

Estimasi Pose Kepala Menggunakan Histogram of Oriented gradients dan Multiclass Support Vector Machine

Estimating Head Pose Using Histogram of Oriented Gradients and Multiclass Support Vector Machine

Ambiya Wikarsa¹, Rosa Ariani Sukamto², Yaya Wihardi³

Prodi Studi Ilmu Komputer Departemen Pendidikan Ilmu Komputer Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam Universitas Pendidikan Indonesia

Bandung, Indonesia

¹ambiya.wikarsa@student.upi.edu, ²rosa.ariani@upi.edu, ³yayawihardi@upi.edu

Abstrak— Pose kepala mengindikasikan serta memvisualisasi seseorang akan atensi dan ketertarikan akan sesuatu, hal itu memainkan peranan penting di berbagai macam aplikasi. Dengan banyaknya jumlah kelas dari pose kepala membuat tugas dalam mengestimasi ini merupakan tugas yang sulit. Dalam penelitian ini metode yang digunakan dalam mengestimasi pose kepala adalah Histogram of Oriented Gradients dan Multiclass Support Vector Machine. Histogram of Oriented Gradient digunakan sebagai ekstraksi fitur kepada kepala gambar yang akan diestimasi menggunakan fungsi dalam OpenCV dan Multiclass Support Vector Machine dijalankan sebagai pengestimasi pose kepala menggunakan fungsi dari Scikit-learn. Head pose database dari INRIA Rhône-Alpes 2004 digunakan sebagai Training dan Testing, memiliki jumlah gambar sebanyak 2790 buah lalu dibagi menjadi 93 kelas untuk pose kepala menghasilkan 30 gambar pose per kelas. Dilakukan pengujian dengan menggunakan cross validation sebanyak 5-folds rata-rata akurasi yang didapat adalah 22,5% dengan rata-rata dari fi-score (0,21), precision (0,23), recall (0,22).

Kata Kunci: *estimasi pose kepala, Histogram of Oriented gradient, Multiclass Support Vector Machine*

Abstract— Head poses indicates and visualizes someone's attention and interest in something, this plays an important role in various applications. The large number of classes of head poses makes its estimation a difficult task. In this study, the methods used in estimating head pose are the Histogram of Oriented gradients and the Multiclass Support Vector Machine. The Histogram of Oriented Gradient is used as feature extraction for the head images which will be estimated using the functions in OpenCV and the Multiclass Support Vector Machine is run as an estimation of the head pose using the Scikit-learn function. Head pose database from INRIA Rhône-Alpes 2004 is used as a Training and Testing, consisting of 2790 images then divided into 93 classes for head poses resulting in 30 poses per class. The test was conducted using cross validation of 5-folds, the average accuracy obtained was 22.5% with an average of fi-score (0.21), precision (0.23), recall (0.22).

Keywords: *estimated head pose, Histogram of Oriented gradient, Multiclass Support Vector Machine*

I. PENDAHULUAN

Seiring dengan cepatnya perkembangan TIK saat ini, semua hal dapat dijadikan informasi salah satunya kepala. kepala manusia yang merupakan bagian penting dari tubuh manusia yang mana analisis terhadap pergerakannya dapat merepresentasikan banyak hal, salah satunya pose kepala. salah satu informasi yang di dapat pose kepala adalah dapat mengetahui arah ketertarikan akan sesuatu. penelitian akan pose kepala sudah banyak dilakukan dengan hasil yang bermacam-macam [1]. dalam tulisan ini topik yang akan di angkat adalah estimasi pose kepala.

“Dimanakah orang tersebut ?” adalah pertanyaan yang biasanya dijadikan sebuah masalah dalam penelitian terkait topik deteksi wajah. Dalam estimasi pose kepala pertanyaan yang kita sampaikan dalam penelitian ini adalah “Kemana orang tersebut melihat ?”. berdasarkan asumsi bahwa pose kepala sangat berkorelasi dengan arah pandangan visual [2]. Menentukan arah bidang visual, arah pandangan seseorang biasanya terpantul dari matanya. Tetapi, mengestimasi langsung pandangan seseorang dari bagian mata dalam gambar wajahnya akan sulit karena membutuhkan deteksi wajah. Untungnya, masalah tersebut dapat diselesaikan dengan jalan lain. Berdasarkan penelitian dalam bidang psikologis, menyatakan bahwa arah pandang seseorang dapat diindikasikan oleh orientasi kepala, kepala manusia contohnya pose kepala [3]. Pose kepala dapat di estimasi dari gambar wajah lebih baik dari hanya sekitar bagian mata saja, dimana dalam informasi yang didapat lebih banyak kepala. Karena itu, estimasi pose kepala seseorang dapat diperkirakan dari kepala.

Pose kepala dapat dilihat sebagai salah satu tipe bahasa tubuh, gerak isyarat yang dihasilkannya dapat menjadi salah satu komunikasi nonverbal. Berdasarkan penelitian [4], dalam menjelaskan dan menunjukkan sebuah masalah

dalam penemuan visual focus of attention dari berbagai jumlah orang berkeliaran. Melacak lokasi tubuh, kepala, dan pose kepala mereka, menghasilkan salah satunya bahwa arah terpantulnya visual focus of attention seseorang, menyiratkan ketertarikan dia dengan pertahatiannya pada sebuah target visual dan reaksi pada ransangan visual.

Dalam penerapannya, aplikasi yang menggunakan estimasi pose kepala sebagai sebagai kuncinya atau input masuknya. Mengembangkan aplikasi yang dalam interaksinya kepada aplikasi tersebut tidak menggunakan tangan atau menggantikan kontrolnya seperti mengendalikan cursor mouse untuk komputer yang hanya membaca pose kepala saja. Pose kepala juga bisa digunakan untuk memahami tingkah laku manusia, seperti pengawasan atensi pengendara atau analisis terhadap pengunjung yang lewat terhadap sebuah iklan yang berada di luar ruangan. Analisis tingkah laku dalam interaksi sosial seseorang dalam berinteraksi dengan menganalisis pose kepala sebagai pemahaman dalam interaksi antar mereka. Selain itu pose kepala paling sering digunakan dalam virtual reality [1].

Untuk dapat membuat aplikasi pose kepala tersebut dibutuhkan sebuah cara untuk dapat membaca dan mengidentifikasi pola pada setiap pose kepala yang diidentifikasi. Salah satu strategi untuk mengidentifikasi pola tersebut adalah dapat menggunakan machine learning. Machine learning merupakan cara dimana komputer atau mesin dapat belajar sendiri untuk dapat memecahkan masalah. Bilamana untuk dapat memecahkan masalah tersebut mesin membutuhkan latihan untuk dapat menyelesaikan sebuah masalah..

Support Vector Machine salah satu teknik yang terdapat dalam machine learning akan digunakan dalam penelitian untuk estimasi pose kepala ini. sebelum masuk kedalam Support Vector Machine gambar dilakukan ekstraksi fitur untuk dapat menghasilkan pola yang dapat dengan mudah dibacanya. Teknik ekstraksi fitur yang digunakan adalah Histogram of oriented gradients.

II. ESTIMASI POSE KEPALA

A. Histogram of oriented gradients

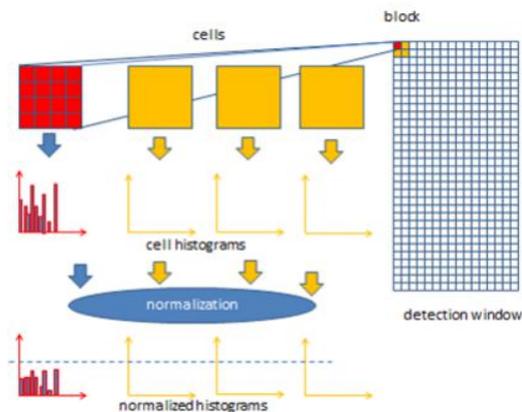
Histogram of oriented gradients (HOG) adalah deskriptor fitur yang digunakan sebagai deteksi objek dalam *computer vision* dan pengolahan citra. Deskriptor HOG adalah teknik perhitungan kemunculan orientasi gradien dalam lokasi pose dari sebuah gambar, atau *region of interest* (ROI) (Gambar 1) [5].

Implementasi dari algoritma deskriptor HOG dapat di urutkan sebagai berikut [5] :

- 1) Membagi gambar menjadi sebuah wilayah kecil yang terhubung disebut cells, untuk setiap cell hitung arah histogram gradien atau orientasi tepi untuk pixel dalam cells.
- 2) Diskritkan setiap cell menjadi sudut bins berdasarkan dari orientasi gradien,
- 3) Setiap cells pixel berkontribusi berat gradien untuk di sesuaikan dengan sudut bin
- 4) Kelompok dari cell yang berdekatan bisa di anggap sebagai wilayah spatial disebut blocks. Pengelompokan dari cells menjadi sebuah blocks adalah basis untuk pengelompokan dan normalisasi histogram.
- 5) Kelompok normalisasi dari histogram di representasikan dengan block histogram. Set dari block histogram ini merepresentasikan deskriptor

Komputasi dari dekriptor HOG sebaiknya mengikuti konfigurasi parameter sebagai berikut [5] :

- 1) Masks untuk derivatif dan gradien
- 2) Geometri dari pembelahan gambar menjadi cells dan mengelompokkan cells menjadi block.
- 3) Block overlapping
- 4) Normalisasi parameter



Gambar 1 ilustrasi proses HOG

Rekomendasi nilai parameter untuk HOG [18] :

- 1) 1D derivatif maks terpusat [-1,0,+1]
- 2) Deteksi ukuran window adalah 64x128
- 3) Ukuran cell adalah 8 x 8
- 4) Ukuran block adalah 16x16 (2x2 cells)

B. Support Vector Machine

Support Vector Machine(SVM) adalah mesin pembelajaran untuk masalah klasifikasi dua kelompok, yang secara konseptual mengimplementasikan ruang input *non-linearly* yang dipetakan kedalam ruang fitur berdimensi tinggi [6]. *Linear decision surface* dibangun dari hal tersebut. SVM diterapkan dalam berbagai macam

aplikasi, klasifikasi sinyal, identifikasi kanker, pengenalan wajah, gangguan bicara, dan bioinformatikan.

Citra pertama kali difilter dan di ekstraksi fitur untuk dapat diterapkan dengan linear SVM, karena output yang diharapkan adalah sebuah klasifikasi pose kepala. Berdasarkan pada sample pose kepala rata-rata yang disimpan. Karena itu, linearly separable SVM bisa digunakan untuk mengklasifikasi data dalam penelitian.

Saat pertama kali dikenalkan SVM hanya dapat mengklasifikasi data ke dalam dua kelas [7]. Dalam pengimplemantasian SVM multikelas terdapat banyak pilihan seperti menggabungkan beberapa SVM biner atau menggabungkan ke dalam sebuah bentuk permasalahan optimasi.

Beberapa metode SVM yang digunakan untuk permasalahan klasifikasi multi kelas [8]:

1) *One-Versus-One*: suatu sample data akan diuji pada setiap pasang kelas. Setiap pasangan kelas memiliki bidang pemisah yang masing-masing. Oleh karena itu, akan terdapat $c * (n - 1) / 2$ atau C_2^n bidang pemisah, dengan n banyak kelas. Diantara pasangan tersebut, kelas yang paling banyak terpilih oleh sample data tersebut merupakan kelas sampel data yang bersangkutan. Jika tidak ada kelas yang terbanyak, maka dilakukan analisis terhadap nilai fungsi tujuan sehingga diperoleh pola pengklasifikasian.

2) *One-Versus-All*: pada metode ini terdapat n kali klasifikasi SVM, dengan n adalah banyak kelas. SVM melakukan dua kelas antara satu kelas dengan kelas-kelas lain yang dipandang sebagai satu kelas. Kelas untuk suatu sample data dapat langsung ditentukan dengan metode ini. Ketika sample data tidak dimasukkan ke dalam kelompok yang berisi sekumpulan kelas, tetapi ke dalam satu kelas yang spesifik, maka kelas tersebut adalah kelas dari sample data yang bersangkutan.

III. METODE

Data penelitian didapat dari Head pose Image database INRIA (Gambar 2) Rhône-Alpes 2004 [11]. Database pose kepala ini terdiri dari 2790 monocular citra wajah dari 15 orang dengan banyak variasi pose kepala dari -90 sampai +90 derajat. Yang mana untuk setiap orang, 2 dari 93 jenis pose di sediakan. Tujuan setiap orang memiliki 2 gambar yang berbeda adalah untuk dapat digunakan sebagai train dan test algoritma. Desain penelitian yang digunakan dalam penelitian ini adalah one group pretest—posttest. Desain penelitian ini hanya menggunakan satu kelompok saja yang diberi perlakuan, sehingga tidak memerlukan kelompok kontrol.



Gambar 2 database pose kepala INRAData Masukan

Data yang digunakan pada penelitian adalah citra dari wajah manusia yang sedang menggerakkan kepalanya meraka. Data citra yang didapat berformat .JPG. dengan seting latar belakang ada ruangang dengan pencahayaan yang baik.

Pose kepala yang dideteksi merupakan pose kepala yang telah di kelompokkan menjadi 93 jenis pose. Data citra yang didapat dilakukan pra proses secara terpisah pemotongan citra wajah secara manual dengan tujaun mendapat semua citra wajah secara jelas

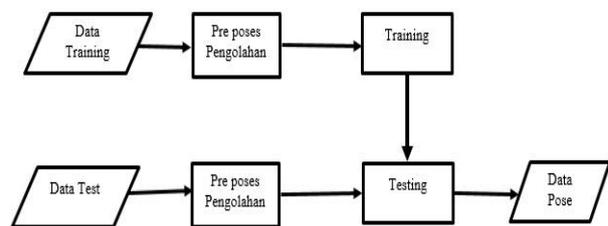
analisis dan perancangan sistem yang akan dibangun dengan tahapan perancangan sistem yang akan di buat sebagai mana pada Gambar 3.

1) *Pra proses pengolahan*: Pada proses ini pengolahan citra agar dapat di baca pada sistem, dalam hal ini citra mengalami konversi warna menjadi grayscale dan resize agar ukuran manjadi seragam setelah itu pada citra di lakukan ekstraksi fitur menggunakan HOG

2) *SVM*: Digunakan SVM dengan strategy one-versus-one sebagai klasifikasi yang digunakan dengan dimana di setiap citra yang dilatih dan ditest memiliki kemungkinan lebih dari satu kelas.

a) *Training*: Tahapan ini adalah tahap pelatihan data yang nantinya akan menghasilkan sebuah model yang nantinya akan menajdi dasar untuk melakukan tahapan testing

b) *Testing*: Tahapan ini adalah tahap pengecekan dengan model yang telah diperoleh dari training. Testing ini menghasilkan output berupa citra dengan pindaian muka dengan perubahan pose wajah pada test.



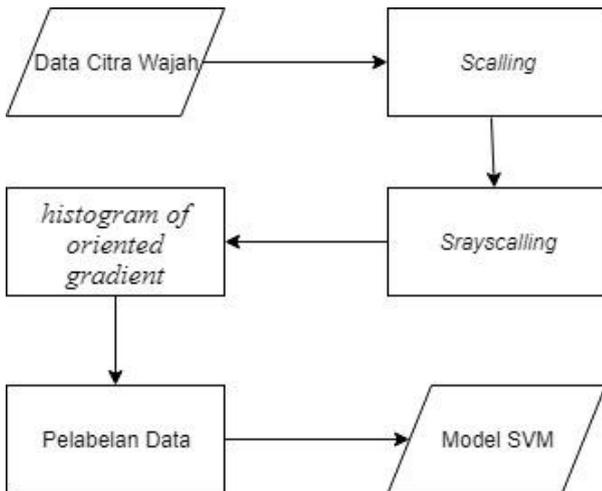
ambar 3 Proses analisis dan perancangan

IV. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pada tahap awal dari pembuatan sistem ini adalah pembuatan model, bilamana akan dilakukan pelatihan dari dari 2790 citra wajah yang telah disediakan terlebih dahulu. Dalam pelatihan dilakukan dengan SVM dengan tahapan pra proses dari mulai *scalling* sampai dengan ekstraksi fitur menggunakan HOG.

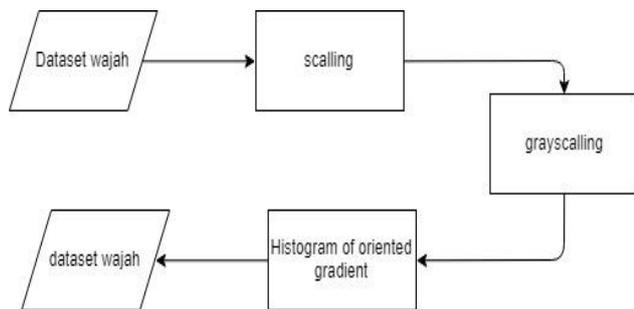
Berikut Gambar 4 adalah Alur pembuatan model dengan data citra wajah yang digunakan:



Gambar 4 Diagram Alur Pembuatan Model

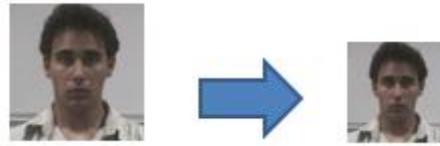
Selanjutnya tahap Pre proses, bertujuan untuk mempersiapkan citra data training dan data testing yang di proses oleh sistem agar informasi yang terkandung didalamnya dapat digunakan untuk proses pengolahan selanjutnya.

Alur pra proses data dapat digambarkan dalam diagram alur pada Gambar 5:



Gambar 5 Alur pra proses data

Proses *scalling* yaitu proses merubah ukuran citra sama dengan ukuran citra yang di inginkan dengan data pada citra tersebut tetap sama.pada proses scalling data set diubah ukurannya menjadi ukuran 100x100 Gambar 6 menunjukkan hasil citra dengan proses scalling.



Gambar 6 Proses scalling

Proses grayscaling adalah proses mengubah citra berwarna menjadi citra grayscale. Perubahan grayscale ini merubah channel gambar yang mana channel gambar berwarna adalah 3 dan grayscaling adalah 1. Gambar 7 menunjukkan hasil poses grayscaling



Gambar 7 Proses grayscaling

Dalam proses ekstraksi fitur, metode yang digunakan adalah Histogram Of Oriented Gradients(HOG). Parameter yang digunakan dalam ekstraksi fitur adalah winsize = (100,100), blockSize = (8,8), blockStride = (4,4), cellSize = (4,4), nbins = 9. Gambar 8 menunjukkan penggambaran hasil proses HOG.

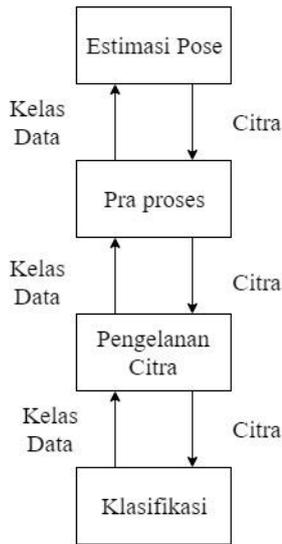


Gambar 8 proses ekstraksi fitur

Proses pelabelan dalam setiap data citra memiliki label yang akan digunakan sebagai parameter untuk training dan merupakan symbol dari tiap kelas. Dalam proses image datase yang telah di sediakan di kelompokan berdasarkan tipe posenya.

Proses klasifikasi ini bertujuan untuk mengambil ciri pada citra data training dan citra yang diuji. Pengambilan ciri pada citra data training dan citra yang diuji menggunakan metode Support Vector Machine. Pseudocode untuk metode ini menggunakan library yang disediakan oleh Scikit-Learn. Parameter yang digunakan adalah C = 1, decision_function_shape = ovo ,chahe_size = 2000, max_iter = 10000, dan gamma = auto.

Perancangan arsitektur aplikasi yang dibangun dimaksudkan untuk menjelaskan secara umum tentang sistem yang dibuat. Dimulai dengan jenis sistem yang dibangun, terdapat proses utama yaitu pengenalan dengan masukan gambar dan juga proses pendukung yaitu pra proses dan klasifikasi data. Berikut Gambar 9 merupakan rancangan arsitektur dari aplikasi.



Gambar 9 Alur Sistem

B. Pengujian

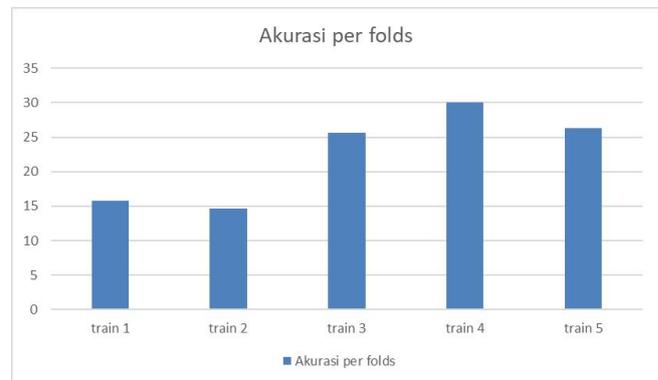
Dalam penelitian ini terdapat beberapa skenario yang akan dilakukan dalam pengujian. Pengujian ini dimaksud untuk mengetahui akurasi sistem estimasi pose kepala menggunakan metode Support Vector Machine. Berikut langkah-langkah yang dilakukan dalam pengujian :

1. Dataset terdiri dari 93 kelas data yaitu 93 jenis pose kepala. Masing masing kelas data terdapat 30 data citra yang menjadikan dataset berjumlah 2790 data citra.
2. Menghitung rata-rata keakuratan menggunakan cross validation sebanyak 5 folds. Dataset dibagi lima bagian untuk lima kali pengujian tiap literasinya. Masing—masing bagian berjumlah 558 data yang di ambil secara berututan sesuai jenis label yang mencakup semua kelas.
3. Pengujian pertama menggunakan bagian fold pertama sebagai data testing, dari training hasil pembagian dari fold pertama.
4. Pengujian kedua menggunakan bagian fold kedua sebagai data testing, dari training hasil pembagian dari fold kedua.
5. Pengujian ketiga menggunakan bagian fold ketiga sebagai data testing dari training hasil pembagian dari fold ketiga.
6. Pengujian keempat menggunakan bagian fold keempat sebagai data testing dari training hasil pembagian dari fold keempat.
7. Pengujian kelima menggunakan bagian fold kelima sebagai data testing dari training hasil pembagian dari fold kelima.
8. Pengujian tiap objek.

Masing-masing data pengujian dihitung akurasi nya

Pada Tabel 1 atas menunjukan akurasi hasil dari pengenalan menggunakan 5 folds cross validation. Dengan data training yang telah disediakan dihasilkan rata-rata benar sebesar 22,5% dan rata-rata waktu proses

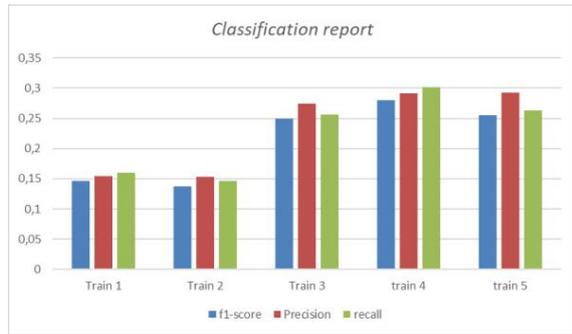
			Waktu Proses (detik)
1	1	15,77%	206
2	2	14,69%	439
3	3	25,62%	671
4	4	30,1%	903
5	5	26,34%	1135
Rata-rata		22,5%	670,8



Gambar 10 Grafik rata-rata pengujian

Tabel 2 menunjukan hasil *classification report* untuk tiap hasil pengujian, dengan hasil *f1-score* tertinggi (0,28), *precision* tertinggi (0,29), dan *recall* tertinggi (0,30), serta rata-rata untuk *f1-score* adalah (0,21), *precision* adalah (0,23), dan *recall* adalah (0,22). Pada grafik yang ditampilkan pada Gambar 11 terlihat perubahan dari nilai yang *f1-score*, *precision*, dan *recall* pada saat melakukan *training*.

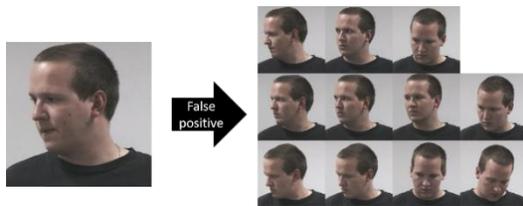
		precision	recall
Train 1	0,15	0,15	0,16
Train 2	0,14	0,15	0,15
Train 3	0,25	0,27	0,25
Train 4	0,28	0,29	0,30
Train 5	0,25	0,29	0,26
Rata-rata	0,21	0,23	0,22



Gambar 11 Grafik *classification report* per fold

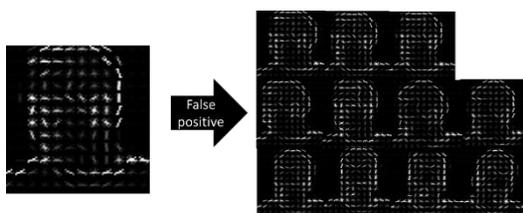
A. Pembahasan

Dari semua citra wajah yang telah diuji ternyata tidak semua kelas pose kepala dapat dikenali dengan baik. Dilihat dari hasil *classification report*, kelas +90+0 selalu memiliki nilai *f1-score*, *precision*, dan *recall* tertinggi, kalau dilihat secara visual beberapa citra dari kelas tersebut cenderung berbeda dengan data citra kepala yang lain. Untuk kelas yang mendapatkan nilai *classification report* kecil seperti -15+45, -15+60, -15+75, dan -15+90 bila dilihat secara visual terdapat perbedaan yang cukup kecil membuat sistem sering *false positive* seperti yang ditampilkan pada Gambar 12.



Gambar 12 kelas -15+60 memberi false positive pada kelas lain

Dengan melihat *confusion matrix* kelas -15+60 memberikan banyak *false positive* pada kelas lain, yang mana *true positive* untuk kelas ini 0. Seperti pada gambar diatas kelas -15+60 memberi *false positive* terhadap 11 kelas, dengan membandingkan hasil ekstraksi fitur terlihat bahwa dalam beberapa kelas, perbedaan pada masing-masing kelas tersebut terlihat kecil membuat sulit untuk dibedakan jika data *training* yang digunakan sedikit seperti yang diperlihatkan pada Gambar 13.



Gambar 13 perbandingan kelas -15+60 hasil ekstraksi HOG dengan kelas yang lain

V. KESIMPULAN

Berdasarkan hasil implementasi dan pengujian metode Support Vector Machine Untuk klasifikasi dan analisis pose kepala, Maka dapat diambil kesimpulan sebagai berikut :

- 1) Dari hasil pengujian yang dilakukan akurasi pengklasifikasian dalam estimasi pose kepala menggunakan head pose database dari INRA Rhône-Alpes yang mana terdapat 93 kelas sebanyak 22%.
- 2) Dengan akurasi sebesar 22% untuk estimasi pose kepala dengan metode Support Vector Machine sebagai klasifikasi dan Histogram of oriented gradients sebagai ekstraksi fitur yang di implementasikan dalam system tidak dapat digunakan di dunia nyata.
- Adapun saran yang diberikan untuk penelitian selanjutnya diantaranya adalah :
 - 1) Menggunakan dan menambah pemakaian algoritma lain untuk membandingkan dan meningkatkan akurasi pengklasifikasian.
 - 2) Menggunakan metode lain untuk ekstraksi fitur selain Histogram of oriented gradients.
 - 3) Menyusun dan menganalisis ulang data training sebagai masukan dengan menambah setiap varian dari tiap kelasnya agar meningkatkan akurasi pengklasifikasian, atau mengurangi jumlah kelas yang digunakan dengan menggolongkan yang lebih baik. Dari dataset yang digunakan.

REFERENSI

- [1] S. Wu, J. Liang dan J. Ho, "Head pose estimation and its application in TV viewers' behavior analysis," dalam *Canadian Conf. on Electrical and Computer Engineering (CCECE)*, Vancouver, Canada, 2016.
- [2] R. Stiefelhagen, M. Finke, J. Yang dan A. Waibe, "From gaze to focus of attention," dalam *In Visual Information and Information Systems*, Springer, 1999, p. 765–772.
- [3] S. R. Langton, H. Honeyman dan E. Tessler, "The influence of head contour and nose angle on the perception of eye-gaze direction," dalam *Perception & psychophysics*, 2004, p. 66(5):752–771.
- [4] K. Smith, S. O. Ba, J.-M. Odobez dan D. Gatica-Perez, "Tracking the visual focus of attention for a varying number of wandering people," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 7, p. 1212–1229, 2008.
- [5] INTEL, "Intel® Integrated Performance," dalam *Developer Reference, Volume 2: Image Processing*, INTEL, 2019, pp. 949 - 955.
- [6] C. Cortes dan V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273-297, 1995.
- [7] L. B. D. Cahyo, "IMPLEMENTASI METODE SUPPORT VECTOR MACHINE UNTUK MELAKUKAN KLASIFIKASI PADA DATA BIOINFORMATIKA. YOGYAKARTA : UNIVERSITAS ISLAM INDONESIA," 2018.
- [8] P. W. Angriyasa, "KLASIFIKASI TEKANAN DALAM RONGGA KEPALA MENGGUNAKAN SUPPORT VECTOR MACHINE

SEQUENTIAL,” UNIVERSITAS INDONESIA, DEPOK , 2011.

- [9] G. Bradski dan A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*, O’Reilly Media, Inc, 2008.
- [10] P. e. al, “Scikit-learn: Machine Learning in Python,” *JMLR*, vol. 12, pp. pp. 2825-2830, 2011.
- [11] N. Gourie, D. Hall dan J. L. Crowley, “Estimating Face Orientation from Robust Detection of Salient Facial Features,” dalam *International Workshop on Visual Observation of Deictic Gestures*, Cambridge, UK, 2004.