

# Identifying Down Syndrome Cases by Combined Use of Face Recognition Methods

Vlad Dima<sup>1</sup>, Anca Ignat<sup>1(✉)</sup>, and Cristina Rusu<sup>2</sup>

<sup>1</sup> Faculty of Computer Science, “Alexandru Ioan Cuza” University of Iași,  
str. General Berthelot nr. 16, 700483 Iasi, Romania  
{vlad.dima,ancai}@info.uaic.ro

<sup>2</sup> “Grigore T. Popa” University of Medicine and Pharmacy,  
str. Universității nr. 16, 700115 Iasi, Romania  
abcrusu@yahoo.com

**Abstract.** Down syndrome is a chromosomal disorder, people affected by this disease having very specific facial characteristics. In this paper we adapt some well-known face recognition methods (Local Binary Patterns, Eigenfaces) and test their ability to distinguish between a Down syndrome face and a normal one in digital images. For classification  $k$ NN ( $k$  Nearest Neighbor) and Support Vector Machines (SVM) with polynomial kernel of third degree and Radial Basis Function (RBF) are employed. We test the methods using FERET, CAS-PEAL, LFW, AT&T for normal face images and a collection of Down faces gathered from the Internet. Accuracy, precision, recall and specificity are computed in order to evaluate the classification results.

**Keywords:** Down syndrome · PCA · Eigenfaces · Local Binary Patterns · Discrete Wavelet Transform ·  $k$ NN · SVM

## 1 Introduction

Down syndrome is the most common chromosomal abnormality defined by typical facial appearance, short fingers/toes and moderate/severe intellectual disability. These features are frequently associated with cardio-vascular and gastro-intestinal defects. The most characteristic facial traits include: round and flat face, upslanting palpebral features (mongoloid eyes), epichantal fold (skin fold in the inner eyelid angle), Brushfield spots (spotted iris), small nose with flattened root, small mouth (usually open, with protruding tongue) and small, low set ears [1, 2].

Down syndrome identification in digital images of faces can be useful in computer-aided diagnosis, human–computer interaction systems or demographic studies.

Down syndrome persons have very distinctive facial traits, that is why it is easy for humans to distinguish between a Down syndrome person and a healthy person just by looking at their faces. In this paper we investigate how effective are some computer vision techniques for face recognition, to make the same decision: Down/non Down face.

The subject of face recognition is one of the major research direction in image processing and computer vision. Some of the main directions of research in this field are described in [3, 4]. The problem of Down syndrome detection was also approached. In [5] the authors studied the problem of automatic identification of other (than Down) genetically disordered patients. In [6, 7] they use for Down syndrome identification a Hierarchical Constrained Local Model with Independent Component Analysis first for detecting facial landmarks and then to extract features around these points Local Binary Patterns (LBP) are employed. They used SVMs (both linear and with Radial Basis Function kernel),  $k$ NN, random forest and Linear Discriminant Analysis as classifiers. The tests were made using images of young children (50 Down/50 healthy) of different ethnicity, taken under variable illumination, but in a controlled environment. The best result was an accuracy of 95.6% with the SVM-RBF. In [8] they also extract information about fiducial face points and then use contourlets and LBP to extract also textural features around the landmarks. The tests were made using children images, making two type of evaluations: one with 24 normal patients and 24 with mixed genetic syndromes (81.3% accuracy in detecting Down syndrome) and the second with 24 normal faces and 24 Down faces (97.9% accuracy). In [9] the authors worked with 1–12 years old children images (15 healthy, 15 with Down syndrome), performing a standardization step before extracting features using Gabor wavelets and employed  $k$ NN and linear SVM's for classification, the best result being an accuracy of 97.34%.

In our approach we used “real-life” images of persons with Down syndrome, of different ages and tested them using normal faces of mixed ethnicity, or using only Chinese. We tested LBP, a mixture of Discrete Wavelet Transform (DWT) with LBP, and Eigenfaces with Principal Components projection features using  $k$ NN, polynomial and RBF SVMs. For Down syndrome we selected 50 images from the Internet and for normal faces we used classical face recognition datasets (FERET [10], CAS-PEAL [11], LFW [12]). The influence of noise (inherent to real-life images) was also studied.

## 2 Face Characterization Features

There are many ways to extract features for face recognition. Among them are the Eigenfaces method [13] and Local Binary Patterns (LBP) [14, 15]. In the Eigenfaces algorithm, first the face images are vectorized and the mean value of these images is subtracted. A matrix is formed with these vectors as columns. Then the eigenvalues and eigenvectors of the covariance matrix are computed. The eigenvectors are arranged by sorting the corresponding eigenvalues in descending order. The eigenvectors are then employed to project new images on the linear subspace generated by them, thus forming a feature vector.

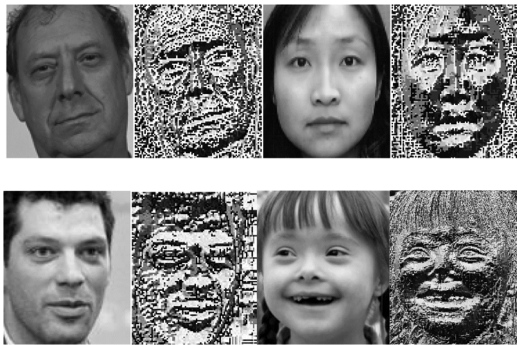
Local Binary Patterns [14] is a technique initially developed for texture characterization but can also be applied for face recognition [15]. A binary pattern for each pixel in a grayscale image is computed by comparing the intensities of its 8 neighbors with the intensity of the considered pixel. If the neighbor's intensity is bigger, a 1 is put in the binary pattern associated with the pixel, otherwise a 0. Thus, for each pixel an 8-bit pattern is computed and usually is converted to a 0–255 decimal value. One obtains a

matrix of binary patterns associated with the original image. In order to extract LBP features, the image is divided in blocks and histograms of the binary patterns are computed for each subimage. The feature vector is formed by concatenating these histograms. It was observed that among the 256 binary patterns that can be associated to a pixel, those that have only at most two 0–1 or 1–0 transitions are more important. These patterns are called uniform and there are 58 such patterns. If one computes histograms counting only the uniform patterns and the non-uniform ones are all counted in one bin, the size of histograms reduces from 256 to 59 without losing precisions in results, and thus the feature vectors are 4 times shorter.

The Wavelet Transform [16] was a solution to the shortcoming of the fact that the Fourier Transform coefficients provided no local information on frequencies. In 1988 Mallat and Meyer introduced the wavelet multiresolution concept, then Daubechies gave an algorithm for computing orthogonal wavelets with compact support, and Mallat deduced the fast wavelet transform, all these steps made the wavelet transform a very useful tool for signal processing. At one level of Discrete Wavelet (DW) decomposition, 4 filtered images are computed, half the size of the processed image: HH (high-high filtering both rows and columns of the image), HL (highpass filter is applied to rows and a lowpass one to columns), LH (low-high filtering rows/columns) and LL (lowpass filtering both rows and columns).

We use these methods to build feature vectors for Down syndrome identification in the following ways:

1. The uniform binary patterns are computed for an image, then the image is divided in  $4 \times 4$  or  $8 \times 8$  blocks, for each block the histogram of the uniform pattern is computed. The feature vector is formed by concatenating these 16 or 64 histograms. In Fig. 1 are some examples of images and their uniform binary pattern correspondent.



**Fig. 1.** Examples of face images from FERET, CAS-PEAL, LFW, and with Down syndrome and their corresponding uniform LBP images

2. First apply a 2 levels DW decomposition, for the LL subimage from the second level of wavelet decomposition we build the feature vector as described in 1 (using  $8 \times 8$  blocks).

3. We first computed the eigenfaces for the AT&T ([17], Fig. 2) face database, then we projected the image on the most significant of these eigenfaces. The weights obtained from the projection step form the feature vector.



**Fig. 2.** Eigenfaces from AT&T dataset

4. For the AT&T faces, the  $8 \times 8$  uniform LBP feature vectors are computed; for these feature vectors the principal components are calculated; for an image to be classified the final feature vector is obtained by projecting the  $8 \times 8$  uniform LBP feature vector of the image on some of the most important principal components.

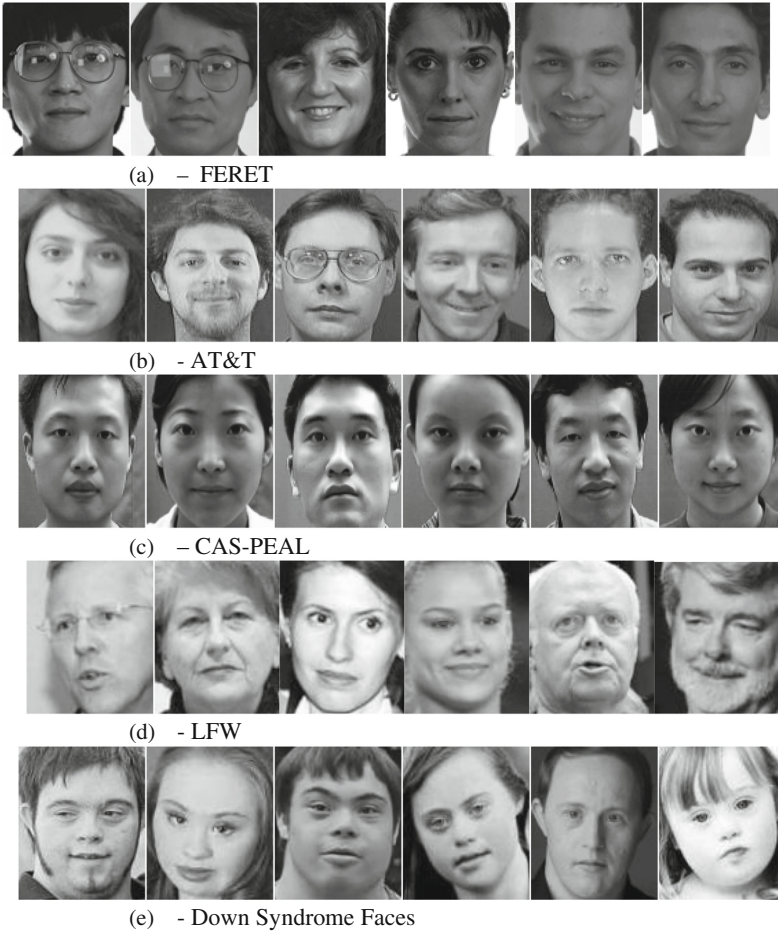
All the feature vectors were normalized: for the LBP ones, the histogram for each block was divided by the city-block norm, and for the projection weights we used the Euclidean norm for normalization.

The LBP method is known to be very sensitive to noise, that is the reason of considering applying first a DWT before extracting features with LBP method or projecting the face images on a set of eigenfaces or principal components.

As classifiers, the Support Vector Machines (SVM) with 3-rd degree polynomial kernel, or Radial Basis Function (RBF) kernel and  $k$ NN with  $k = 3$  and Euclidean norm, were employed.

For testing the feature vectors, face images from FERET [10, 18], CAS-PEAL [11], and LFW [12] were employed. We selected randomly, from the Internet 50 images of faces for persons with Down syndrome.

In Fig. 3 are some examples of images we used in the tests.



**Fig. 3.** Samples of face images from FERET, AT&T, CAS- PEAL, LFW, Down syndrome datasets

We evaluated the classification results using the following measures:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$Specificity = \frac{TN}{TN + FP} \quad (4)$$

where TP, TN, FP, FN stands for “true positive”, “true negative”, “false positive” “false negative”. For example, “true positive” is the number of Down faces identified as having Down syndrome, “false positive” the number of non-Down faces identified by the classifier as having Down syndrome. The accuracy is a measure of the overall success of the classification process, it is the proportion of the correctly identified images from all images. The recall (also called sensitivity) and specificity measure the proportion of the correctly identified Down/non-Down images from all Down/non-Down images, respectively. The precision is the ratio of the correctly identified Down images from all those images identified as Down.

### 3 Results

Our numerical tests were performed using MATLAB and OpenCV in Java. The face databases that we worked with have the following characteristics. The AT&T [17] database has 400 images of size  $92 \times 112$ , for 40 distinct persons, for each person 10 images with different facial expressions, under different light conditions and with or without glasses. The FERET (Face REcognition Technology, [10, 18]) have 14126 images for 1199 persons, with different facial expressions and positions, mixed ethnicity, taken in controlled environment. The size of images in FERET is  $256 \times 384$ . The CAS-PEAL database is a dataset of Chinese persons consisting of 30900 images for 1040 subjects, taken under varying pose, expression, accessory and lighting (PEAL) conditions, the size of images is  $360 \times 480$ . The LFW (Labeled Faces in the Wild) is a collection of 13000 face images collected from the Internet (such that Viola and Jones [19] face detector worked on them) for 5479 individuals, the size of the images is  $250 \times 250$ . Our criterion for selecting the images (Down or non-Down) was also the Viola and Jones face detector working on the images.

All the images from the AT&T dataset were employed for building 10304 ( $92 \times 112$ ) eigenfaces or 3776 principal components. In the projection computations we only used the most relevant 1500 of these eigenfaces and 125 principal components. From FERET database 900 frontal images were selected, from CAS-PEAL 300 (frontal and normal) and from LFW 70. The Viola and Jones face detector was applied to all the selected images and then feature vectors were computed. Because of the eigenfaces projection method 3. described in the previous section, all the faces were resized to  $92 \times 112$ , the common size of the AT&T images.

We performed several tests on these databases. The tests were carried out by using 60 (for FERET or CAS-PEAL) or 70 (for LFW) non-Down images and the 50 Down faces. Thus, for FERET 15 tests were made, for CAS-PEAL, 5 tests and 1 test for LFW images. As cross validation we used “leave-one-out”. In the tables with results, LBP ( $4 \times 4$  or  $8 \times 8$ ) means that the uniform binary patterns were used to extract features for blocks of size  $4 \times 4$  or  $8 \times 8$ , respectively. EP stands for Eigenfaces Projection method. LBP(PCA) is the  $8 \times 8$  uniform LBP feature extractor followed by the projection on the principal components on AT&T as described in Sect. 2, reducing the dimension of the feature vectors to 125. For the DWT we used 22 tap biorthogonal Daubechies filters.

The feature vectors have the following dimensions: 944 for LBP ( $4 \times 4$ ), 3776 for LBP ( $8 \times 8$ ) and DWT + LBP, 1500 for EP, 125 for PCA (Principal Component Analysis) reduced dimension LBP ( $8 \times 8$ ).

In Tables 1, 2 and 3 are the (mean) results for the (1)–(4) measures of classification. We expected low values for the Recall parameter due to the fact that the selected Down faces images are low quality and are not standardized in any way. The good classification results we get for FERET and CAS-PEAL are due to this difference in image quality. The images in FERET and CAS-PEAL were taken in controlled conditions, with the same camera. The 50 Down selected images are rather noisy, and had different sizes, the re-dimensioning process enhanced the ‘defects’ in these images. We suspect that in the case of CAS-PEAL classification, in fact a Chinese/non-Chinese identification process was performed rather than Down/non-Down one. The results we obtain for the LFW database give better information on the capacity of distinguishing between Down and non-Down face images of the considered features, because these images were selected in the same way.

**Table 1.** Mean values for Accuracy (A), Precision (P), Recall (R), Specificity (S) for 15 tests with 60 FERET faces and 50 Down faces

	SVM (polynomial, 3-rd degree)				3NN			
	A	P	R	S	A	P	R	S
LBP ( $4 \times 4$ )	98.18%	99.3% <sup>1</sup>	96.66%	99.44%	83.45%	99.24%	64.13%	99.55%
DWT + LBP	<b>98.72%</b>	99.33%	<b>97.86%</b>	99.44%	71.33%	<b>100%</b>	36.93%	<b>100%</b>
EP	96.84%	98.22%	94.8%	98.55%	<b>94.66%</b>	98.27%	<b>89.86%</b>	98.66%
LBP ( $8 \times 8$ )	95.65%	<b>99.49%</b>	90.94%	<b>99.58%</b>	91.51%	99.56%	81.76%	99.56%
LBP (PCA)	98.54%	99.19%	97.6%	99.33%	77.34%	<b>100%</b>	50.47%	<b>100%</b>

**Table 2.** Mean values for Accuracy (A), Precision (P), Recall (R), Specificity (S) for 5 tests with 60 CAS-PEAL faces and 50 Down faces

	SVM (polynomial, 3-rd degree)				3NN			
	A	P	R	S	A	P	R	S
LBP ( $4 \times 4$ )	<b>100%</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>	85.81%	<b>100%</b>	68.8%	<b>100%</b>
DWT + LBP	99.63%	<b>100%</b>	99.2%	<b>100%</b>	90.72%	<b>100%</b>	79.6%	<b>100%</b>
EP	97.63%	98.74%	96%	99%	<b>93.27%</b>	97.77%	<b>87.2%</b>	98.33%
LBP ( $8 \times 8$ )	98.72%	99.60%	97.6%	99.66%	89.45%	<b>100%</b>	76.8%	<b>100%</b>
LBP (PCA)	99.63%	99.60%	99.60%	99.66%	84.90%	<b>100%</b>	66.8%	<b>100%</b>

**Table 3.** Accuracy (A), Precision (P), Recall (R), Specificity (S) for 70 LFW faces and 50 Down faces

	SVM (polynomial, 3-rd degree)				3NN			
	A	P	R	S	A	P	R	S
LBP (4 × 4)	90.83%	95.34%	82%	97.14%	76.66%	92.3%	48%	97.14%
DWT + LBP	<b>93.33%</b>	<b>97.72%</b>	<b>86%</b>	<b>98.57%</b>	72.5%	<b>100%</b>	34%	<b>100%</b>
EP	90.83%	91.48%	<b>86%</b>	94.28%	80.83%	75.47%	<b>80%</b>	81.42%
LBP (8 × 8)	89.16%	97.43%	76%	<b>98.57%</b>	<b>82.5%</b>	93.93%	62%	97.14%
LBP (PCA)	91.66%	97.61%	82%	<b>98.57%</b>	77.5%	<b>100%</b>	46%	<b>100%</b>

We studied the effect of noise on the considered features by performing the following test: we trained the classifier using the original images and the same images where Gaussian noise with zero mean and 0.0005 variance was added. We tested using images with added Gaussian noise with zero mean and 0.003 variance. The tests were made in a “leave-one-out” manner, by training the classifier without the corresponding image we test. We only tested the FERET and CAS-PEAL images using the SVM (for 3NN the results are unusable). The results are in Tables 4 and 5.

**Table 4.** Mean values for Accuracy (A), Precision (P), Recall (R), Specificity (S) for 15 tests with 60 FERET faces and 50 Down faces—Gaussian noise with zero mean and 0.0005 variance for training and 0.003 variance for test

	SVM (polynomial, 3-rd degree)			
	A	P	R	S
LBP (4 × 4)	84.06%	97.62%	67.06%	98.22%
DWT + LBP	89.81%	<b>99.3%</b>	78.26%	<b>99.44%</b>
EP	96.84%	98.22%	94.8%	98.55%
LBP (PCA)	<b>98.54%</b>	99.19%	<b>97.6%</b>	99.33%

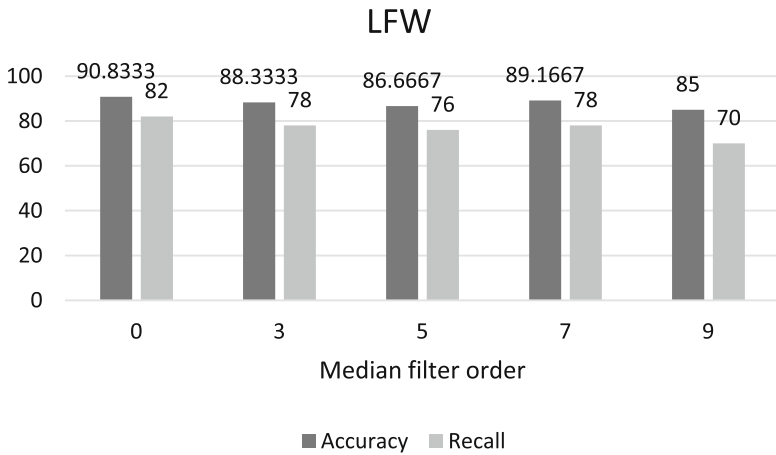
**Table 5.** Mean values for Accuracy (A), Precision (P), Recall (R), Specificity (S) for 5 tests with 60 CAS-PEAL faces and 50 Down faces—Gaussian noise with zero mean and 0.0005 variance for training and 0.003 variance for test

	SVM (polynomial, 3-rd degree)			
	A	P	R	S
LBP (4 × 4)	70.36%	60.98%	<b>100%</b>	45.66%
DWT + LBP	82.18%	72.16%	<b>100%</b>	67.33%
EP	97.63%	98.74%	96%	99%
LBP (PCA)	<b>99.09%</b>	<b>99.6%</b>	98.4%	<b>99.66%</b>



We can observe from these tables that the LBP does not perform well in noisy situations (which is well known), the DWT manages to deal better with the noise, but by far, the best results are obtained by using the projection methods on the principal components or eigenfaces of the AT&T images (the results are almost unchanged from the normal, no noise case). The recall becomes lower for the FERET database, the noise enhancing the ‘defects’ of the Down images and decreasing their identification rate. For the CAS-PEAL situation, the specificity decreases showing that, in fact, the features are classifying Chinese/non-Chinese.

We used DWT or eigenfaces projection in order to deal with noise. We made some tests by first applying median filters of different sizes (3, 5, 7, 9) in order to reduce the noise and then computing LBP (4 × 4) feature vectors. Tests were made for Down and LFW images. The results can be seen in Fig. 4. The results tend to decrease when the size of the filter increases, as expected, with the exception of the 7 sized median filter and all the results are worse than those obtained by using the DWT and eigenfaces projection methods.



**Fig. 4.** Accuracy and Recall for LBP median filtered 70 LFW, 50 Down images

We used a 3-rd degree polynomial kernel for SVM because it needed no parameter tuning as the RBF kernel needs and the results are good. Of course, in some cases we obtain better results with RBF SVMs. For example, for the LFW test without noise we get for the LBP case a better accuracy result, 91.66%, with lower precision (93.47%) and higher recall (86%). For the DWT + LBP and SVM-RBF case we obtain overall improved results with accuracy of 95.83%, precision 97.87%, 92% recall and 98.57% specificity. For the EP case we didn’t manage to find the appropriate parameter for the RBF kernel in order to improve the polynomial SVM result.

## 4 Conclusions

In the present paper we tested four feature extraction methods for Down syndrome identification in digital images. The first is the classical Local Binary Patterns method, the second is a combination of Discrete Wavelet Transform and LBP, the third method is an eigenfaces projection method using AT&T as eigenfaces provider and the fourth is a projection method on principal components of the LBP features of AT&T. All these methods give good results in distinguishing between faces with Down syndrome and healthy ones, but in the presence of noise projection methods are the best (especially the projection on the principal components of the LBP of AT&T faces) followed by the combination DWT and LBP. The feature vectors were tested using only frontal face images from the well-known face recognition datasets FERET, CAS-PEAL, LFW and some Down images collected from Internet, with SVM and 3NN as classifiers.

For the projection methods, we need to further test the influence of the dataset from which the eigenfaces or principal components are built. Other visually visible genetic disorders will be approached in a future work.

## References

1. Jones, K.L., Jones, M.C., del Campo, M.: *Smith's Recognizable Patterns of Human Malformation*, 7th edn, pp. 7–13. Elsevier Saunders, Philadelphia (2013)
2. Cassidy, S., Allanson, J.: *Management of Genetic Syndromes*, 3rd edn, pp. 309–335. Wiley Blackwell, New York (2010)
3. Zhao, W., Chellappa, R.P., Phillips, P.J., Rosenfeld, A.: Face recognition: a literature survey. *ACM Comput. Surv.* **35**(4), 399–458 (2003)
4. Jafri, R., Arabnia, H.R.: A survey of face recognition techniques. *JIPS* **5**(2), 41–68 (2009)
5. Boehringer, S., Guenther, M., Sinigerova, S., Wurtz, R.P., Horsthemke, B., Wiczorek, D.: Automated syndrome detection in a set of clinical facial photographs. *Am. J. Med. Genet. A.* **155A**(9), 2161–2169 (2011)
6. Zhao, Q., Okada, K., Rosenbaum, K., Zand, D.J., Sze, R., Summar, M., Linguraru, M.G., Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N.: Hierarchical constrained local model using ICA and its application to down syndrome detection. In: *Proceedings of MICCAI 2013, Part II: 16th Int. Conf. Nagoya, Japan, September 22–26*, pp. 222–229 (2013)
7. Zhao, Q., Rosenbaum, K., Sze, R., Zand, D.J., Summar, M., Linguraru, M.G.: Automated Down syndrome detection using facial photographs. In: *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka*, pp. 3670–3673 (2013)
8. Zhao, Q., Rosenbaum, K., Okada, K., Zand, D.J., Sze, R., Summar, M., Linguraru, M.G.: Down syndrome detection from facial photographs using machine learning techniques. In: *SPIE Medical Imaging, Medical Imaging 2013: Computer-Aided Diagnosis*, pp. 867003–867003, 28 February 2013
9. Saraydemir, Ş., Taşpınar, N., Eroğul, O., Kayserili, H., Dinçkan, N.: Down syndrome diagnosis based on gabor wavelet transform. *J. Med. Syst.* **36**(5), 3205–3213 (2012)
10. Phillips, P.J., Wechsler, H., Huang, J., Rauss, P.: The *FERET* database and evaluation procedure for face recognition algorithms. *Image Vis. Comput. J.* **16**(5), 295–306 (1998)

11. Gao, W., Cao, B., Shan, S., Chen, X., Zhou, D., Zhang, X., Zhao, D.: The CAS-PEAL large-scale chinese face database and baseline evaluations. *IEEE Trans. Syst. Man. Cybern. (Part A)* **38**(1), 149–161 (2008)
12. Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E.: Labeled faces in the wild: a database for studying face recognition in unconstrained environments. University of Massachusetts, Amherst, Technical Report 07–49, October 2007
13. Turk, M.A., Pentland, A.P.: Face recognition using eigenfaces. *Computer vision and pattern recognition*. In: *Proceedings CVPR 1991*, pp. 586–591 (1991)
14. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
15. Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: application to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(12), 2037–2041 (2006)
16. Mallat, S.: *A Wavelet Tour of Signal Processing*. Academic Press, Cambridge (2008)
17. Samaria, F.S., Harter, A.C.: Parameterisation of a stochastic model for human face identification. In: *Proceedings of the Second IEEE Workshop on Applications of Computer Vision*, 1994, Sarasota, FL, pp. 138–142 (1994)
18. <http://www.nist.gov/itl/iad/ig/feret.cfm>
19. Viola, P., Jones, M.J.: Rapid object detection using a boosted cascade of simple features. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 511–518 (2001)