Diving Deep onto Discriminative Ensemble of Histological Hashing & Class-Specific Manifold Learning for Multi-class Breast Carcinoma Taxonomy

Sawon Pratiher[1] Subhankar Chattoraj[2]

Presented by:
Tauheed Ahmed[1]

[1]Indian Institute of Technology, Kharagpur, India
[2]Techno India University, West Bengal, India

Outline

- Introduction
- Literature Survey
- Proposed methodology : Proposed Methodology
- Dataset Details : BreakHis
- Proposed methodology : Pre-Processing
- Proposed methodology : Class Specific Manifold Learning
- Proposed methodology : Image Hashing
- Proposed methodology : Discriminative DNN
- Results : Accuracy, Specificity and Sensitivity
- Results : Comparison with State-of-the-Art
- Conclusion and Future Scope
- References

Histopathological Image Analysis

Importance of Histopathological Image

- Histopathological image is taken as the gold standard by the clinician for diagnosis of cancer.
- Rich underlying tissue structure is easier for cancer diagnosis

Reason of Analysis

- Tool to assist clinician in diagnosing breast cancer patients.
- Early stage diagnosis prevent from deaths in women.

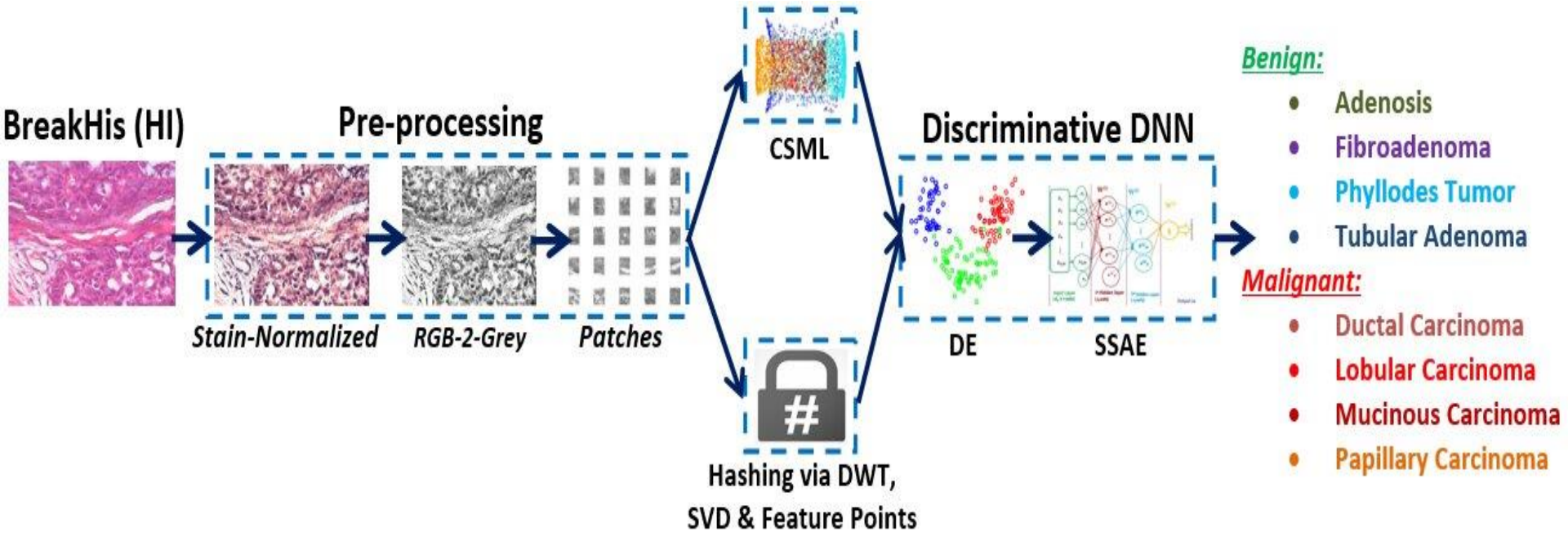
Motivation, Problems and Goals

Motivation	WHO shows breast cancer (BC) as the second most cause of cancer-related deaths in females with an estimate of 627,000 deaths, impacting more than 2.1 million women every year. Computer aided diagnostic (CAD) system for BC diagnosis is important due to the shortage of pathologists in developing countries
Problems	Different CAD system aided deep learning & CNN has been proposed earlier but lacks the proper discrimination of the benign and malignant classes from the underlying complex tissue structure of HI.
Goals	Develop a CAD system capable of identifying BC at higher recognition rate for both binary and multi-class problem. Eliminating the use of high performance computing like GPU which is not readily available in developing countries pathological center.

Literature Survey for Breast Cancer Classification

Author	Research Methodology
Gupta. et. al.	Ensemble Classifier model trained on colour-texture image descriptor (C-TIP).
Fabio. et. al.	Whole slide images extracted at multiple magnification and deep features extracted via DeCAF having multiple feature vectors (MFV).
Wei. et. al.	Vector of locally aggregate descriptors (VLAD) on grassman manifold (GM).
Yang. et. al.	Class structure based (CSD) deep CNN (CSD-CNN) for multi-class classification from multiple whole slide images patches with an augmented multiple instance learning (MIL) layer.
Gupta. et. al.	Pre-trained inception V3 model & VGG19 has been utilized for potent feature vector (FV) computation.
Morillo. et. al.	Multilevel pyramidal information using XGboost & PCA and pre-trained DenseNet was used for training.

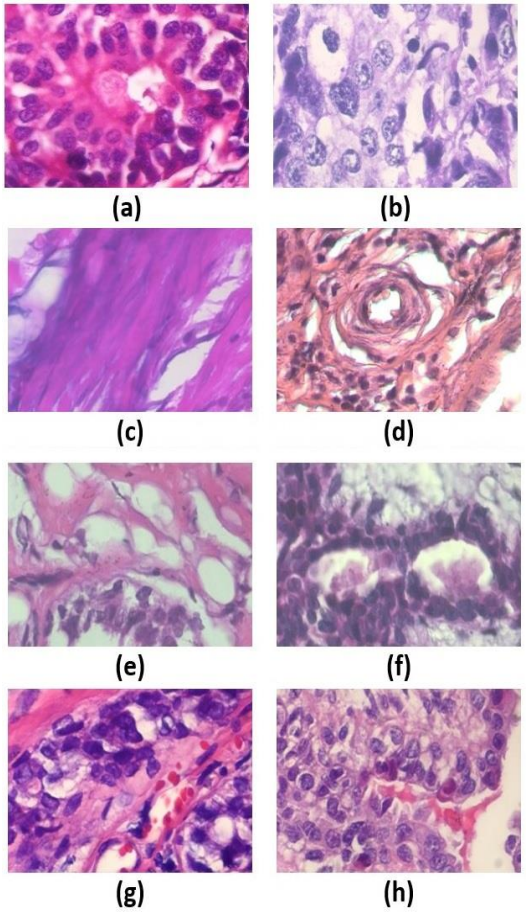
Proposed Methodology



CSML: Class-specific Manifold Learning
DWT: Discrete Wavelet Transform
SVD: Singular Value Decomposition
SSAE: Stacked Sparse AutoEncoder
DE: Discriminative Feature Ensemble via **DCA** (Discriminant Correlation Analysis)

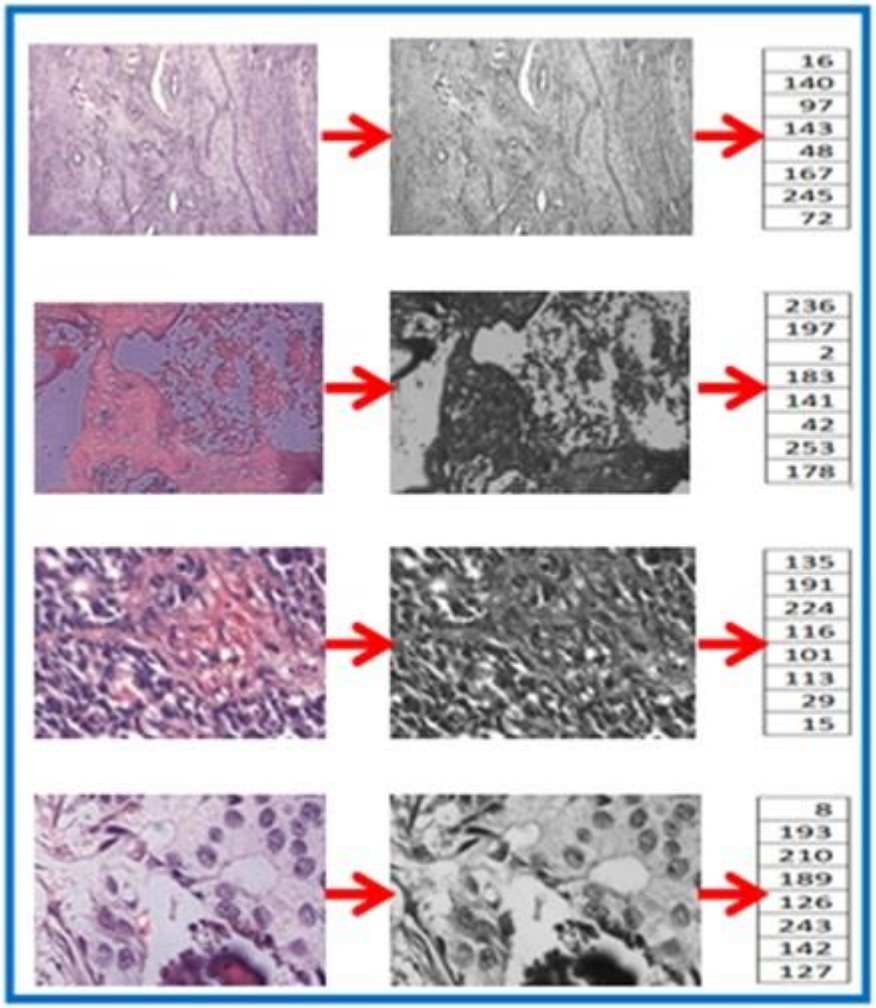
BreaKHis Dataset

CLASS	Sub-Class	Magnification Factor				Total	Patients
		40X	100X	200X	400X		
M	DC	864	903	896	788	3451	**58**
	LC	156	170	163	137	626	
	MC	205	222	196	169	792	
	PC	145	142	135	138	560	
B	A	114	113	111	106	444	
	F	253	260	264	237	1014	**24**
	TA	109	121	108	115	453	
	PT	149	150	140	130	569	
Total		**1995**	**2081**	**2013**	**1820**	**7909**	**82**



(a) Ductal carcinoma (DC), (b) Lobular carcinoma (LC), (c) Mucinous carcinoma (MC), (d) Papillary carcinoma (PC), (e) Adenosis, (f) Fibroadenoma, (g) Tubular adenoma (TA), (h) Phyllodes tumor (PT). (a) to (d) & (e) to (g) corresponds to malignant and benign class respectively

Pre-Processing of HI dataset



Stain normalization proposed comprises of **four** modules that is:

- Strain matrix approximation
- Color deconvolution (CD)
- Channel statistics mapping (nonlinear)
- Image reconstruction.

Proposed methodology for Class Specific Manifold Learning

- Fast alternative of Isomap a.k.a., **Landmark ISOMAP (L-ISOMAP)** for Graph manifold learning is utilized by shortest path computation using Dijkstra's algorithm [22].
- Each neighbor weight computation which signifies the effect of the data point of the neighbor.
- Based on the computed weight low dimension embedding of the data is constructed.

$$m^T(p,q) = \frac{1}{2}\left({F_{pq}}^2 - e_p \frac{1}{m}\sum_l {H^2}_{pl} \right)$$

H^2 is geodesic distance matrix H is element-wise squared &e_p is eigen vector

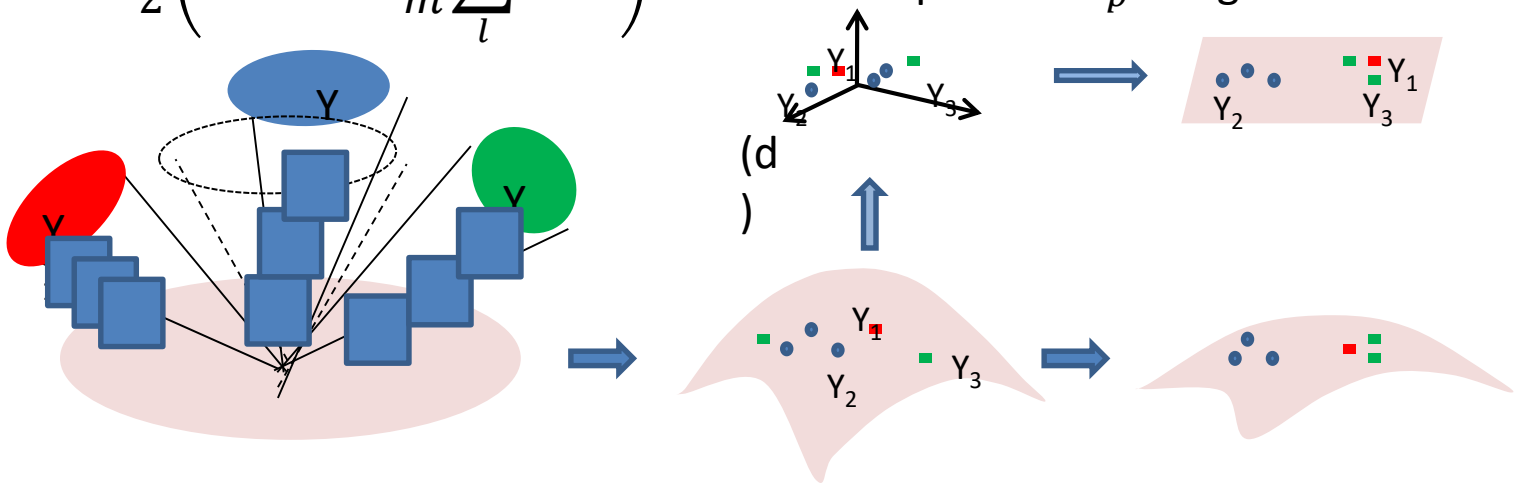


Fig. Representative Class Specific Manifold Learning (CSML)

[22] H. Shi, B. Yin, X. Zhang, Y. Kang and Y. Lei, "A landmark selection method for L-Isomap based on greedy algorithm and its application," 2015 54th IEEE Conference on Decision and Control (CDC), Osaka, 2015, pp. 7371-7376. doi: 10.1109/CDC.2015.7403383

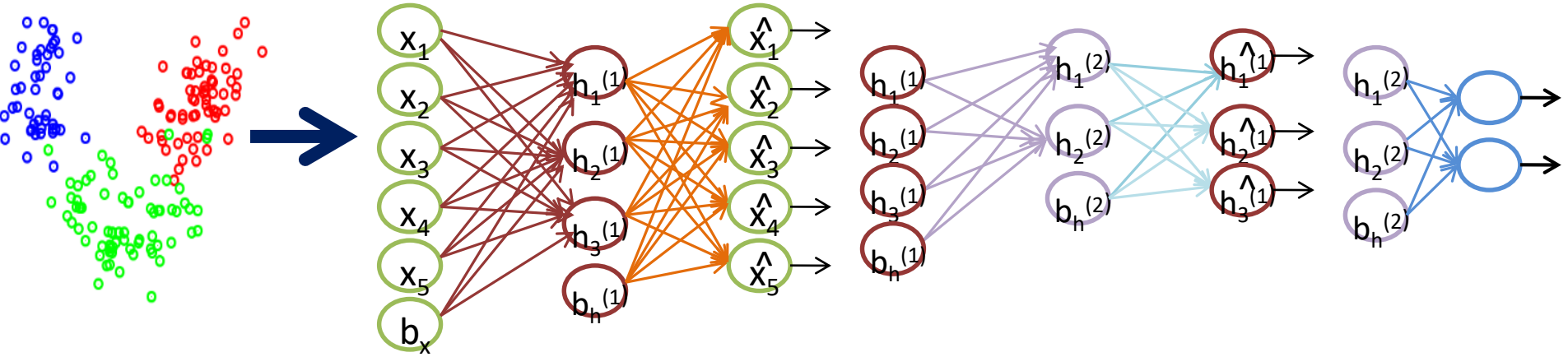
Histological Hashing for Local Malignancy Signatures

- **Discrete Wavelet Transform (DWT) Based Image Hash:** Robust & compact hash via 2D DWT on HI, which decomposes into four sub-bands. Both edge or high frequency information & coarse stable low-frequent coefficients are perceived via DWT coefficients [23].
- **Hashing Via Singular Value Decomposition (SVD):** SVD based image hashing has robust tolerance to small rotational changes until 10· & is translation in-variance & incorporates low rank approximation of original normalized sub-image of HI & non-correlated directional feature space encoding [24].
- **Feature Point Based Image Hashing:** Statistical image features like Hessian affine, maximally stable extremal region (MSER) detectors, Harris corner detector & feature points based on end-stopping behavior like end-stopped Wavelets such as Morlets preserves the momentous image geometry feature-space constraints of 2-D HI pixels & mapped to 1- D feature vector which is compressed to generate the hash vector [25].

[23] Tang, Zhenjun, et al. ”Robust image hashing via colour vector angles and discrete wavelet transform.” IET Image Processing 8.3 (2013): 142-149.
[24] Kozat, Suleyman Serdar, Ramarathnam Venkatesan, and Mehmet Kivanc¸ Mihc¸ak. ”Robust perceptual image hashing via matrix invariants.” 2004 International Conference on Image Processing, 2004. ICIP’04. Vol. 5. IEEE, 2004.
[25] Monga, Vishal, and Brian L. Evans. ”Perceptual image hashing via feature points: performance evaluation and tradeoffs.” IEEE transactions on Image Processing 15.11 (2006): 3452-3465.

Cascaded DCA and SSAE aided discriminative DNN ensemble

- Discriminative feature ensemble of holistic CSML & shallow hash signatures is fused via discrimination correlation analysis [27].
- Stacked Sparse Autoencoder (SSAE) [28] involves optimal parameter θ = (V; ak; ay) computation by minimizing the error between the model input and output. Rudimentary details can be traced from [21].



DE: Discriminative Feature Ensemble via **DCA** (Discriminant Correlation Analysis) followed by **SSAE** (Stacked Sparse AutoEncoder)

[21] Bose, Tulika, Angshul Majumdar, and Tanushyam Chattopadhyay. "Machine Load Estimation Via Stacked Autoencoder Regression." 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2018.

[27] Haghighat, Mohammad, Mohamed Abdel-Mottaleb, and Wadee Alhalabi. "Discriminant correlation analysis for feature level fusion with application to multimodal biometrics." Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on. IEEE, 2016.

[28] Qi, Yu, et al. "Robust feature learning by stacked autoencoder with maximum correntropy criterion." Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on. IEEE, 2014.

Experimental Results & Discussion

Hyper-parameter Tuning

- A maximum epoch of 500 & number hidden layer neurons in the first and second autoencoders (AE) is optimally kept at 500 & 300 respectively.
- Further, L_2 weight regularization, sparsity regularization & sparsity proportion in the two AE are set to 0.001, 4 & 0.15.
- $tanh$ activation function (AF) in the hidden units is connected to by fully connected layer & classification done via a softmax layer.
- Piece-wise back propagation learning via stochastic gradient descent (SGD) algorithm is envisaged with a learning rate of 10^{-4}.
- The initial random weights drawn uniformly from $left$ $[-0{:}1;\ 0{:}1\ ]$, gradient decay factor of *0.2*, momentum of *0.6*, learning rate drop period (LRDP) of 5.

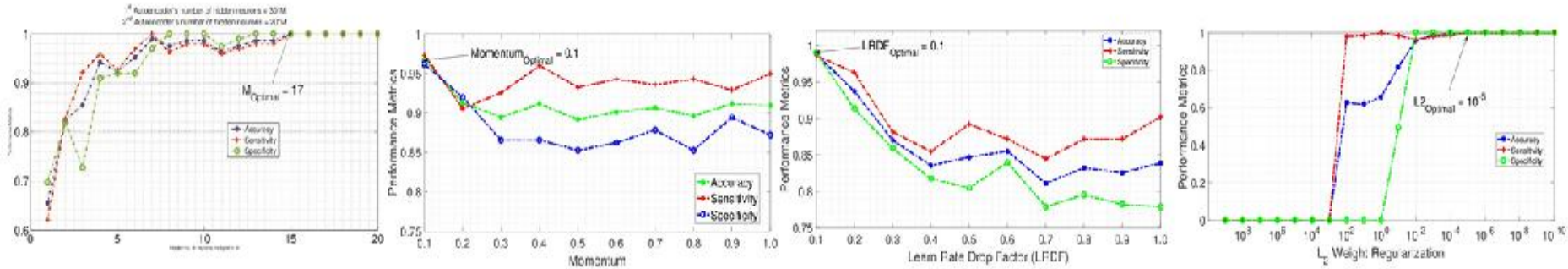


Fig: Hyper-parameter tuning in DNN. (a) hidden size, (b) momentum, (c) learn rate drop factor, (d) L_2 weight regularization

Experimental Results & Discussion

Continued…

Performance Evaluation

- Each of these feature vectors are randomized to eschew the bias of the training parameters & prevent over-fitting. 70% samples are selected for training & the rest 30% is used for testing.
- Standard evolution metrics was utilized for performance evaluation.

$$SN = \frac{TP}{TP + FN},$$

$$SP = \frac{TN}{TN + FP}$$

$$AC = \frac{TP + TN}{TP + TN + FP + FN}$$

Where, TP = true positive, TN = true negative, FP = false positive & FN = false negative respectively.

Experimental Results & Discussion

Continued…

- Binary and Multi-class Classification using the proposed method.

Class	PM (%)	Magnification Factor			
		40X	100X	200X	400X
Binary-class	AC	99.1	98.7	99.3	98.4
	SN	100	100	100	100
	SP	98.3	97.9	98.5	96.8
Multi-class	AC	95.1	95.7	95.8	95.2

Table II: Binary-class & Multi-class classification performance

- 100% sensitivity for all magnification factor ensures that all malignant classes are recognized correctly.
- Pathologist may invest more time for identifying benign cases with our system.
- Proposed CSML & HI hash extraction, discriminative ensemble & DNN learning effectively learns the underlying tissue structure of HI in intrinsic quasi-isometric low embedded space.
- Few literature available in the multi-class classification of BC from HI

Experimental Results & Discussion

Continued…

- Class-specific Classification using the proposed method.

CLASS	Sub-Class	Magnification Factor			
		40X	100X	200X	400X
M	DC	96.7	97	97.6	96.9
	LC	93.8	94.7	97.6	96.9
	MC	94.4	95.8	96.6	94.9
	PC	93.1	95.2	93	93.1
B	A	93.5	93.9	97	94.1
	F	95	95.6	94.9	95.1
	TA	94.5	94.1	94.3	94.4
	PT	94.4	95.1	95.2	94.7

Table III: Class-specific classification performance

- The Proposed method is further examined towards the class-specific classification of BC to demonstrate its efficacy towards practical use from a clinical perspective.
- High recognition rate and lower computational training time suggests its use for deployment in developed countries.

BC classification results of the proposed & state-of-art methods for Binary-Class

Ref, Years	Feature + Method	Magnification Factor			
		40X	100X	200X	400X
[13], 2016	CNN, FR	85.6	83.5	82.7	80.7
[8], 2016	CLBP, SVM	77.4	76.4	70.2	72.8
[9], 2017	C-TID, EC	87.2	88.2	88.8	85.8
[10], 2017	CNN, DeCAF, MFV	84.6	84.8	84.2	81.6
[11], 2017	GML, BI-LSTM	96.2	97.2	97.1	95.4
[14], 2017	VLAD, GM	91.8	92.1	91.4	90.2
[15], 2017	I-EM, CFV, CNN, GMM	87.7	87.6	86.5	83.9
[16], 2018	CSDCNN	95.8	96.9	96.7	94.9
[17], 2018	DCNN	95.1	96.3	96.9	93.8
[18], 2018	Dense SIFT, SURF, BOW, LCLC	98.33	97.12	97.85	96.15
[19], 2018	FV, CSE	87.5	88.6	85.5	85.0
This Work	**CSML, Hashing, DNN**	**99.1**	**98.7**	**99.3**	**98.4**

BC classification results of the proposed & state-of-art methods for Multi-Class

Ref, Years	Feature + Method	Magnification Factor			
		40X	100X	200X	400X
[15], 2018	CSDCNN	92.8	93.9	93.7	92.9
[18], 2018	SIFT + BoW	41.80	38.56	49.75	38.67
	SURF + BoW	53.07	60.80	70.00	51.01
	DSIFT + LLC	60.58	57.44	70.00	49.96
	SURF + LLC	80.37	63.84	74.54	54.70
	DSIFT, BOW + SVM	18.77	17.28	20.16	17.49
	SURF, BOW + SVM	49.65	47.00	38.84	29.50
	DSIFT, LLC + SVM	48.46	49.44	43.97	32.60
	SURF, LLC + SVM	55.80	54.24	40.83	37.20
	CNN, SVM-RBF	75.43	71.20	67.27	65.12
This Work	**CSML, Hashing, DNN**	**95.1**	**95.7**	**95.8**	**95.2**

BC classification results of the proposed & state-of-art methods for Multi-Class

Class	Ref, Years	SC	Magnification Factor			
			40X	100X	200X	400X
Malignant	[17], 2018	DC	91.51	90.77	91.14	92.74
	This Work	DC	**96.7**	**97**	**97.6**	**96.9**
	[17], 2018	LC	78.72	54.90	63.27	56.10
	This Work	LC	**93.8**	**94.7**	**92.8**	**93.1**
	[17], 2018	MC	70.49	82.09	61.02	70.59
	This Work	MC	**94.4**	**95.8**	**96.6**	**94.9**
	[17], 2018	PC	67.44	83.72	57.50	68.29
	This Work	PC	**93.1**	**95.2**	**93**	**93.1**
Benign	[17], 2018	A	85.29	79.41	84.85	90.63
	This Work	A	**93.5**	**93.9**	**97**	**94.1**
	[17], 2018	F	86.84	91.03	91.14	77.46
	This Work	F	**95**	**95.6**	**94.9**	**95.1**
	[17], 2018	TA	75.56	93.33	76.19	82.05
	This Work	TA	**94.5**	**94.1**	**94.3**	**94.4**
	[17], 2018	PT	76.19	63.89	62.50	58.82
	This Work	PT	**94.4**	**95.1**	**95.2**	**94.7**

Conclusion and Future Work

- Deep discriminative CAD for multi-class **(between benign & malignant sub-classes)** BC characterization is introduced.
- The method implements deep contextual grading of hybrid holistic level CSML representations & local hash signatures of HI.
- Superior performance as compared to the existing state-of-the-art.
- In particular, it shows high specificity towards malignant sub-classes.
- Currently, the method is being escalated to include deeper structures using graph CNN & sequential contextual learning with other tissue images to investigate the diagnostic modality.

Acknowledgment

- The authors would like to thank **Fabio Alexandre Spanhol et.al.,** fo the publicly available BreaKHis dataset.

- **Mr. Tauheed Ahmed** for presenting this work on behalf of us.

THANK YOU

- Sawon Pratiher (Email: sawon@iitkgp.ac.in)
- Subhankar Chattoraj (Email: chattorajsubhankar@gmail.com)

References

[1] https://www.wcrf.org/dietandcancer/cancer-trends/ breast-cancer-statistics
[2] https://www.who.int/cancer/prevention/diagnosis-screening/ breast-cancer/en/ \
[3] Stenkvist, Bjrn, et al. ”Computerized nuclear morphometry as an objective method for characterizing human cancer cell populations”. Cancer research 38.12 (1978): 4688-4697.
[4] Boucheron, Laura E., B. S. Manjunath, and Neal R. Harvey. ”Use of imperfectly segmented nuclei in the classification of histopathology images of breast cancer.” Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on. IEEE, 2010.
[5] Kowal, Marek, et al. ”Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images”. Computers in biology and medicine 43.10 (2013): 1563-1572.
[6] Lacquet, F. A., et al. ”Slide preparation and staining procedures for reliable results using computerized morphology.” Archives of andrology 36.2 (1996): 133-138
[7] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot and B. Yener, ”Histopathological Image Analysis: A Review,” in IEEE Reviews in Biomedical Engineering, vol. 2, pp. 147-171, 2009. doi: 10.1109/RBME.2009.2034865
[8] Spanhol, Fabio A., et al. ”A dataset for breast cancer histopathological image classification.” IEEE Transactions on Biomedical Engineering 63.7 (2016): 1455-1462.
[9] Gupta, Vibha, and Arnav Bhavsar. ”Breast Cancer Histopathological Image Classification: Is Magnification Important?.” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017.

References

[10] Fabio A. Spanhol, Paulo R. Cavalin y, Luiz S. Oliveira, Caroline Petitjean, and Laurent Heutte, ”Deep Features for Breast Cancer Histopathological Image Classification”,2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC).

[11] S. Pratiher, S. Chattoraj, S. Agarwal and S. Bhattacharya, ”Grading Tumor Malignancy via Deep Bidirectional LSTM on Graph Manifold Encoded Histopathological Image,” 2018 IEEE International Conference on Data Mining Workshops (ICDMW), Singapore, Singapore, 2018, pp. 674-681. doi: 10.1109/ICDMW.2018.00104

[12] Dimitropoulos K, Barmpoutis P, Zioga C, Kamas A, Patsiaoura K, Grammalidis N (2017), ”Grading of invasive breast carcinoma through Grassmannian VLAD encoding”. PLoS ONE 12(9): e0185110. https: //doi.org/10.1371/journal.pone.0185110

[13] F. A. Spanhol, L. S. Oliveira, C. Petitjean and L. Heutte, ”Breast cancer histopathological image classification using Convolutional Neural Networks,” 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, 2016, pp. 2560-2567. doi: 10.1109/IJCNN.2016.7727519

[14] Song Y., Chang H., Huang H., Cai W. (2017) ”Supervised Intraembedding of Fisher Vectors for Histopathology Image Classification”, Medical Image Computing and Computer-Assisted Intervention - MICCAI 2017 Lecture Notes in Computer Science, vol 10435. Springer, Cham

B. Wei, K. Li, S. Li, Y. Yin, Y. Zheng, Z. Han, ”Breast cancer multi-classification from histopathological images with structured deep learning model Sci Rep, 7 (2017), p. 4172, https://doi.org/10.1038/ s41598-017-04075-z.

[16] Vo, Duc My, Ngoc-Quang Nguyen, and Sang-Woong Lee.”Classification of breast cancer histology images using incremental boosting convolution networks.” Information Sciences 482 (2019): 123-138.

[17] D. Bardou, K. Zhang and S. M. Ahmad, ”Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks,” in IEEE Access, vol. 6, pp. 24680-24693, 2018. doi: 10.1109/ACCESS.2018.2831280

References

[18] Yang Song, Hang Chang, Yang Gao, Sidong Liu, Donghao Zhang, Junen Yao, Wojciech Chrzanowski & Weidong Cai, ”Feature learning with component selective encoding for histopathology image classification” 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018).

[19] Vibha Gupta, Arnav Bhavsar ”Sequential Modeling of Deep Features for Breast Cancer Histopathological Image Classification”, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2018, pp. 2254-2261.

[20] Sanchez-Morillo D., Gonzalez J., Garcıa-Rojo M., Ortega J. (2018) Classification of Breast Cancer Histopathological Images Using KAZE Features. In: Rojas I., Ortuno F. (eds) Bioinformatics and Biomedical ˜ Engineering. IWBBIO 2018. Lecture Notes in Computer Science, vol 10814. Springer, Cham.