

Pengenalan Ekspresi Wajah Menggunakan *Local Binary Pattern (LBP)*

Facial Expression Recognition Using Local Binary Pattern (LBP)

Adi Saputra (1103081084)

Fakultas Informatika

Telkom University

adisaputra@telkomuniversity.ac.id

Tjokorda Agung Budi W, S.T, M.T

Fakultas Informatika

Telkom University

ABSTRAK

Pengenalan ekspresi wajah merupakan subjek yang penelitian yang menarik. Pengaplikasiannya membuka kesempatan hadirnya suatu bentuk interaksi yang lebih alami antara manusia dan komputer.

Dalam paper ini dibangun sebuah sistem pengenalan ekspresi wajah menggunakan metode *Local Binary Pattern (LBP)* sebagai metode ekstraksi ciri. LBP memiliki perhitungan yang cepat karena algoritmanya yang sederhana, namun tetap dapat menghasilkan ciri wajah secara efisien. Percobaan dilakukan menggunakan dataset JAFFE. Klasifikasi menggunakan metode kNN dan SVM (*Support Vector Machine*) dengan skema *one-against-one*.

Hasil pengujian menunjukkan bahwa metode LBP dapat menghasilkan ciri diskriminatif dari ekspresi wajah. Dari percobaan yang dilakukan didapatkan akurasi tertinggi sebesar 84% untuk skema pengujian *person-dependent* dengan konfigurasi operator LBP ($p=8, r=8$) dan segmentasi 16 *region*.

Kata Kunci: - Pengenalan Ekspresi Wajah, LBP, kNN, SVM

1. PENDAHULUAN

Ekspresi wajah dalam kehidupan manusia sehari-hari digunakan sebagai bentuk respon alami manusia yang menggambarkan perasaan yang dirasakan oleh seseorang dalam berinteraksi dengan suatu hal tertentu. Dalam interaksi antar sesama manusia ekspresi digunakan sebagai bagian dari komunikasi. Di lain hal, perkembangan teknologi dari masa ke masa mengalami lompatan besar ketika suatu bentuk interaksi baru diimplementasikan seperti penggunaan tombol, layar monitor, GUI, layar sentuh, hingga perintah suara. Kunci perkembangan teknologi berada pada pengguna, oleh karena itu interaksi komputer adalah bagian penting dalam perkembangan teknologi dan mengerti respon dari pengguna menjadi sangat penting. Salah

satu pendekatan yang dilakukan untuk mengerti respon dari pengguna adalah dengan sebuah sistem pengenalan ekspresi wajah. Dengan mengerti ekspresi dari pengguna, komputer dapat dibuat seolah-olah dapat mengerti perasaan manusia.

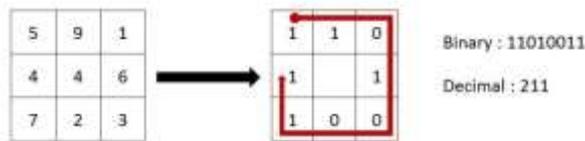
Penelitian tentang ekspresi wajah telah dimulai sejak tahun 1978 [1] dan banyak diteliti di area *computer science* pada tahun 90an dan berlangsung hingga sekarang. Sistem pengenalan ekspresi dua tahapan penting, yaitu representasi ciri wajah dan desain pengklasifikasian. Representasi ciri wajah didapatkan dengan mengekstraksi inputan berupa citra digital dengan menggunakan metode tertentu. Adapun metode ekstraksi ciri yang sudah digunakan untuk membangun sistem pengenalan ekspresi wajah diantaranya metode-metode handal dan terkenal seperti *Gabor Wavelet Analysis (GWA)*, *Principal Component Analysis (PCA)* dan *Linear Discriminant Analysis (LDA)* telah berhasil diimplementasikan dan mendapatkan akurasi yang baik tergantung dari metode pemrosesan data serta metode klasifikasi yang digunakan. Namun metode-metode ini masing-masing memiliki kelemahan. GWA kompleks secara komputasi [2]. PCA sebagai sebuah metode *unsupervised-learning* gagal mengekstrak ciri diskriminatif dari data berdimensi besar [2]. LDA sebagai metode *supervised-learning* memiliki batasan mendasar, yaitu jumlah maksimum *embedded feature* harus lebih sedikit dari jumlah kelas data dikarenakan kurangnya peringkat antar-kelas pada *scatter* matriks [3].

Local Binary Pattern (LBP) pertama kali digunakan pada *texture analysis* [4]. Kemudian LBP pernah juga digunakan sebagai metode untuk *face detection* [5] dan *face recognition* [6]. Penggunaan metode LBP pada paper ini dimaksudkan untuk menghasilkan sistem pengenalan ekspresi wajah yang efisien dari segi algoritma dan komputasi namun tetap memberikan hasil yang dapat bersaing dengan metode-metode yang sudah digunakan sebelumnya.

2. LOCAL BINARY PATTERN

Operator LBP pertama kali dikenalkan oleh Ojala et al [13], dan telah terbukti sebagai deskriptor tekstur yang tangguh. Operator LBP melabeli piksel-piksel dari sebuah citra dengan melakukan proses thresholding ketetanggaan 3x3 dari masing-masing piksel sebagai nilai tengah dan mengubah hasilnya menjadi nilai biner, dan 256-bin LBP histogram digunakan sebagai *texture descriptor*. Bilangan biner yang dihasilkan (disebut *Local Binary Pattern* atau *LBP codes*) mengkodekan lokal primitif termasuk variasi dari lengkungan sisi, bintang, area datar, dsb.

Batasan dari operator dasar LBP (yaitu 3x3 *neighborhood* tidak dapat menangkap ciri dominan yang memiliki struktur dengan skala besar). Karena itu kemudian operator dikembangkan untuk menggunakan tingkat ketetanggaan dengan ukuran yang berbeda [14]. Menggunakan *circular neighborhood* dan *bilinear interpolation* nilai piksel memungkinkan beragam ukuran radius dan jumlah pengambilan titik sampel. Lihat *figure 5* untuk contoh LBP operator, dimana (P,R) menunjukkan jumlah pengambilan titik sampel P dari radius R dari sebuah *circular symmetric neighbor set*.



Gambar 1. Perhitungan dasar LBP [1]

Operator LBP P,R menghasilkan 2^P nilai output berbeda, berdasarkan dari 2^P *binary pattern* yang berbeda yang dapat dibentuk oleh P piksel dalam *neighbor set*. Telah terlihat bahwa bin tertentu mengandung informasi yang lebih dari yang lainnya [14]. Oleh karenanya dimungkinkan untuk menggunakan hanya subset dari 2^P *Local Binary Pattern* untuk mendeskripsikan tekstur citra. Ojala et al [14] menyebut *fundamental pattern* ini sebagai *uniforms pattern*. Sebuah *local binary pattern* disebut *uniform* apabila mengandung paling banyak dua *bitwise transition* dari 0 ke 1 dan sebaliknya pada sirkular *binary string*. Sebagai contoh, 00000000, 001110000, dan 1110001 adalah *uniform pattern*. Telah diteliti bahwa *uniform pattern* terhitung hampir 90% dari jumlah *pattern* keseluruhan pada (8,1) *neighborhood* dan sekitar 70% pada (16,2) *neighborhood* di *texture image* [14]. Mengumpulkan *patterns* yang memiliki lebih dari 2 transisi ke dalam *single bin* menghasilkan sebuah LBP operator, diindikasikan dengan $LBP_{P,R}^{u2}$ dengan jumlah bin kurang dari 2^P . Sebagai contoh jumlah label pada *neighborhood* 8 piksel dihasilkan 256 label sedangkan untuk LBP^{u2} dihasilkan 59.

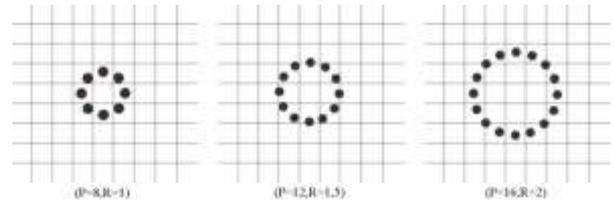
Setelah melabeli citra dengan *local binary pattern*. Histogram dari citra $f(x,y)$ dapat didefinisikan sebagai,

$$H_i = \sum_{x,y} I(f_i(x,y) = i), \quad i = 0, \dots, n - 1$$

Dimana n adalah jumlah label biner yang dihasilkan oleh LBP operator dan

$$I(A) = \begin{cases} 1 & A \text{ is true} \\ 0 & A \text{ is false} \end{cases}$$

LBP histogram ini mengandung informasi tentang distribusi *local micro-pattern*, seperti sisi-sisi, titik-titik, dan area datar, di seluruh permukaan citra, sehingga secara statistik dapat mendeskripsikan karakteristik yang terdapat pada citra.



Gambar 2. Tiga contoh extended LBP [1]

Citra wajah dapat dilihat sebagai komposisi *micro-pattern* yang mana dapat dideskripsikan oleh LBP secara efektif. Oleh karena itu, secara intuitif LBP dapat digunakan untuk merepresentasikan citra wajah [5]. Sebuah histogram LBP dihitung dari keseluruhan citra wajah hanya mengkodekan kemunculan *micro-pattern* tanpa mengindikasikan lokasi dari *pattern* tersebut.

Untuk mengikutsertakan informasi bentuk dari wajah, ekstraksi LBP dilakukan pada citra wajah yang dibagi sama rata menjadi *region-region* yang lebih kecil R_0, R_1, \dots, R_m . Histogram yang diekstrak dari tiap-tiap *sub region* di-concat menjadi sebuah histogram ciri tertambah secara spasial yang didefinisikan dengan

$$H_{i,j} = \sum_{x,y} I\{f_i(x,y) = i\} I\{(x,y) \in R_j\}$$

dimana $i=0, \dots, n - 1, j=0, \dots, m - 1$.

Histogram ciri yang terbentuk merepresentasikan *local texture* and bentuk global dari wajah. beberapa parameter dan di optimalisasi untuk menghasilkan ekstraksi ciri yang lebih baik. Pertama yaitu operator LBP, dan kedua jumlah pembagian region.

3. PENGENALAN EKSPRESI WAJAH

Sistem pengenalan ekspresi wajah secara garis besar memiliki komposisi proses yang sama dengan sistem rekognisi secara umum, yaitu akuisisi data, ekstraksi ciri, dan klasifikasi. Berikut ini proses-proses yang terdapat dalam sistem.

3.1 Loading Dataset

Proses loading dataset adalah proses yang pertama kali dilakukan dimana kumpulan citra wajah dibaca oleh sistem.

Informasi citra wajah dengan format warna *grayscale* disimpan ke dalam matriks 2 dimensi.



Gambar 3. Image Grayscale

3.2 Preprocessing

Proses *preprocessing* bertujuan untuk menghasilkan inputan citra wajah yang seragam. Dimana *object* non-wajah disisihkan dan hanya objek wajah yang masuk ke sistem. Lebih lanjut inputan sistem berupa citra wajah dengan ukuran 256x256 pixel akan dipotong seragam menjadi citra berukuran 110x150 pixel dengan mengambil area wajah saja. Penentuan area cropping dilakukan dengan acuan titik tengah dari jarak antar mata. Koordinat titik tengah dari jarak antar mata dicari dengan menggunakan proses *face detection* dan *Eyes detection* dengan algoritma Viola-Jones. Implementasi algoritma Viola-Jones dilakukan dengan menggunakan modul program yang terdapat di Matlab 2015a. *Eyes detection* dibutuhkan untuk mencari koordinat mata untuk selanjutnya dicari titik tengah antar jarak kedua mata. Sedangkan penggunaan *face detection* sebelum *eyes detection* dimaksudkan untuk mempersempit area pencarian koordinat mata terhadap citra, dikarenakan pencarian koordinat mata melalui *eyes detection* secara langsung pada gambar inputan menghasilkan pendeteksian mata yang salah.

3.2.1 Face Detection

Face detection atau deteksi wajah bertujuan untuk mendeteksi keberadaan objek berupa wajah pada sebuah citra. Inputan berupa citra dan output yang dihasilkan berupa koordinat wajah didalam citra apabila benar ditemukan wajah didalamnya. Penggunaan *Face Detection* Sebelum *Eyes Detection* bertujuan agar *Eyes detection* lebih akurat.

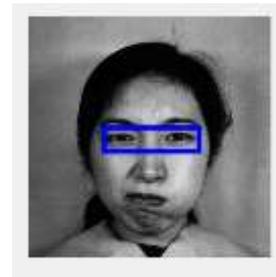


Gambar 4. Deteksi wajah

3.2.2 Eyes Detection

Eyes detection atau deteksi pasangan mata bertujuan untuk mendeteksi keberadaan pasangan mata di dalam sebuah citra.

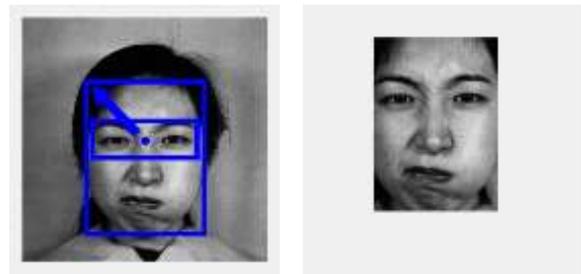
Proses *Face Detection* sebelumnya membuat area pencarian pasangan mata terfokus ke wajah. Proses ini menghasilkan *output* koordinat pasangan mata.



Gambar 5. Eye Pair Detection

3.2.3 Image Cropping

Image Cropping atau pemotongan citra bertujuan mengambil sebagian informasi berupa citra yang lebih kecil. Koordinat potong diambil dari titik tengah pasangan mata yang di perluas menjadi segi empat dengan ukuran 110x150 *pixel*. Hasil dari *image cropping* ini menjadi *output* dari proses *preprocessing* yang selanjutnya menjadi ciri yang akan diekstrak.



Gambar 6. Posisi cropping mengacu pada posisi titik tengah jarak antar mata; Citra wajah ter-crop

3.3.3 Ekstraksi Ciri Menggunakan Local Binary Pattern

Proses ini menjadi fokus dalam pengerjaan tugas akhir ini. proses ini bertujuan untuk mendapatkan ciri dari inputan citra berupa histogram yang kemudian direpresentasikan menjadi matriks 1xn. Nilai-nilai yang dihasilkan dalam *vector* ciri dinormalisasi ke dalam rentang 0-1.

3.3.4 kNN Classifying

Proses ini bertujuan untuk melakukan klasifikasi data uji. Sebagai *weak classifier*, kNN dalam pengerjaan tugas akhir ini diposisikan sebagai pembanding untuk melihat akurasi tingkat diskriminasi dari ciri yang dihasilkan dari metode LBP, sebelum kemudian dilanjutkan dengan metode klasifikasi SVM.

3.3.5 SVM Training

Proses ini merupakan langkah awal klasifikasi menggunakan metode SVM. Proses *training SVM* bertujuan untuk mendapatkan informasi *hyperplane*, *support vector*, dsb yang direpresentasikan ke dalam sebuah *struct*. *Struct* didapat dari pelatihan *binary class* atau dua kelas yang berbeda. Dalam percobaan ini akan dilakukan

klasifikasi untuk 7 kelas yang berbeda, untuk itu dibutuhkan skema klasifikasi *SVM multiclass*.

3.3.6 SVM Classifying

Proses klasifikasi SVM dilakukan menggunakan *struct* yang dihasilkan dari proses *training SVM*. Di sini diimplementasikan skema *SVM multiclass* yang diimplementasi adalah *one-against-one*, dimana tiap kelas di klasifikasi secara biner dengan tiap kelas lainnya. Hasil dari klasifikasi yang ada kemudian di rangking. Kelas yang paling banyak menjadi kelas hasil.

3.3.7 Perhitungan Akurasi

Perhitungan akurasi dilakukan untuk melihat tingkat keberhasilan sistem dalam mengenali ekspresi dari inputan wajah. Akurasi menjadi salah satu poin penilaian keberhasilan sebuah sistem. Dalam percobaan yang dilakukan tingkat akurasi diukur dari hasil klasifikasi dari data uji dibandingkan dengan kelas uji sesungguhnya. Jumlah benar dihitung dan dibagi dengan banyaknya data uji. Kemudian dikalikan 100%.

4. PENGUJIAN

4.1 Data Set

Dataset yang digunakan dalam tugas akhir ini adalah JAFFE dataset. Sebuah database yang berisi 213 citra digital dari 7 ekspresi wajah (6 ekspresi wajah dasar + 1 ekspresi netral) diperagakan oleh model-model wanita Jepang. Setiap citra telah di-*rating* ke dalam 6 *emotion adjective* oleh 60 subject individu berkebangsaan jepang. Database dibuat dan dirancang oleh Michael Lyons, Miyuki Kamachi, dan Jiro Gyoba. Foto-foto diambil di Departemen Psikologi di Universitas Kyushu. [12]



Gambar 7. Contoh Citra Ekspresi Dalam JAFFE Dataset

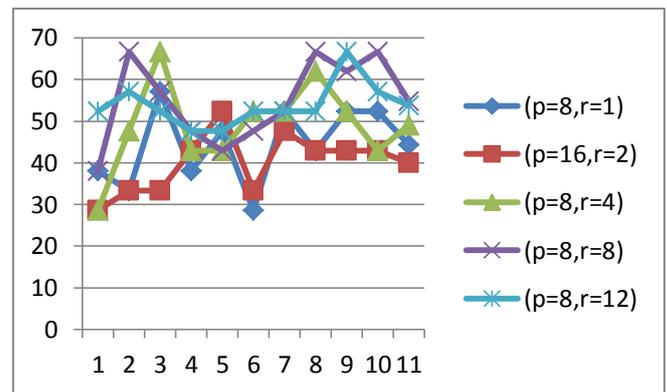
4.2 Person-Dependent

Skema pengujian ini bertujuan untuk melihat performa sistem secara umum. Dalam skema pengujian ini digunakan skema *10-fold cross validation test*. Dalam skema ini sampel dibagi menjadi 10 subset secara random, kemudian 1 subset digunakan sebagai data uji dan 9 subset sisanya digunakan sebagai data latih. Selanjutnya secara bergiliran 1 subset digunakan sebagai data uji dan sisanya digunakan sebagai data latih. Proses pelatihan dan

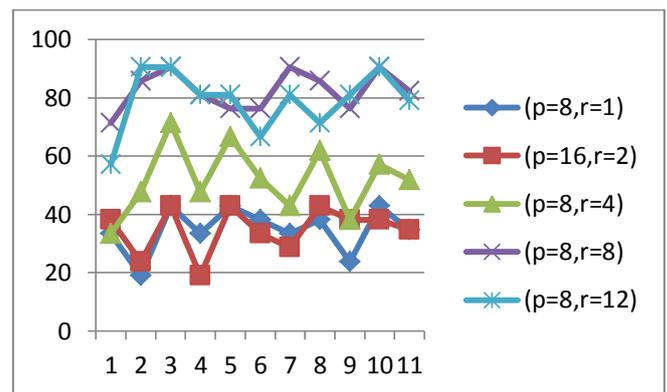
pengujian dilakukan hingga semua subset telah digunakan sebagai data uji. Akurasi didapat dari rata-rata akurasi pengujian dari semua subset. Skema pengujian ini digunakan di dalam setiap skenario pengujian yang dilakukan.

4.2.1 Eksperimen: Mengubah Operator Radius dan Pengambilan Titik Sampel

Pada pengujian ini akan dilihat pengaruh konfigurasi operator LBP dari segi penggunaan radius dan variasi pengambilan titik sampel. Pada skenario ini dilakukan perhitungan LBP dengan radius 1,4,8,12 dan pengambilan titik sampel 8, dan 16 titik.



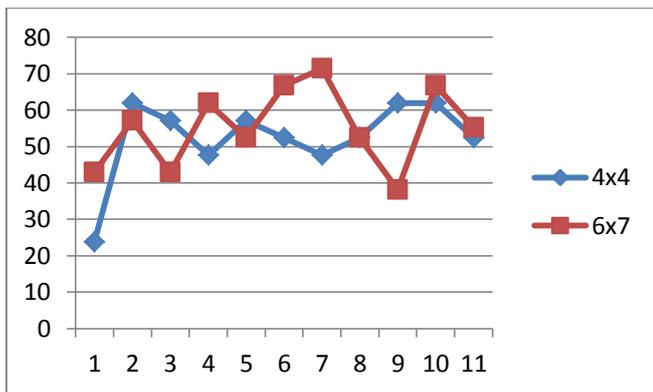
Gambar 8. Hasil Pengujian 1 Dengan Classifier kNN



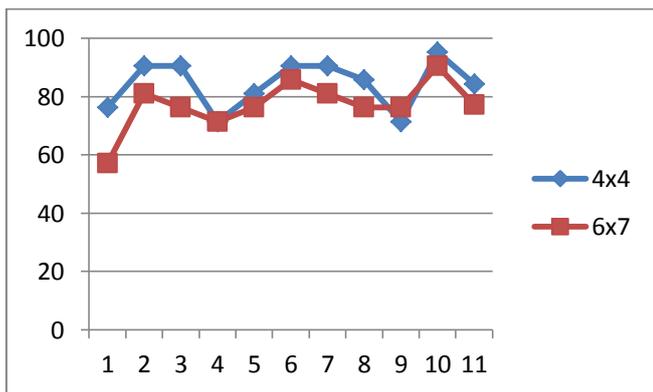
Gambar 9. Hasil Pengujian 1 Dengan Classifier SVM

4.2.2 Eksperimen: Segmentasi Citra Inputan

Tujuan dari pengujian ini adalah untuk menguji dan menganalisis efektifitas dari pembagian *region* wajah. Pada pengujian ini, pertama dilakukan pembagian *region* di area wajah sebanyak 16 *region*, 4 baris, 4 kolom, yang membagi objek-objek diwajah seperti dahi, mata, hidung, pipi, dan mulut. Kedua, seperti yang dilakukan di[1], dilakukan pembagian *region* di area wajah sebanyak 42 *region*, 7 baris, 6 kolom, sehingga didapat masing-masing *region* sebesar 18x21px dari ukuran area wajah keseluruhan (setelah preprocessing) sebesar 110x150px. Pada pengujian ini akan digunakan konfigurasi LBP dengan akurasi terbaik dari pengujian sebelumnya.



Gambar 10. Hasil Pengujian 2 Dengan Classifier kNN



Gambar 11. Hasil Pengujian 2 Dengan Classifier SVM

4.3 Person-Independent

Skema pengujian ini bertujuan untuk mengevaluasi kehandalan sistem dalam mengenali ekspresi dari subjek yang berbeda. Dalam skema pengujian ini sample dibagi berdasarkan individu. Dalam dataset JAFFE terdapat 10 individu yang kemudian dibagi untuk data latih dan data uji dengan komposisi 8 data uji, 2 data latih. Skema pengujian *person-independent* cenderung sulit untuk dilakukan dengan informasi, dalam hal ini data latih yang terbatas. Dalam tugas akhir ini skema pengujian *person-independent* hanya bersifat komplementer untuk menguji tingkat kehandalan sistem pengenalan ekspresi yang telah dibangun.

Table 1. Hasil Pengujian Person-Independent

| Rasio Perbanding Data Uji per Data Latih | Akurasi (%) |
|--|-------------|
| 2:8 | 65.1 |
| 2:7 | 62.8 |
| 2:6 | 58.1 |

5. KESIMPULAN

Paper ini menyajikan pengenalan ekspresi wajah menggunakan metode *Local Binary Pattern* yang berdasarkan percobaan berhasil mengekstraksi ciri yang diskriminatif dari citra wajah yang terdapat pada dataset JAFFE. Lebih lanjut ditemukan bahwa penggunaan operator radius dan jumlah pengambilan titik sampel yang tepat pada metode *Local Binary Pattern* dapat menambah kualitas ciri yang dihasilkan. Segmentasi pada citra wajah berdasarkan objek-objek wajah seperti alis, mata, hidung, dan mulut juga berhasil menambahkan ciri diskriminatif pada ciri yang dihasilkan.

6. REFERENSI

- [1] T. Ojala, M. Pietikäinen, and D. Harwood (1994), "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions", Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR 1994), vol. 1, pp. 582 - 585.
- [2] C Shan, S Gong, and P McOwan, "Robust facial expression recognition using local binary patterns", Proc. IEEE International Conference on Image Processing, 2005, pp. 370-373.
- [3] P N Belhumeur, J P Hespanha, and D J Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No.7, 1997, pp. 711-720.
- [4] T. Ojala, M. Pietikäinen, and D. Harwood, "A Comparative study of texture measures with classification based on feature distribution," Pattern Recognition, vol.29, no.1, 1996.
- [5] T.Ahonen, A. Hadid, and M. Pietikinen, "Face Recognition with local binary pattern," in ECCV, 2004, pp.469-481.
- [6] A. Hadid, M. Pietikinen, and T.Ahonen, "A Discriminative feature space for detecting and recognizing faces," in IEEE CVPR, June 2004, pp. 797 - 804.
- [7] M.Lyons. "The Japanese Female Facial Expression (JAFFE) Database." n.p, n.d. Web. 18 Nov 2014.
- [8] G. Donato, M.S. Bartlett, J.C. Hager, P. Ekman, T.J. Sejnowski, "Classifying Facial Actions", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 21, No. 10, pp. 974-989, 1999.

- [9] Matsumoto, Keltner, Shiota, Frank, & O'Sullivan, "Sequential Dynamics of Culturally Moderated Facial Expressions of Emotion," APS, Psychological Science, 2008.
- [10] E. Wood, Richard. 2008. "Digital Image Processing Third Edition".
- [11] Stephen Johnson (2006). "Stephen Johnson on Digital Photography". O'Reilly. ISBN 0-596-52370-X.
- [12] M.Lyons. "The Japanese Female Facial Expression (JAFFE) Database.", n.p, n.d. Web. 18 Nov 2014.
- [13] T. Ojala, M. Pietikäinen, and D. Harwood (1994), "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions", Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR 1994), vol. 1, pp. 582 - 585.
- [14] T Ojala, M Pietik inen, and T M Enp, "Multiresolution gray scale and rotation invariant texture analysis with local binary patterns", IEEETransactions on Pattern Analysis and MachineIntelligence, Vol. 24, No. 7, 2002, pp. 971-987.
- [15] Y. Tian, T Kanade, J. Cohn, "Recognition Action Units for Facial Expression Analysis". IEEETransactions on Pattern Analysis and MachineIntelligence, Vol. 23, No. 2, 2001.
- [16] Pantic, Maja. "Facial Expression Recognition." - Springer. Springer US, 27 May 2014. Web. 1 Sept. 2015.
- [17] Vladimir N. Vapnik. "The nature of statistical learning theory". Springer-Verlag New York, Inc., New York, NY, USA, 1995.
- [18] M.S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, J. Movellan, "Recognizing facial expression: machine learning and application to spotaneous behavior", in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005.
- [19] M. Bartlett, G. Littlewort, I. Fasel, R. Movellan, "Real time face detection and facial expression recognition: development and application to human-computer interaction", in: CVPR Workshop on CVPR for HCI, 2003.
- [20] V.N. Vapnik, "Statistical Learning Theory", Wiley, New York, 1998.
- [21] S Zhang, X Zhao, B Lei et al., "Facial Expression Recognition Based on Local Binary Patterns and Local Fisher Discriminant Analysis", WSEAS TRANSACTIONS on SIGNAL PROCESSING.
- [22] R. Kohavi. "A Study of CrossValidation and Bootstrap for Accuracy Estimation and Model Selection". IJCAI, 1995
- [23] Yang, M. H. (2004). "Recent advances in face detection". IEEE ICPR 2004 Tutor.
- [24] K. Beyer , J. Goldstein , R. Ramakrishnan , U. Shaft. "When Is "Nearest Neighbor" Meaningful?".International Conference on Database Theory,1999.
- [25] L.Sørensen, S.B.Shaker, M. de Bruijne. "Texture Classification in Lung CT Using Local Binary Patterns". MICCAI. 2008.
- [26] C.Shan, S.Gong, P.McOwen. "Facial expression recognition based on Local Binary Patterns: A comprehensive study". Image and Vision Computing 27 .2009