Biomed. Eng.* 58 (5) (2011) 1421–1429.
- [12] F. Schnorrenberg, C.S. Pattichis, K.C. Kyriacou, C.N. Schizas, *Computer-aided detection of breast cancer nuclei*, *IEEE Trans. Inf. Technol. Biomed.* 1 (2) (1997) 128–140.
- [13] C.M. Guittet, M.J. Varga, P.G. Ducksbury, *Scoring estrogen and progesterone receptors expression based on image analysis*, 2010, US Patent 7,646,905.
- [14] A.A. Thihe, M.J. Chng, P.H. Tan, S. Fook-Chong, *Immunohistochemical expression of hormone receptors in invasive breast carcinoma: correlation of results of h-score with pathological parameters*, *Pathology* 33 (1) (2001) 21–25.
- [15] S.K. Mohsin, H. Weiss, T. Havighurst, G.M. Clark, M. Berardo, L.D. Roanh, T.V. To, Q. Zho, R.R. Love, D.C. Allred, *Progesterone receptor by immunohistochemistry and clinical outcome in breast cancer: a validation study*, *Modern Pathol.* 17 (12) (2004) 1545–1554.
- [16] D.A. Turbin, S. Leung, M.C. Cheang, H.A. Kennecke, K.D. Montgomery, S. McKinney, D.O. Treaba, N. Boyd, L.C. Goldstein, S. Badve, et al., *Automated quantitative analysis of estrogen receptor expression in breast carcinoma does not differ from expert pathologist scoring: a tissue microarray study of 3,484 cases*, *Breast Cancer Res. Treat.* 110 (3) (2008) 417–426.
- [17] J. Lotz, J. Olesch, B. Müller, T. Polzin, P. Galuschka, J. Lotz, S. Heldmann, H. Laue, M. González-Vallinas, A. Warth, et al., *Patch-based nonlinear image registration for gigapixel whole slide images*, *IEEE Trans. Biomed. Eng.* 63 (9) (2016) 1812–1819.
- [18] M. Veta, P.J. van Diest, R. Kornegoor, A. Huisman, M.A. Viergever, J.P. Pluim, *Automatic nuclei segmentation in H&E stained breast cancer histopathology images*, *PLoS One* 8 (7) (2013) e70221.
- [19] M. Saha, S. Agarwal, I. Arun, R. Ahmed, S. Chatterjee, P. Mitra, C. Chakraborty, *Histogram based thresholding for automated nucleus segmentation using breast imprint cytology*, in: *Advancements of Medical Electronics*, Springer, 2015, pp. 49–57.
- [20] M. Saha, C. Chakraborty, I. Arun, R. Ahmed, S. Chatterjee, *An advanced deep learning approach for Ki-67 stained hotspot detection and proliferation rate scoring for prognostic evaluation of breast cancer*, *Sci. Rep.* 7 (2017).
- [21] J.-H. Xue, D.M. Titterton, *t-tests, f-tests and Otsu's methods for image thresholding*, *IEEE Trans. Image Process.* 20 (8) (2011) 2392–2396.
- [22] Y. Zhou, H. Chang, K.E. Barner, B. Parvin, *Nuclei segmentation via sparsity constrained convolutional regression*, in: *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*, IEEE, 2015, pp. 1284–1287.
- [23] F. Kromp, S. Taschner-Mandl, M. Schwarz, J. Blaha, T. Weiss, P.F. Ambros, M. Reiter, *Semi-automated segmentation of neuroblastoma nuclei using the gradient energy tensor: a user driven approach*, in: *Seventh International Conference on Machine Vision (ICMV 2014)*, International Society for Optics and Photonics, 2015. 94451K–94451K.
- [24] J. Xu, L. Xiang, Q. Liu, H. Gilmore, J. Wu, J. Tang, A. Madabhushi, *Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images*, *IEEE Trans. Med. Imag.* 35 (1) (2016) 119–130.
- [25] O. Ronneberger, P. Fischer, T. Brox, *U-net: convolutional networks for biomedical image segmentation*, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2015, pp. 234–241.
- [26] J. Oscanoa, F. Doimi, R. Dyer, J. Araujo, J. Pinto, B. Castaneda, *Automated segmentation and classification of cell nuclei in immunohistochemical breast cancer images with estrogen receptor marker*, in: *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, IEEE, 2016, pp. 2399–2402.
- [27] R. Vijayashree, P. Aruthra, K.R. Rao, *A comparison of manual and automated methods of quantitation of oestrogen/progesterone receptor expression in breast carcinoma*, *J. Clin. Diagnost. Res.* 9 (3) (2015) EC01.
- [28] A. Grote, M. Abbas, N. Linder, H.H. Kreipe, J. Lundin, F. Feuerhake, *Exploring the spatial dimension of estrogen and progesterone signaling: detection of nuclear labeling in lobular epithelial cells in normal mammary glands adjacent to breast cancer*, *Diagnost. Pathol.* 9 (1) (2014) 1.
- [29] R. Mofidi, R. Walsh, P. Ridgway, T. Crotty, E. McDermott, T. Keaveny, M. Duffy, A. Hill, N. O'Higgins, *Objective measurement of breast cancer oestrogen receptor status through digital image analysis*, *Eur. J. Surg. Oncol. (EJSO)* 29 (1) (2003) 20–24.
- [30] E. Rexhepaj, D.J. Brennan, P. Holloway, E.W. Kay, A.H. McCann, G. Landberg, M.J. Duffy, K. Jirstrom, W.M. Gallagher, *Novel image analysis approach for quantifying expression of nuclear proteins assessed by immunohistochemistry: application to measurement of oestrogen and progesterone receptor levels in breast cancer*, *Breast Cancer Res.* 10 (5) (2008) R89.
- [31] B. Chaudhury, M. Zhou, D.B. Goldgor, L.O. Hall, R.A. Gatenby, R.J. Gillies, J.S. Drukeinis, *Using features from tumor subregions of breast DCE-MRI for estrogen receptor status prediction*, in: *Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference on*, IEEE, 2014, pp. 2624–2629.
- [32] A. Mouelhi, M. Sayadi, F. Fnaiech, *A novel morphological segmentation method for evaluating estrogen receptors' status in breast tissue images*, in: *Advanced Technologies for Signal and Image Processing (ATSIP), 2014 1st International Conference on*, IEEE, 2014, pp. 177–182.
- [33] A. Basavanahally, S. Ganesan, M. Feldman, N. Shih, C. Mies, J. Tomaszewski, A. Madabhushi, *Multi-field-of-view framework for distinguishing tumor grade in ER+ breast cancer from entire histopathology slides*, *IEEE Trans. Biomed. Eng.* 60 (8) (2013) 2089–2099.
- [34] F. Cloppet, A. Boucher, *Segmentation of complex nucleus configurations in biological images*, *Pattern Recognit. Lett.* 31 (8) (2010) 755–761.
- [35] V. Badrinarayanan, A. Kendall, R. Cipolla, *SegNet: A deep convolutional*

- encoder-decoder architecture for scene segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* PP (99) (2017). 1–1 doi: 10.1109/TPAMI.2016.2644615.
- [36] A. Kendall, V. Badrinarayanan, R. Cipolla, Bayesian SegNet: Model uncertainty in deep convolutional encoder-decoder architectures for scene understanding (2015). arXiv: 1511.02680.
- [37] J. Wei, Y. Xia, Y. Zhang, M3net: a multi-model, multi-size, and multi-view deep neural network for brain magnetic resonance image segmentation, *Pattern Recognit.* 91 (2019) 366–378.
- [38] S. Chen, C. Ding, M. Liu, Dual-force convolutional neural networks for accurate brain tumor segmentation, *Pattern Recognit.* 88 (2019) 90–100.
- [39] Y. Han, J.C. Ye, Framing u-net via deep convolutional framelets: application to sparse-view ct, *IEEE Trans. Med. Imag.* 37 (6) (2018) 1418–1429.
- [40] Y. Man, Y. Huang, J.F.X. Li, F. Wu, Deep q learning driven ct pancreas segmentation with geometry-aware u-net, *IEEE Trans. Med. Imag.* (2019).
- [41] J. Zhang, J. Du, H. Liu, X. Hou, Y. Zhao, M. Ding, Lu-net: an improved u-net for ventricular segmentation, *IEEE Access* 7 (2019) 92539–92546.
- [42] S. Li, M. Dong, G. Du, X. Mu, Attention dense-u-net for automatic breast mass segmentation in digital mammogram, *IEEE Access* 7 (2019) 59037–59047.
- [43] S. Lian, L. Li, G. Lian, X. Xiao, Z. Luo, S. Li, A global and local enhanced residual u-net for accurate retinal vessel segmentation, *IEEE/ACM Trans. Comput. Biol. Bioinform.* (2019).
- [44] K. Sirinukunwattana, S.E.A. Raza, Y.-W. Tsang, D.R. Snead, I.A. Cree, N.M. Rajpoot, Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images, *IEEE Trans. Med. Imag.* 35 (5) (2016) 1196–1206.
- [45] Y. Zheng, Z. Jiang, F. Xie, H. Zhang, Y. Ma, H. Shi, Y. Zhao, Feature extraction from histopathological images based on nucleus-guided convolutional neural network for breast lesion classification, *Pattern Recognit.* 71 (2017) 14–25.
- [46] H. Chen, X. Qi, L. Yu, Q. Dou, J. Qin, P.-A. Heng, DCAN: deep contour-aware networks for object instance segmentation from histology images, *Med. Image Anal.* 36 (2017) 135–146.
- [47] L. Hou, V. Nguyen, A.B. Kanevsky, D. Samaras, T.M. Kurc, T. Zhao, R.R. Gupta, Y. Gao, W. Chen, D. Foran, et al., Sparse autoencoder for unsupervised nucleus detection and representation in histopathology images, *Pattern Recognit.* 86 (2019) 188–200.
- [48] F. Li, H. Qiao, B. Zhang, Discriminatively boosted image clustering with fully convolutional auto-encoders, *Pattern Recognit.* 83 (2018) 161–173.
- [49] J. Calvo-Zaragoza, A.-J. Gallego, A selectional auto-encoder approach for document image binarization, *Pattern Recognit.* 86 (2019) 37–47.
- [50] M.F. Jamaluddin, M.F. Fauzi, F.S. Abas, J.T. Lee, S.Y. Khor, K.H. Teoh, L.M. Looi, Cell classification in er-stained whole slide breast cancer images using convolutional neural network, in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2018, pp. 632–635.
- [51] R.F. Chang, H.H. Chen, Y.C. Chang, C.S. Huang, J.H. Chen, C.M. Lo, Quantification of breast tumor heterogeneity for ER status, HER2 status, and TN molecular subtype evaluation on DCE - MRI, *Magnet. Resonance Imag.* 34 (6) (2016) 809–819.
- [52] H. Ma, Y. Lu, P.A. Marchbanks, S.G. Folger, B.L. Strom, J.A. McDonald, M.S. Simon, L.K. Weiss, K.E. Malone, R.T. Burkman, et al., Quantitative measures of estrogen receptor expression in relation to breast cancer-specific mortality risk among white women and black women, *Breast Cancer Res.* 15 (5) (2013) R90.
- [53] K. Joensuu, M. Leidenius, M. Kero, L.C. Andersson, K.B. Horwitz, P. Heikkilä, ER, PR, HER2, Ki-67 and CK5 in early and late relapsing breast cancer—reduced CK5 expression in metastases, *Breast Cancer* 7 (2013) 23.
- [54] J.A. Kraus, D.J. Dabbs, S. Beriwal, R. Bhargava, Semi-quantitative immunohistochemical assay versus oncotype DX® qRT-PCR assay for estrogen and progesterone receptors: an independent quality assurance study, *Modern Pathol.* 25 (6) (2012) 869–876.
- [55] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, *Cognit. Model.* 5 (3) (1988) 1.
- [56] H. Bourlard, Y. Kamp, Auto-association by multilayer perceptrons and singular value decomposition, *Biol. Cybernet.* 59 (4) (1988) 291–294.
- [57] L. Lu, X. Zhang, K. Cho, S. Renals, A study of the recurrent neural network encoder-decoder for large vocabulary speech recognition., in: *Interspeech*, 2015, pp. 3249–3253.
- [58] F. Xing, Y. Xie, L. Yang, An automatic learning-based framework for robust nucleus segmentation, *IEEE Trans. Med. Imag.* 35 (2) (2016) 550–566.
- [59] Z. Fang, J. Liu, Y. Li, Y. Qiao, H. Lu, Improving visual question answering using dropout and enhanced question encoder, *Pattern Recognit.* 90 (2019) 404–414.
- [60] L.B. Kinsel, E. Szabo, G.L. Greene, J. Konrath, G.S. Leight, K.S. McCarty, Immunocytochemical analysis of estrogen receptors as a predictor of prognosis in breast cancer patients: comparison with quantitative biochemical methods, *Cancer Res.* 49 (4) (1989) 1052–1056.
- [61] D. Giri, V. Dangerfield, R. Lonsdale, K. Rogers, J. Underwood, Immunohistology of oestrogen receptor and D5 antigen in breast cancer: correlation with oestrogen receptor content of adjacent cryostat sections assayed by radioligand binding and enzyme immunoassay., *J. Clin. Pathol.* 40 (7) (1987) 734–740.
- [62] D. Giri, S. C. Dundas, J. Nottingham, J. E Underwood, Oestrogen receptors in benign epithelial lesions and intraduct carcinomas of the breast: an immunohistological study, *Histopathology* 15 (6) (1989) 575–584.
- [63] V.J. Tuominen, S. Ruotoistenmäki, A. Viitanen, M. Jumppanen, J. Isola, Immunoratio: a publicly available web application for quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and Ki-67, *Breast Cancer Res.* 12 (4) (2010) 1.



puterized Medical Imaging and Graphics, Scientific Reports, Journal Microscopy, etc.



Indu Arun completed her MD in Pathology from Government Medical College, Kottayam, Kerala in 2007. Upon completion of her MD, she worked as a Research Fellow in the Flow Cytometry Lab, National Cancer Institute (NCI), National Institutes of Health, Bethesda, U.S.A. In 2009, she moved to Kerala where she did Senior Residency in the Department of Pathology, Government Medical College, Trichur. She then did a specialty fellowship in Oncopathology at Tata Medical Center Kolkata and subsequently continued in the institution as a consultant in 2012. Her special interests include Lymphomas and applications of flow cytometry in the diagnosis of lymphomas, Breast cancer, and Head and Neck cancer.



a breast clinic. She moved to Tata Medical Centre in September 2011. Her current interests include treatment of younger women with breast cancer and broadening awareness and education about breast cancer.



Sanjoy Chatterjee completed MBBS from Calcutta Medical College before doing his MRCP (Gen Medicine), FRCP, and FRCR (Clinical Oncology) from London. Thereafter he completed Post Graduate Certificate in Medical Education and was given the Fellowship of Higher Education Academy of the UK. During his training, Dr. Chatterjee received various awards including the Gold Medal and Duke of Edinburgh prize. He has 10 years of international exposure working in institutions like Ninewells Hospital, Dundee, Beatson Oncology Centre, Glasgow and Western General Hospital, Edinburgh. He has also done Fellowships in IMRT/IGRT in Royal Marsden Hospital, London and Christie Hospital Manchester and completed original studies and projects in Head and Neck and Breast cancers. Before joining TMCT, Dr. Chatterjee was working as Consultant Clinical Oncologist in Northern Centre for Cancer Care, Freeman Hospital, Newcastle upon Tyne, UK and was a member of the council of the British Association of Head and Neck Oncologists (BAHNO).



Chandan Chakraborty is an Associate Professor at the School of Medical Science and Technology, Indian Institute of Technology Kharagpur, India. He completed his Masters from IIT Bombay and Ph.D. from IIT Kharagpur. His major research interests include medical image processing, statistical pattern classification, machine learning, computer/smartphone -aided diagnosis and prognosis. He received awards including ISCA YoungScientist Award, DAE Young Investigator Award, IBM Faculty Award, etc.