

Prognostic Prediction of Histopathological Images by Local Binary Patterns and RUSBoost

Jarle Urdal*, Kjersti Engan*, Vebjørn Kvikstad†, Emilius A.M. Janssen†‡

*Department of Electrical Engineering and Computer Science, University of Stavanger, Norway

Email: {jarle.urdal, kjersti.engan}@uis.no

†Department of Pathology, Stavanger University Hospital, Norway

‡Department of Mathematics and Natural Science, University of Stavanger, Norway

Abstract—A high recurrence rate, and progression to higher stages are observed for patients diagnosed with urothelial carcinoma (previously known as transitional cell carcinoma). Low prognostic value of the current grading systems result in extensive follow-up of patients for multiple years after first diagnosis. Although, the aid of computer systems for prognosis prediction of superficial urothelial carcinomas have been proposed, earlier analyses have been focused on using morphological features of cells and attributes describing the patient. In this study, we propose a system to aid in the prediction of prognostic information based on a texture analysis of histopathological images of superficial urothelial carcinoma. The analyses are conducted using the local binary pattern (LBP) and local variance (VAR) operators followed by a RUSBoost classifier. A dataset of 42 patients, consisting of 13 patients without recurrence, 14 with recurrence but not progression and 15 patients with progression are studied. Using a leave-one-out cross-validation, an accuracy of 70% and sensitivity of 84% is achieved.

I. INTRODUCTION

Bladder cancer is the 6th most common cancer in the world, with 429.000 new incidents reported in 2012 [1]. While several types of bladder cancer exists, urothelial carcinoma is the most common [2]. Superficial tumours are diagnosed in 66 – 75% of new patients on first presentation of the carcinoma [3], and are confined to the epithelium and subepithelial connective tissue and does not invade the muscles. Grading of tumours are currently done by analysing multiple features from the extracted cell-tissue by a trained histopathologists. Based on these features, a grade is given using the WHO73 (1973 World Health Organization classification of papillary urothelial neoplasms) or WHO04 (2004 World Health Organization classification of papillary urothelial neoplasms) guidelines. As this grading is done manually, interobserver variability has been a concern and has been reported as a challenge for both grading systems [2]. Prognostic values of the grading systems are low, with only a significant difference ($p = 0.04$) found for progression free survival between PNLMP and high grade (WHO04) [4]. An extensive follow-up is therefore given to patients diagnosed with superficial urothelial carcinoma, as recurrences are found in 50 – 70% of patients, and progression to a higher stage in 10 – 30% of patients [2].

Multiple systems for automatic carcinoma detection and diagnoses prediction to aid histopathologists and clinicians in cancer diagnosis has previously been proposed, an overview of the field is given by He et al [5]. A system for automatic

segmentation of eight different types of tissue using multiple texture descriptors has been proposed by Kather et al [6], achieving an accuracy of 87.4%. A system for prognostic prediction of urothelial carcinoma was proposed by Spyridonos et al [7]. Based on a probabilistic neural network using morphological and features computed from the nucleus, Spyridonos et al achieved a total accuracy of 71.7% for prediction of recurrence. A different approach was explored by Borgi et al with a diagnostic system based on associate classification. Using 24 attributes to describe each patient (age, sex, smoking habits etc.), the proposed method correctly identified 38.53% of patients without recurrence and 68.83% of patients with recurrence [8]. The use of image processing and machine learning techniques does, however, introduce the opportunity of using texture analysis on entire histopathological images, without the required segmentation to identify each cell nucleus, resulting in a considerable simplification of the analysis. In this paper, we use local binary pattern (LBP) on histopathological images for prognostic purposes in urothelial carcinoma without segmenting individual cells. LBP have previously been applied with promising results on segmentation of histopathological images [6], other medical applications [9], [10], as well as other applications [11], [12].

The aim of this study is to design a framework to aid in the prognostic prediction of urothelial carcinoma by using a texture analysis on histopathological images. The texture analysis will be conducted using the LBP and local variance (VAR) texture descriptors.

A. Data material

The data set used in this study are previously used by Mangrud et al [13] and consist of 42 cases of primary superficial urothelial carcinoma diagnosed at the Department of Urology and Pathology at Stavanger University Hospital during the period January 1. 2002 - December 31. 2006. The study was approved by the Norwegian Regional Ethics Committee (REK Vest, #106/09) before the start of the study. With approval from REK Vest, informed consent was not obtained as the tissue samples had already been removed for diagnostic and treatment purposes [2]. Tumor tissue was fixed in 4% buffered formaldehyde, dehydrated and embedded in paraffin, and four μm thick sections stained with Haematoxylin-Erythrosine-Saffron (HES) was used [2].

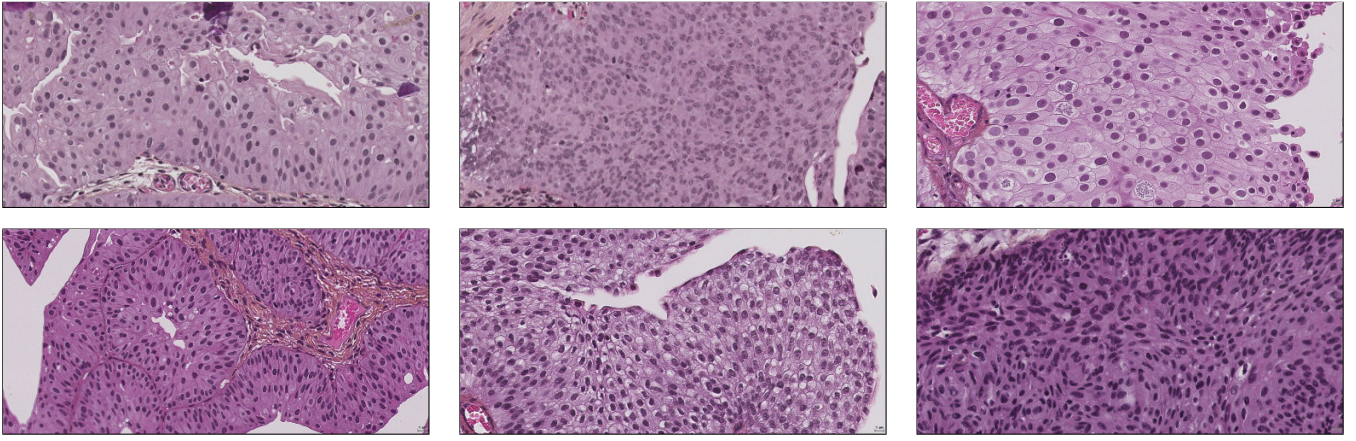


Fig. 1. Example images of intersample variations between prognoses and within each prognosis. No-recurrence (left column), recurrence (middle column), and progression (right column).

The original study considered follow-up information within a five year period. This study will however consider all available follow-up information, as progression occurred for one of the 42 patients after five years and one month. The data set then consist of 15 cases with progression and 27 cases without progression, table I, with two cases for each prognosis shown in Figure 1.

TABLE I
PROGNOSTIC FOLLOW-UP INFORMATION

	No recurrence	Recurrence	
		No progression	Progression
# of patients	13	14	15

Follow-up information was acquired from the Department of Pathology at Stavanger University Hospital.

The histopathological images are scanned at a high resolution, image processing using complete images were not considered feasible due to the high computational requirement. As a consequence of this, smaller sections (at a resolution of $1918 \times 921 \times 24\text{bit}$) using a $400\times$ magnification were chosen in collaboration with a trained pathologist with the goal of illustrating cell maturation from the connective tissue to the outer cell layer. Regions with artefacts due to heating damage or other external causes were avoided. Folded tissue was included in one image, as the chosen section was otherwise a good representation of the tumour.

II. METHODS

A. Local binary pattern

Local binary pattern provides a robust way of describing local texture in a neighbourhood [14]. A label describing each pixel in the image is computed using the neighbourhood defined by P points on the circle with radius R from the pixel itself. Each neighbour are thresholded using the value

of the center pixel. By assigning weights to each neighbour, a numerical value describing the texture is obtained.

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p, \quad s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (1)$$

Where g_c and g_p denotes the center and a neighbouring pixel. The operator can be extended to facilitate for rotation invariant and uniform textures [15].

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP_{P,R}) \leq 2 \\ P + 1 & \text{otherwise} \end{cases} \quad (2)$$

Where $U(\cdot)$ describes number of transitions in the binary label. The local variance of a neighbourhood, denoted $VAR_{P,R}$, is given by

$$VAR_{P,R} = \frac{1}{P} \sum_{p=0}^{P-1} (g_p - \mu)^2, \quad \mu = \frac{1}{P} \sum_{p=0}^{P-1} g_p \quad (3)$$

When keeping the number of neighbouring pixels fixed while increasing the radius, an increasingly sparse representation of the image is obtained. A multi-resolution LBP combined with Gaussian filtering was proposed by Mäenpää et al to overcome this challenge [16]. The method uses an exponentially growing radius and non-overlapping effective areas. This multi-resolution method is adopted here. The radius for each scale is, however, manually chosen instead of using the exponentially growing radius proposed by Mäenpää.

B. RUSBoost

Creating an effective classification model using an imbalanced data set can be challenging, as traditional methods tend to favour the largest classes [17]. Multiple algorithms have previously been proposed to alleviate this challenge, including methods utilizing data resampling and boosting techniques [18]. Data sampling eliminate the class imbalance by introducing new examples in the smaller classes, or removing examples in the larger classes. Boosting, however, creates a

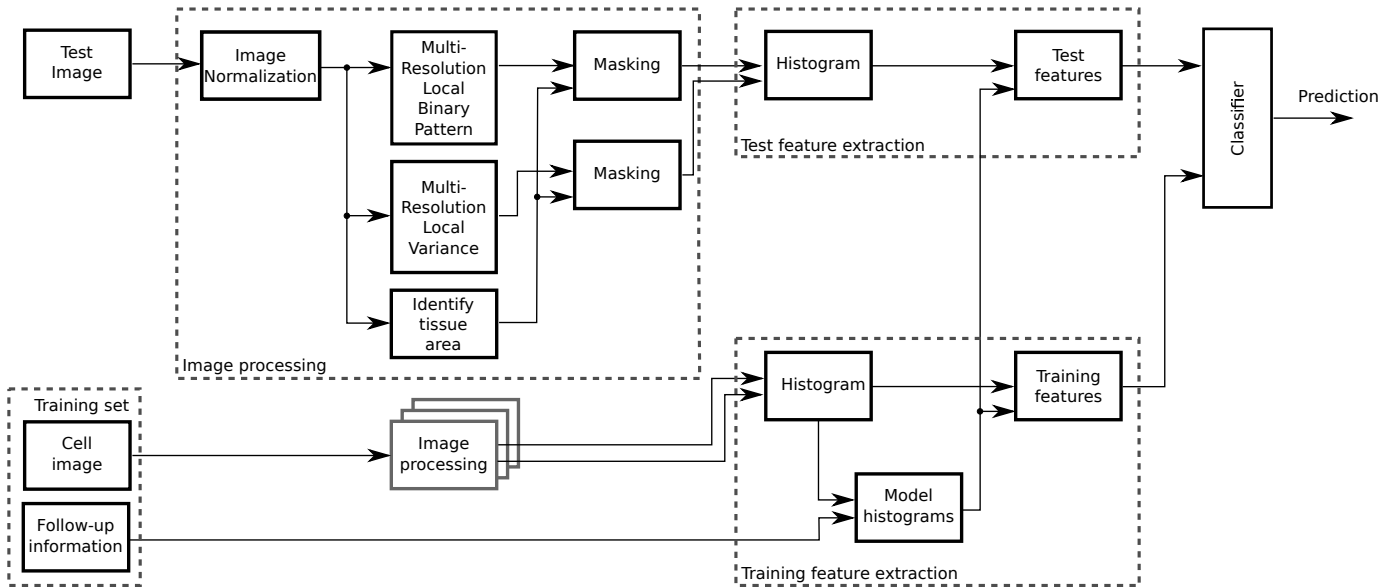


Fig. 2. Block diagram of the proposed method. Input images are normalized, texture operators are applied and regions of tissue in the images identified. A histogram is created for all pixels identified as tissue. Features are computed as the difference between the image histogram, and model histograms computed from the training.

collection of weak learners, where each weak learner perform better than random guessing. These are in turn used to create a more complex strong learner to describe the data set [19]. SMOTEBoost, proposed by Chawla et al [20] implement both techniques. The method utilize a data oversampling technique, where new synthetic features are added between existing features in a continuous feature space, in combination with AdaBoost [19]. The drawback of this method is, however, the increased computational requirement as a consequence of the introduced synthetic examples. An alternative method, RUSBoost, was proposed by Seiffert et al [17] to reduce the computational requirement of SMOTEBoost. The method utilize a random under sampling technique in combination with AdaBoost. The primary drawback of using under sampling techniques, is the loss of information. The combination of under sampling and boosting overcomes this drawback, as the removed examples are likely to be used in another iteration of the boosting technique. The method also achieve similar performance as SMOTEBoost [17].

III. PROPOSED SYSTEM

A block diagram of the proposed system is shown in Figure 2. In the following section, the different parts of the system will be explained.

A. Image processing

The section describes the operations contained in the *image processing* block in Figure 2. Variations in thickness of the cell sections and amount of HES used, affects both shade and brightness of the tissue. These variations are clearly apparent in the example images, Figure 1. As these parameters are independent of the prognosis of the patient, their effect in this

analysis are minimized by converting all images to grayscale, and equalizing pixel values to span the range $[0, 1]$.

An automatic masking algorithm for identification of regions covered by tissue are implemented using Otsu's method [21]. To obtain the major trends in the image, the mask are dilated using disk shaped structure element of size 9. Holes in the mask with a size ≤ 4000 pixels are removed using a flood-fill operation, remaining regions in the mask with a size ≤ 4000 pixels are removed using a flood-fill operation on the inverted mask. The mask is finally eroded with the same 9 pixel disk that was used for dilation.

We use multi-resolution LBP including a variable size Gaussian low-pass filter, with 95% of its mass within a circle with radius $g_r = \pi R/P$, according to the method presented by Mäenpää et al [16]. The variable size Gaussian low-pass filter is applied to the image before computing LBP and VAR at each scale.

B. Feature extraction

The section describes the operations contained in the *test feature extraction* and *training feature extraction* blocks in Figure 2. For each scale R_r , a histogram describing all pixels identified as tissue is created for both operators. Models are defined for each prognosis using histograms computed from the training set at scale R_r . The mean histogram for prognosis ω is found by

$$\mu_{\omega} = \frac{1}{n} \sum_{i=1}^n h_{\omega,i} \quad (4)$$

where $h_{\omega,i}$ describes the histogram of image i with prognosis ω , and n the number of patients in the training set with the given prognosis. The mean histogram, μ_{ω} , is a $1 \times B$ vector,

where B is the number of bins in $h_{\omega,i}$. Variance for each bin in the histogram is found by

$$\sigma_{\omega}^2 = \frac{1}{n} \sum_{i=1}^n (h_{\omega,i} - \mu_{\omega})^2 \quad (5)$$

where n is the number of patients and \circ^2 the element-wise squared, and σ_{ω}^2 is a $1 \times B$ vector. Using Chi-square, the similarity between the image histogram and models are computed for each prognosis.

$$\chi_{k,\omega}^2 = \sum_{j=1}^B \frac{(h_k(j) - \mu_{\omega}(j))^2}{\sigma_{\omega}^2(j)} \quad (6)$$

where $h_k(j)$ describes bin j of the histogram of input image k , B the number of bins in h_k , $\mu_{\omega}(j)$ bin j of the mean histogram, and $\sigma_{\omega}^2(j)$ bin j of the variance of prognosis ω .

Multiple scales, R_r , are then combined using individual normalized weights $w \in [0, 1]$.

$$f = \sum_{r=1}^R w_r \chi_r^2 = \mathbf{w}^T \chi^2 \quad (7)$$

Where \mathbf{w}^T is a $1 \times R$ weight vector and χ^2 is a column vector of χ_r^2 . The weight $\mathbf{w} = [1, 0.89]$ was chosen through an exhaustive search using the chosen scales.

A combined vector, using both LBP and local variance in combination with all known prognoses, results in the feature vector.

$$f_i = \begin{bmatrix} \mathbf{w}^T \chi_{LBP, \omega_1}^2 \\ \mathbf{w}^T \chi_{LBP, \omega_2}^2 \\ \mathbf{w}^T \chi_{LBP, \omega_3}^2 \\ \mathbf{w}^T \chi_{Var, \omega_1}^2 \\ \mathbf{w}^T \chi_{Var, \omega_2}^2 \\ \mathbf{w}^T \chi_{Var, \omega_3}^2 \end{bmatrix} \quad (8)$$

Where prognoses are denoted ω_1 for *no recurrence*, ω_2 for *recurrence but not progression*, and ω_3 for *progression*.

C. Classifier

An apparent class imbalance exists when predicting either no-recurrence or recurrence of patients using the chosen data set. By expanding the data set as more patients are diagnosed, this imbalance will prevail as the various prognoses rates are different. To alleviate this imbalance, the RUSBoost classifier is chosen.

IV. EXPERIMENTS AND RESULTS

The algorithm for identifying tissue in the images were developed using one of the images in the data set. Truth masks were manually created by the authors of this paper, it should therefore be taken as an approximation of the actual truth masks. Validation of the masks on the other 41 images achieve a mean accuracy of 98.7%, sensitivity of 99.4% and specificity of 92.9%. Accuracy is defined as the proportion of correctly classified examples in all classes, sensitivity the proportion in the positive class and specificity the proportion in the negative class.

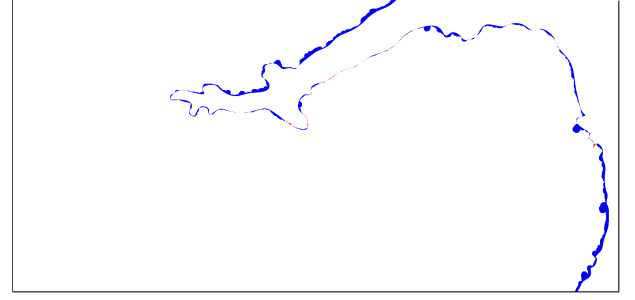


Fig. 3. Masking error using the automatic identification of tissue, histopathological image shown in Figure 1 (second row, middle column). Regions manually annotated as tissue, but not identified by the automatic masking is shown in blue. The opposite is shown in red.

Due to the size of the data set, classification were conducted using a leave-one-out cross-validation. Initial testing to determine system parameters were done using an exhaustive search, using the parameters $R = \{3, 5, 7, 9\}$ and $P = \{8, 16, 24\}$. A two-class problem, using histogram models based on all known prognoses to predict no-recurrence/recurrence was chosen. The optimal neighbourhood was found to be $R = \{3, 5\}$ $P = 8$ with the corresponding weights $w_1 = 1$, $w_2 = 0.89$. For experiments using only the LBP or VAR operator, a total of three features were used. In the experiment using both operators, a total of six features were used.

A wide range of weak learners (1-200) was tested to determine the optimal region. As the classifier uses random undersampling, the accuracy, sensitivity, and specificity were computed as the mean of 48 classifications for each number of weak learners. The highest accuracies were found when using 90 weak learners, and these are the results shown in table II. Performance of existing systems are shown in table III. Patients with recurrence are defined as the positive class, and patients without recurrence as the negative class.

TABLE II
SYSTEM PERFORMANCE OF PREDICTING RECURRENCE USING 90 WEAK LEARNERS. VAR AND LBP ARE COMPUTED USING $R=\{3,5\}$ AND $P=8$.

	Accuracy	Sensitivity	Specificity
LBP & VAR	72.08 ± 2.30	84.36 ± 2.30	44.68 ± 3.78
LBP	71.63 ± 2.71	83.98 ± 2.71	44.07 ± 4.12
VAR	63.94 ± 0.84	61.85 ± 0.84	68.59 ± 4.15

TABLE III
SYSTEM PERFORMANCE OF PREDICTING RECURRENCE IN EXISTING SYSTEMS.

	features	Accuracy	Sensitivity	Specificity
Borgi [8]	24	51.0	68.83%	38.53%
Spyridonos [7]	4	72.82%	72.3%	71.1%

V. DISCUSSION

The proposed algorithm for automatic identification of tissue in histopathological images perform very well, however in regions where a smooth transition occur between areas

with and without tissue, the algorithm struggle to identify the outmost cell layer.

Prognosis prediction using features computed using only the LBP operator achieves a sensitivity of 84.34% and a specificity of 44.68%, an increased identification rate of both outcomes compared to the system proposed by Borgi et al [8], table III. The overall system accuracy is similar to the result presented by Spyridonos [7], table III, but the proposed method in this paper does, however, not require nuclei segmentation to define features, thus reducing the complexity of the system. The proposed method also use fewer features. In addition, the higher identification rate of patients with recurrence achieved in this experiment are desired, as this will reduce the amount of patients receiving insufficient follow-up treatment. Prediction using features computed using the VAR operator, achieves an identification rate above 60% for each class. No apparent increase in performance is observed when using both feature descriptors, compared to using LBP alone. We assume this to be a consequence of the limited data set used in this analysis.

The proposed system does not achieve an identification rate high enough to replace traditional manual grading. However, we consider the system useful, as it can be used as an aid for histopathologists and clinicians in combinations with the traditional manual analysis, as the interobserver variability is removed.

VI. CONCLUSION

This study propose a system to aid in the prediction of recurrence of urothelial carcinoma with the use of a local texture analysis on histopathological images. The results suggest that both local binary pattern, and to some extent local variance, can be used to identify patients with recurrence without the need of segmenting cells. Local binary pattern achieve 84% correct identification of patients with recurrence, while the identification rate of patients without recurrence are approximately 45%. While local variance achieve above 60% correct identification for each class. The use of both operators combined does, however, not result in an improvement over LBP alone.

Due to the limited data set used in this analysis, future work is focused on reproducibility of the results using a larger data set. Additionally, the use of multiple sections from histopathological images to describe each patient are considered an interesting direction to increase the data available.

REFERENCES

- [1] J. Ferlay, I. Soerjomataram, M. Ervik, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. Parkin, D. Forman, and F. Bray, "Globocan 2012 v1.0, cancer incidence and mortality worldwide: IARC CancerBase No. 11 [internet]," Lyon, France: International Agency for Research on Cancer; 2013. Available from: <http://globocan.iarc.fr>, accessed on 28/04/2016.
- [2] O. Mangrud, R. Waalen, E. Gudlaugsson, I. Dalen, I. Tasdemir, E. Janssen, and J. Baak, "Reproducibility and prognostic value of WHO1973 and WHO2004 grading systems in TaT1 urothelial carcinoma of the urinary bladder," *PLoS ONE*, vol. 9, no. 1, pp. 1–8, 01 2014.
- [3] L. S. J. Borden, P. E. Clark, and M. C. Hall, "Bladder cancer," *Current Opinion in Oncology*, vol. May 2004 - Volume 16 - Issue 3, pp. 257–262, 2004.
- [4] J. Oosterhuis, R. Schapers, M. Janssen-Heijnen, R. Pauwels, D. Newling, and F. Ten Kate, "Histological grading of papillary urothelial carcinoma of the bladder: prognostic value of the 1998 WHO/ISUP classification system and comparison with conventional grading systems," *Journal of clinical pathology*, vol. 55, no. 12, pp. 900–905, 2002.
- [5] L. He, L. R. Long, S. Antani, and G. R. Thoma, "Histology image analysis for carcinoma detection and grading," *Computer Methods and Programs in Biomedicine*, vol. 107, no. 3, pp. 538–556, 2012.
- [6] J. N. Kather, C.-A. Weis, F. Bianconi, S. M. Melchers, L. R. Schad, T. Gaiser, A. Marx, and F. G. Zöllner, "Multi-class texture analysis in colorectal cancer histology," *Scientific Reports*, vol. 6, 2016.
- [7] P. Spyridonos, D. Glotsos, D. Cavouras, P. Ravazoula, and G. Nikiforidis, "A prognostic-classification system based on a probabilistic nn for predicting urine bladder cancer recurrence," in *Digital Signal Processing, 2002. DSP 2002. 2002 14th International Conference on*, vol. 2, 2002, pp. 1161–1164 vol.2.
- [8] A. Borgi, S. Ounallah, N. Stambouli, S. Selami, and A. B. A. Elgaaid, "Diagnostic system for predicting bladder cancer recurrence using association rules," in *SAI Intelligent Systems Conference (IntelliSys), 2015*, Nov 2015, pp. 93–99.
- [9] S. Morales, K. Engan, V. Naranjo, and A. Colomer, "Detection of diabetic retinopathy and age-related macular degeneration from fundus images through local binary patterns and random forests," in *Image Processing (ICIP), 2015 IEEE International Conference on*, Sept 2015, pp. 4838–4842.
- [10] J. Y. Choi, D. H. Kim, and Y. M. Ro, "Combining multiresolution local binary pattern texture analysis and variable selection strategy applied to computer-aided detection of breast masses on mammograms," in *Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics*, Jan 2012, pp. 495–498.
- [11] J. A. Khorsheed and K. Yurtkan, "Analysis of local binary patterns for face recognition under varying facial expressions," in *2016 24th Signal Processing and Communication Application Conference (SIU)*, May 2016, pp. 2085–2088.
- [12] M. Jasim and M. Hasanuzzaman, "Sign language interpretation using linear discriminant analysis and local binary patterns," in *Informatics, Electronics Vision (ICIEV), 2014 International Conference on*, May 2014, pp. 1–5.
- [13] O. Mangrud, "Identification of patients with high and low risk of progression of urothelial carcinoma of the urinary bladder stage Ta and T1," Ph.D. dissertation, University of Bergen, 2014.
- [14] T. Ojala, M. Pietikinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, no. 21, pp. 51–59, 1996.
- [15] T. Ojala, M. Pietikainen, and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul 2002.
- [16] T. Mäenpää and M. Pietikäinen, "Multi-scale binary patterns for texture analysis," in *Scandinavian Conference on Image Analysis*. Springer, 2003, pp. 885–892.
- [17] C. Seiffert, T. M. Khoshgoftaar, J. V. Hulse, and A. Napolitano, "Rusboost: A hybrid approach to alleviating class imbalance," *IEEE Trans. Syst., Man, Cybern. A*, vol. 40, no. 1, pp. 185–197, Jan 2010.
- [18] G. M. Weiss, "Mining with rarity: A unifying framework," *SIGKDD Explor. Newsl.*, vol. 6, no. 1, pp. 7–19, Jun. 2004.
- [19] Y. Freund and R. E. Schapire, "A desicion-theoretic generalization of on-line learning and an application to boosting," in *European conference on computational learning theory*. Springer, 1995, pp. 23–37.
- [20] N. V. Chawla, A. Lazarevic, L. O. Hall, and K. W. Bowyer, "Smoteboost: improving prediction of the minority class in boosting," in *In Proceedings of the Principles of Knowledge Discovery in Databases, PKDD-2003*, 2003, pp. 107–119.
- [21] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285–296, pp. 23–27, 1975.