

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/258814002>

Down Syndrome Detection from Facial Photographs using Machine Learning Techniques

Article in *Proceedings of SPIE - The International Society for Optical Engineering* · February 2013

DOI: 10.1117/12.2007267

CITATIONS

22

READS

3,743

6 authors, including:



[Qian Zhao](#)

Children's National Medical Center

18 PUBLICATIONS 440 CITATIONS

[SEE PROFILE](#)



[Dina Zand](#)

FDA

52 PUBLICATIONS 1,725 CITATIONS

[SEE PROFILE](#)



[Marshall L Summar](#)

Children's National Medical Center

243 PUBLICATIONS 8,096 CITATIONS

[SEE PROFILE](#)

Down Syndrome Detection from Facial Photographs using Machine Learning Techniques

Qian Zhao^{*a}, Kenneth Rosenbaum^b, Raymond Sze^{a,c}, Dina Zand^b, Marshall Summar^b,
Marius George Linguraru^a

^aSheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Medical Center,
Washington DC, USA 20010;

^bDivision of Genetics and Metabolism, Children's National Medical Center, Washington DC, USA
20010;

^cDivision of Diagnostic Imaging and Radiology, Children's National Medical Center, Washington
DC, USA 20010

*qzhao@cnmc.org; phone 1 202 476-1285; fax 1 202 476-1270

ABSTRACT

Down syndrome is the most commonly occurring chromosomal condition; one in every 691 babies in United States is born with it. Patients with Down syndrome have an increased risk for heart defects, respiratory and hearing problems and the early detection of the syndrome is fundamental for managing the disease. Clinically, facial appearance is an important indicator in diagnosing Down syndrome and it paves the way for computer-aided diagnosis based on facial image analysis. In this study, we propose a novel method to detect Down syndrome using photography for computer-assisted image-based facial dysmorphology. Geometric features based on facial anatomical landmarks, local texture features based on the Contourlet transform and local binary pattern are investigated to represent facial characteristics. Then a support vector machine classifier is used to discriminate normal and abnormal cases; accuracy, precision and recall are used to evaluate the method. The comparison among the geometric, local texture and combined features was performed using the leave-one-out validation. Our method achieved 97.92% accuracy with high precision and recall for the combined features; the detection results were higher than using only geometric or texture features. The promising results indicate that our method has the potential for automated assessment for Down syndrome from simple, noninvasive imaging data.

Keywords: Down syndrome, computer-aided diagnosis, facial image analysis, feature extraction, classification, machine learning.

1. INTRODUCTION

Many congenital syndromes are caused by chromosome abnormality and Down syndrome is one of the most common. Each year, one in every 1,100 babies is born with Down syndrome worldwide; the incidence is even higher in the United States [1, 2]. Since 1979, the number of children born with Down syndrome has seen a steady increase [3]. Down syndrome is caused by an extra copy of chromosome 21 and has a high incidence of related medical complications, such as heart defects, respiratory and hearing problems. The early diagnosis may provide the best clinical management of children with Down syndrome for lifelong medical care that may involve physical and speech therapists, cardiologists, endocrinologists and neurologists.

Down syndrome may be diagnosed either during pregnancy or shortly after birth. Diagnostic and screening testing can be done during pregnancy. At birth, Down syndrome may be diagnosed by the presence of some typical facial appearances and physical characteristics. Some of the features include a small and flattened nose, small ears and mouth, upward slanting eyes, and protruding tongue. To confirm a diagnosis, a chromosome test called a karyotype can be performed. However, these tests are expensive and time-consuming and many remote healthcare centers do not have ready access to these technologies.

Computer-aided diagnosis (CAD) system has the potential to play an important role in clinical genetics. The distinctive facial features of Down syndrome provide an opportunity for automated detection. So far, only a few attempts have been undertaken to detect syndromes using two-dimensional or three-dimensional face recognition and machine learning techniques. One recent study [4] investigated Gabor wavelet features to detect 14 syndromes. The proposed model graph consisted of manually labeled node coordinates and wavelet coefficients associated with each node. The classification results using multinomial logistic regression achieved 21% accuracy for testing data and 60% accuracy for training data. The study is based on previous work by the same authors[5-7]. However, their dataset does not contain a healthy population. Their method performed classification in groups and in pairs instead of separating syndrome cases from the normal ones. For Down syndrome detection, the authors in [8] also used Gabor wavelet transform as a feature extraction method. Principle component analysis and linear discriminant analysis reduced the feature dimension. Classification was implemented with k-nearest neighbor (kNN) and support vector machine (SVM). They reported that the accuracy for kNN and SVM were 96% and 97.3%, respectively. But their method needs manual image standardization, such as rotation and cropping. Moreover, the dataset only consists of 15 Down syndrome cases and 15 normal cases. In [9], the authors used local binary pattern (LBP) as facial features to distinguish Down syndrome from healthy group. The classification was performed by template matching. However, their features are not extracted locally around clinically relevant landmarks and their method also requires manual pre-processing. Finally, for fetal alcohol syndrome (FAS) detection, the authors in [10, 11] summarized the methods for FAS detection using distance measurements derived from stereo-photogrammetry and facial surface imaging. Their method needs special calibration frames for photography acquisition.

In this study, we investigate a novel method to detect Down syndrome using only facial photographs. To capture the typical facial characteristics of Down syndrome, we define seventeen anatomical landmarks and extract geometric features from them. Local texture features around each landmark are also extracted based on the Contourlet transform and local binary patterns. We concatenate geometric and texture features as combined features to fuse more information. A SVM is adopted as classifier to discriminate between Down syndrome and normal cases. To our knowledge, landmark geometric features and local texture features are combined for the first time to represent the facial characteristic using machine learning techniques. From the clinical point of view, we also identify the relevant facial characteristic for the clinical diagnosis of Down syndrome. Leave-one-out validation is performed to evaluate the method in terms of accuracy, precision, and recall. When fully automated, our method will provide an accurate and simple clinical tool to detect syndromes with great potential to support remote diagnosis of genetic diseases in areas without access to a genetics clinic.

2. METHODS

2.1 Data

Two datasets were collected to verify the method. The mixed dataset contains 48 photographs of pediatric patients, including 24 normal cases and 24 abnormal cases with mixed genetic syndromes. The Down syndrome dataset contains 24 photographs for Down syndrome and 24 photographs for a healthy group. Images were acquired using basic photo cameras and consisted of frontal photograph of the head of the patients. Then images were resized around the patients' faces to 256 by 256 pixels. The images have various background, illumination, poses and expression.

2.2 Geometric features

We defined manually 17 clinically relevant facial anatomical landmarks. They covered most of the inner facial features, such as eyes, nose, and mouth. Nineteen pairwise landmark distances were extracted as the geometric features which represent the key clues for syndrome diagnosis, e.g. short palpebral fissures, midfacial hypoplasia, smooth philtrum. The geometric features are divided to horizontal and vertical distances. The horizontal distances were normalized by the distance between eyes (distance between landmark pair (1, 4)), and the vertical distances were normalized by the distance from the middle point between eyebrows (landmark 5) to the lower point of mouth (landmark 17). The anatomical landmarks and the geometric features are shown in Figure 1 (a).

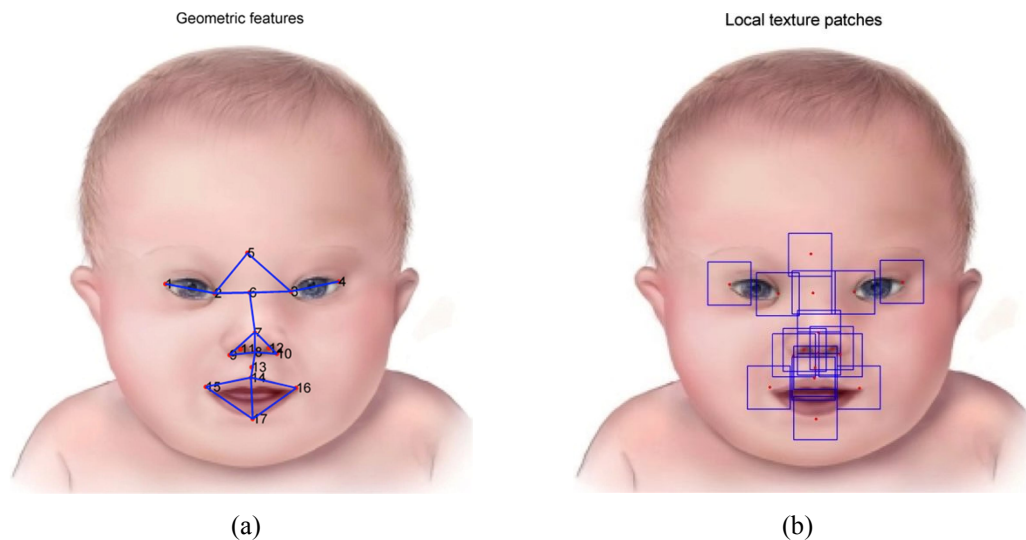


Figure 1. (a) Anatomical landmarks (red points) used in the analysis of facial dysmorphology and geometric features (blue lines) derived from the landmarks; (b) Patches around landmarks for local texture feature extraction.

2.3 Local texture features

To retain the spatial information, texture features were extracted locally for each anatomical landmark on square patches. For pre-processing, the image were first smoothed by a Gaussian filter and converted to gray level. Then a three-level Contourlet transform [12-14] was applied for image multi-scale representation. Contourlets not only maintain the main features of wavelets, but also offer a high degree of directionality and anisotropy.

The Contourlet transform is derived from the discrete domain and can capture smooth contours and edges in any direction. The Contourlet transform is implemented via a two-dimensional filter bank comprising a two-channel quincunx filter bank with fan filters and a shearing operator that decomposes an image into several directional sub-bands at multiple scales. This is accomplished by combining the Laplacian pyramid with a directional filter bank at each scale, shown in Figure 2 (b). As the middle frequency component of the image represents texture better, we reconstructed the image by removing the low frequency component, shown in Figure 2 (a).

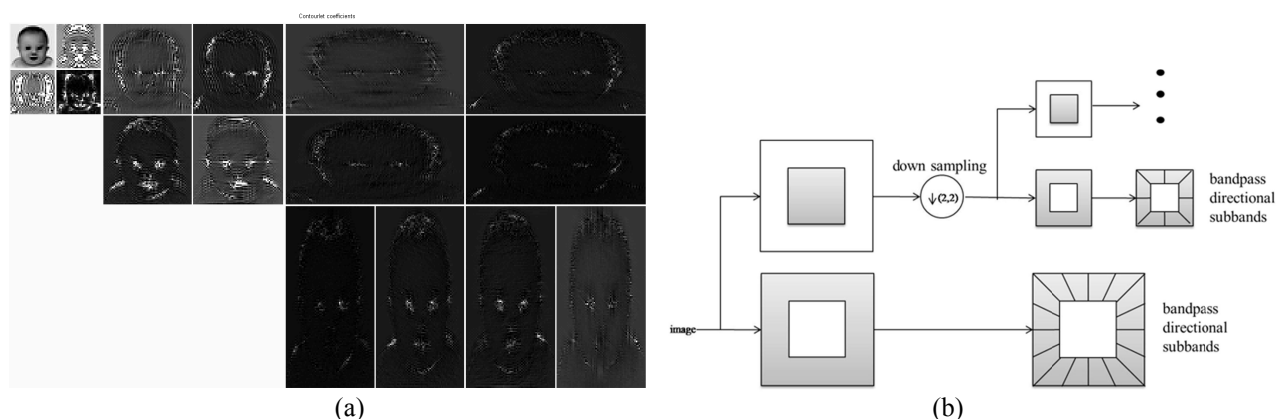


Figure 2. (a) Three-level Contourlet transform. The resolution for the 1st level is 32 by 32, for the 2nd level is 64 by 64, and for the 3rd level is 64 by 128 and 128 by 64; (b) The Contourlet filter banks: first a multiscale decomposition into octave bands by the Laplacian pyramid is computed, and then a directional filter bank is applied to each bandpass channel.

Finally, the local binary pattern (LBP) operator [15] was applied to the reconstructed image and six first-order statistical measurements were extracted from the LBP histogram: the mean, variance, skewness, kurtosis, entropy, and energy. The LBP features have been proven robust against illumination changes and fast to compute. Moreover, they can robustly discern micro-structures, which make them appropriate to our application.

The original LBP method, shown in Figure 3, was first introduced by Ojala et al. [16] as a complementary measure for local image contrast. It operates with eight neighboring pixels using the central pixel as a threshold. The LBP code is then computed by multiplying the threshold values by weights given by powers of two and adding the results in the way described in Figure 3. The LBP of a 3×3 neighborhood produces up to $2^8=256$ local texture patterns, and the 256-bin occurrence LBP histogram computed over a region is then employed for texture description.

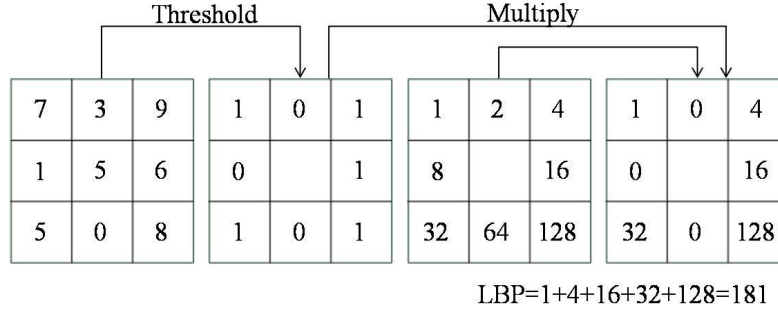


Figure 3. The original LBP operator.

The original LBP was extended to a more general approach, the uniform LBP [16], in which the number of neighboring sample points is not limited:

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

where g_p corresponds to the grey values of P equally spaced pixels on a circle of radius R , g_c is the grey value of the central pixel, and $U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^P |s(g_p - g_c) - s(g_{p-1} - g_c)|$.

The geometric and local texture features have 19 and 102 dimensions, respectively. The 121-dimensional combined features were generated by concatenating the geometric and local texture features. The features were ranked based on the area under the receiver operating characteristic (ROC) curve and the random classifier slope. The optimal dimension for each type of features was found by empirical exhaustive search.

2.4 Classification

A support vector machine (SVM) [17] with radial basis function kernel classifier was employed to analyze the extracted features. SVM is a powerful and robust classifier which can deal with high dimensional data well. The decision function of SVM is

$$f(x) = \langle w, \phi(x) \rangle + b, \quad (2)$$

where $\phi(x)$ is a mapping of sample x from the input space to a high-dimensional space, w is the normal vector to the hyperplane and b represents the offset of the hyperplane from the origin. The optimal values of w and b can be solved by an optimization problem:

$$\text{minimize: } g(w, \xi) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i \quad (3)$$

$$\text{subject to: } y_i (\langle \mathbf{w}, \phi(x_i) \rangle + b) \geq 1 - \xi_i, \xi_i \geq 0, \quad (4)$$

where ξ_i is the i^{th} slack variable and C is the regularization parameter which can be chosen optimally by grid search.

2.5 Validation

Leave-one-out validation was performed for the two datasets separately. We compared the performance of the method using only geometric features, only local texture features and combined features (both geometric and texture), respectively. The results were assessed by accuracy, precision and recall. The area under curve (AUC) of the ROC was also computed for each dataset.

$$\text{Accuracy} = \frac{\text{Correct Classification}}{\text{Total Number of Cases}} = \frac{tp + tn}{tp + tn + fp + fn}, \quad (5)$$

$$\text{Precision} = \frac{\text{Correct Positive Classifications}}{\text{Positive Classifications}} = \frac{tp}{tp + fp}, \quad (6)$$

$$\text{Recall} = \frac{\text{Correct Positive Classifications}}{\text{Number of Positives}} = \frac{tp}{tp + fn}, \quad (7)$$

where tp , tn , fp , and fn correspond to true positive, true negative, false positive, and false negative, respectively.

3. RESULTS

3.1 Mixed dataset

For the mixed dataset, the optimal dimension is 3, 35, 5 for geometric, texture, and combined features, respectively. The results for the mixed dataset are shown in Table 1. The average accuracy for geometric features and local texture features are 77.1% and 75.0%, respectively. The combined features outperform them with 81.3% accuracy. The results show an improvement when combining the landmark spatial information and texture features. Similarly, the precision and recall of the combined features reached 82.6% and 79.2%, respectively. Figure. 4(a) shows the ROC for the detection of genetic syndromes from the mixed dataset. The combined features achieved the largest AUC of 0.837.

Table 1. Comparison of performance among geometric, local texture and combined (geometric and texture) features for the mixed dataset. The accuracy, precision and recall are the average values.

Features	Accuracy	Precision	Recall	AUC
Geometric	0.771	0.842	0.667	0.766
Texture	0.750	0.731	0.792	0.818
Combined	0.813	0.826	0.792	0.837

3.2 Down syndrome dataset

The same experiments were performed on the Down syndrome dataset. For Down syndrome dataset, the optimal dimension for geometric, texture and combined features is 7, 20, and 6, respectively. The results are shown in Table 2. The average accuracy for the combined features is 97.9% which is higher than geometric and texture features only. The precision and recall for the combined features are 100% and 95.8%, respectively. The ROC curve for the detection of Down syndrome is shown in Figure 4(b). The largest AUC is 0.996 for the combined features. From the feature ranking, the top ranked geometric features include the length of nose, the wide of opened mouth, and the distance between eyes. It conforms to the clinical findings.

Table 2. Comparison of performance among geometric, local texture and combined (geometric and texture) features for the Down syndrome dataset. The accuracy, precision and recall are the average values.

Features	Accuracy	Precision	Recall	AUC
Geometric	0.958	0.958	0.958	0.965
Texture	0.771	0.783	0.750	0.854
Combined	0.979	1.000	0.958	0.996

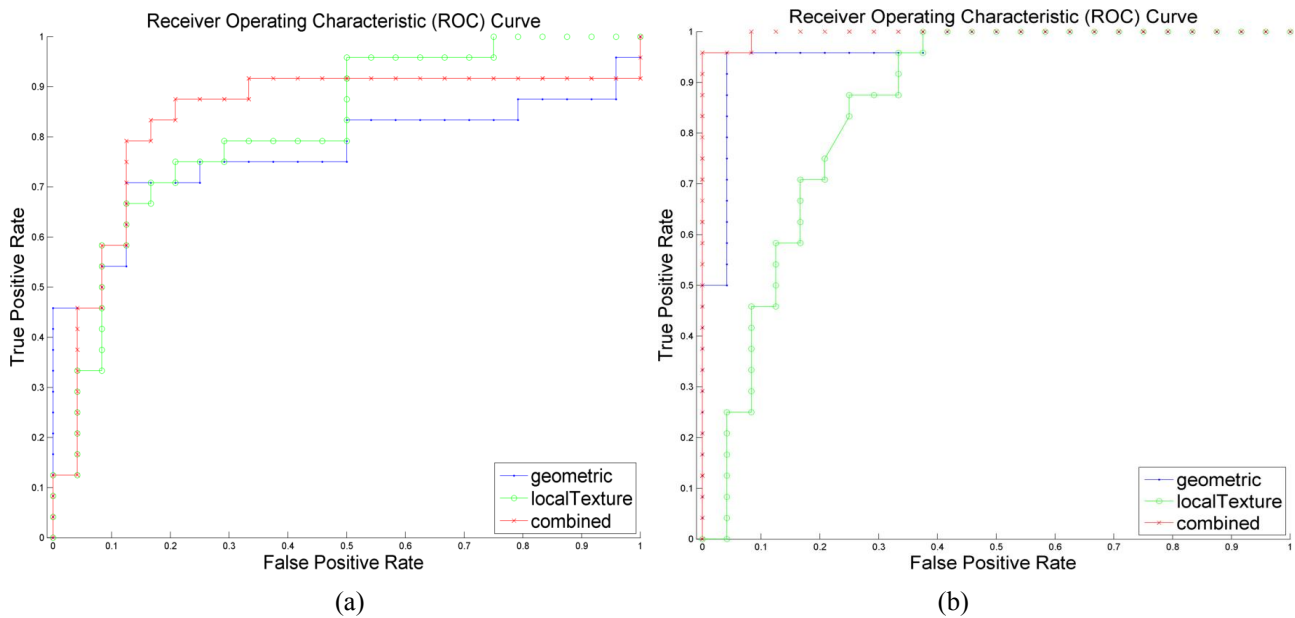


Figure. 4 The ROC curves for (a) the mixed dataset and (b) the Down syndrome dataset.

4. CONCLUSION

A method for the automated detection of Down syndrome using digital facial images was proposed. Geometric features based on anatomical landmarks and local texture features based on Contourlet transform and local binary patterns were combined. After feature selection, a support vector machine classifier was employed to discriminate between the Down syndrome and normal cases. Comparisons among the geometric, texture and combined features were performed by leave-one-out validation. The experimental results showed that the performance of the method has improved when combined features of both facial geometry and texture were used; Down syndrome cases were detected with 97.9% accuracy. The encouraging results also demonstrated that our method has great potential to support the computer-aided diagnosis for Down syndrome from simple photographic data. Data collection is our on-going work for the more comprehensive validation of our method. In future work we will investigate the automated placement of anatomical landmarks.

5. ACKNOWLEDGEMENT

This work is supported in part by the intramural funding of the Sheikh Zayed Institute for Pediatric Surgical Innovation at Children's National Medical Center.

REFERENCES

- [1] G. de Graaf, M. Haveman, R. Hochstenbach *et al.*, "Changes in yearly birth prevalence rates of children with Down syndrome in the period 1986–2007 in the Netherlands," *Journal of Intellectual Disability Research*, 55(5), 462-473 (2011).
- [2] F. K. Wiseman, K. A. Alford, V. L. J. Tybulewicz *et al.*, "Down syndrome—recent progress and future prospects," *Human Molecular Genetics*, 18(R1), R75-R83 (2009).
- [3] [Down Syndrome Cases at Birth Increased].
- [4] S. Boehringer, M. Guenther, S. Sinigerova *et al.*, "Automated syndrome detection in a set of clinical facial photographs," *American Journal of Medical Genetics Part A*, 155(9), 2161-2169 (2011).
- [5] S. Boehringer, T. Vollmar, C. Tasse *et al.*, "Syndrome identification based on 2D analysis software," *Eur J Hum Genet*, 14(10), 1082-1089 (2006).
- [6] H. S. Loos, D. Wiczorek, R. P. Wurtz *et al.*, "Computer-based recognition of dysmorphic faces," *European Journal of Human Genetics*, 11(8), 555-560 (2003).
- [7] T. Vollmar, B. Maus, R. P. Wurtz *et al.*, "Impact of geometry and viewing angle on classification accuracy of 2D based analysis of dysmorphic faces," *European Journal of Medical Genetics*, 51(1), 44-53 (2008).
- [8] Ş. Saraydemir, N. Taşpınar, O. Eroğul *et al.*, "Down Syndrome Diagnosis Based on Gabor Wavelet Transform," *Journal of Medical Systems*, 36(5), 3205-3213 (2012).
- [9] K. Burçin, and N. V. Vasif, "Down syndrome recognition using local binary patterns and statistical evaluation of the system," *Expert Systems with Applications*, 38(7), 8690-8695 (2011).
- [10] T. S. Douglas, and T. E. M. Mutsvangwa, "A review of facial image analysis for delineation of the facial phenotype associated with fetal alcohol syndrome," *American Journal of Medical Genetics Part A*, 152A(2), 528-536 (2010).
- [11] T. E. M. Mutsvangwa, E. M. Meintjes, D. L. Viljoen *et al.*, "Morphometric analysis and classification of the facial phenotype associated with fetal alcohol syndrome in 5- and 12-year-old children," *American Journal of Medical Genetics Part A*, 152A(1), 32-41 (2010).
- [12] M. N. Do, and M. Vetterli, "Contourlets: a directional multiresolution image representation." 1, I-357-I-360 vol.1.
- [13] M. N. Do, and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, 14(12), 2091-2106 (2005).
- [14] M. N. Do, and M. Vetterli, "Contourlets: a new directional multiresolution image representation." 1, 497-501 vol.1.
- [15] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, 29(1), 51-59 (1996).
- [16] T. Ojala, M. Pietikäinen, and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7), 971-987 (2002).
- [17] C. Cortes, and V. Vapnik, "Support-vector networks," *Machine Learning*, 20(3), 273-297 (1995).