

## LITERATUR REVIEW MULTIMODAL AFFECTIVE SPEECH AND EMOTIONAL EXPRESSION

**Muhamad Fatchan<sup>1)</sup>, Andri Firmansyah<sup>2)</sup> Wahyu Hadikristanto<sup>3)</sup>**

<sup>1</sup>Teknik Informatika, STT Pelita Bangsa

<sup>2</sup>Cikarang, Bekasi, 17148

E-mail : fatchan@pelitabangsa.ac.id<sup>1)</sup>, [andrifirmansyah@pelitabangsa.ac.id](mailto:andrifirmansyah@pelitabangsa.ac.id)<sup>2)</sup>,

wahyuhadi@pelitabangsa.ac.id<sup>3)</sup>

### *Abstract*

Affective computing in a new field that is closely related to human and computer interaction, the result of developing intelligent systems that are capable of recognizing human emotional data processing, the rapid development of computational technology in terms of hardware and software demands to meet the service needs of users, various computational technologies which is now easily available like smartphones and gadgets. Giving training skills to computers to be able to better interact with users to handle users' emotional communication aspects, for this reason machine sensing training cannot be separated from advanced service needs for users, this research focuses on advanced computing such as the introduction of facial expression emotions the user continues to develop the issue of recognizing facial effects in humans by extracting features from different human images.

**Keyword :** *Affectif, Introduction, Facial Expression, Artificial Intelligence, Human Computing and Interaction*

### *Abstrak*

Computing affective bidang baru yang berhubungan erat dengan interaksi manusia dan computer, hasil dari pengembangan system cerdas yang mampu mengenali dalam memproses data emosi manusia, semakin cepatnya perkembangan teknologi komputasi baik dari sisi hardware dan software menuntut untuk menjawab kebutuhan layanan terhadap pengguna, berbagai macam teknologi komputasi yang sekarang mudah didapat seperti smartphone dan gadget. Memberikan latihan keterampilan terhadap computer untuk bisa lebih berinteraksi dengan pengguna untuk penanganan aspek komunikasi emosi pengguna, dengan alasan ini pelatihan penginderaan terhadap mesin tidak bisa terpisahkan dengan kebutuhan layanan tingkat lanjut terhadap pengguna, presearch ini focus pada komputasi tingkat lanjut seperti pengenalan emosi ekspresi wajah pengguna untuk terus ikut mengembangkan isu pengenalan efek wajah pada manusia dengan mengekstraksi fitur dari gambar manusia yang berbeda.

**Keyword :** *Pengenalan Affective, Ekpresi Wajah, Kecerdasan Buatan, Interaksi Manusia dan Komputer*

### 1. PENDAHULUAN

Berfokus *affective computing* adalah mengenali pola emosional manusia dalam berinteraksi dengan computer issue ini sudah cukup lama kurang lebih dua dekade sejak kemunculannya di tahun 1990 (Picard, 1997). Dari sisi psikologi emosi di pengaruhi suasana hati [1], dalam *research* ini istilah yang mempengaruhi suasana hati dan emosi akan sering digunakan dalam upaya konsep referensi (ekkekakis, 2013). Paradigma komputasi efeksti telah bergesaer ke *holistic* tentang kecerdasan manusia (Albercht, 2006), emosi wajah adalah factor penting dalam membantu komunikasi untuk bisa menyimpulkan keadaan seseorang seperti gembira, sedih menggunakan ekspresi wajah dan suara dari berbagai survei [2] penyajian *research* yang dihadapi dari latar belakang adalah pengenalan emosi manusia dari ekspresi wajah dan ekspresi

suara, untuk mencapai pada tujuan bagian. Membahas pentingnya komputasi afektif dalam bidang hubungan manusia dan komputer, bagian merinci tujuan dalam mengerjakan *research* ini, terahir menyajikan struktur hasil *research* tentang *affective computing*.

Pengenakan ekspresi wajah bagian bidang terpenting *machine* dalam mengenali espresi wajah secara mendalam secara cepat, untuk expresi manusia melalui gambar wajah dengan mengekstraksi fitu yang ada di gambar wajah dengan *Active Shape Model* (ASM) metode ini sangat populer dalam mengekstraksi firut wajah, secara akurat, kelemahan dari metode ini adalah bergantung pada *factor* kecerahan gambar, ketajaman gambar untuk mendapatkan hasil yang baik maka di gabungkan dengan metode Gaussian dan dalam jurnal saya sitasi mengusulkan

pengenalan ekspresi wajah dengan metode ASM dan Radiasi Basis *Function Network* (RBFN) [3]

Affective Computing untuk peningkatan kinerja interaksi manusia dan computer dengan menggunakan teknik pemrosesan ucapan dengan menggunakan jaringan saraf tiruan yang mendukung pada *Super Vector Machine* (SVM) dan K Nearest Neighbor (KNN) untuk mendeteksi Tujuh keadaan affective. [4]

Tantangan yang dihadapi dalam penelitian ada dua yaitu penginderaan efek wajah dan pengindraan dialog atau speech, meskipun kedua cabang ini berbeda masalah pengenalan kedua model ini harus dipecahkan penelitian ini bertujuan untuk membawa perspektif baru ke affective computing, memecahkan efek wajah yang mempengaruhi dialog maupun sebaliknya dialog mempengaruhi efek wajah. Berikut beberapa uraian tujuan penelitian: menganalisa perkembangan *Affective computing focus* pada pengenalan ekspresi wajah dan dialog, menentukan dan memvalidasi untuk mendapatkan akurasi yang baik dalam penentuan ekspresi wajah berupa citra inputan dalam ekspresi emosi apapun, memberikan strategi dalam memvalidasi secara terperinci, mengevaluasi hasil klasifikasi untuk mendapatkan kekuatan system, menentukan metodologi pada penanganan masalah ekspresi wajah dan ekspresi dialog hingga bisa memadukan dua model yang berbeda, menyediakan mekanisme inputan dari sejumlah fitur yang berasal dari 2 model secara berkala, dan menerapkan metode guna mendeteksi secara otomatis

2. Metode Penelitian

Judul harus jelas dan singkat. Nama penulis dan afiliasinya. berfokus pada dua masalah yang harus di selesaikan dalam research ini yaitu penginderaan wajah dan speech guna membawa perspektif baru tentang gabungan penginderaan wajah dan speech, hasil penelitian ini bisa di jadikan pedoman baru system atau metodologi yang akan di kembangkan, dapat di lihat pada tabel 1.

Gambar wajah yang akan di implementasikan dalam penentuan affect harus mempunya karakter seperti berikut ini :

- Model gambar wajah yang akan di jadikan sebagai data input harus berbasis fitur.
- Efek wajah harus mendiskripsikan ekspresi emosi dengan seperti itu maka statu hasil keluaran sesuai dengan kebutuhan
- Sistem ini harus divalidasi dengan sumber wajah yang umum, yang mampu menganalisa ras, umur, jenis kelamin seperti dengan penggunaan multilayer perceptron

Tabel 1. Tabel Matrik Letaratur Review

Research	Metode 1	Metode 2	Metode 3	Metode 4	Metode 5
Analysis of EEG signals and facial expressions for continuous emotion detection	Long-short-term-memory recurrent neural networks (LSTM-RNN)	Continuous Conditional Random Fields (CCRF)	Support Vector Regression (SVR) and	Long-Short-Term-Memory Recurrent Neural Networks (BLSTM-RNN)	Long-Short-Term-Memory Recurrent Neural Networks (BLSTM-RNN)
Template Adaptation for Face Verification and Identification	VGG-Face deep convolutional neural network	Bilinear-CNN	SVM		
Fuzzy Histogram of Optical Flow Orientations for Micro-expression Recognition	Linier SVM	3 Point KNN	LDA		
Facial Emotional Expressions Recognition Based on Active Shape Model and Radial Basis Function Network	Active Shape Model (ASM)	Radial Basis Function Network (RBFN)			
Pengenalan Emosi Berdasarkan Ekspresi Mikro Menggunakan Metode Local Binary Pattern	KNN	Local binary Pattern			
Affective Computing as Complex Systems Science	Component not methode : Multimodal Sensory Agents Human	Component not methode : Nonlinear Nature of Emotions Meuleman	Component not methode : Emergent Process Property Referring		
Affective Computing using Speech Processing for Call Centre Applications	Neural Network	SVM	KNN classifier.		

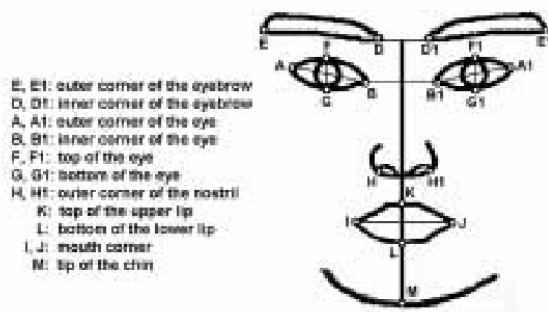
*Matrik Literatur Review* Bagian dari pembahasan review dari topik research yang sama dengan menggunakan metod yang sama maupun yang berbeda, guna menentukan Posisi peneliti untuk bisa menggunakan metode yang berbeda maupun mengningkatkan perform metode dalam peneleitian

3. HASIL DAN PEMBAHASAN

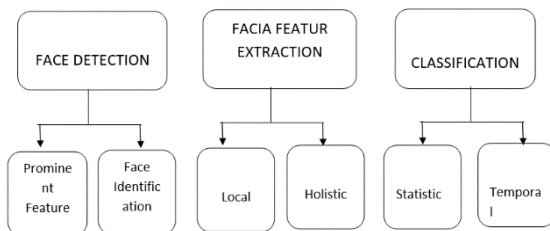
Pada bagian ini menentukan untuk mengklasifikasi secara otomatis pada gambar wajah hususnya pada ekspresi emosional

3.1. Preprocessing

Menyeleksi dan mengekstraksi fitur wajah sebagi inputan Menganalisa ekspresi wajah dengan mencari fitur dan mengukur serangkaian jarak sudut wajah antara satu fitur dengan fitur yang lain untuk diklasifikasikan dari ekspresi wajah yang berbeda – beda dengan pendekatan ini di harapkan bisa mengklasifikasikan jarak dan sudut yang tepat. Dari gambar 1, di dapatkan 19 hasil ekstraksi fitur 2, mewakili pusat mata (titik X dan Y) dan 1 mewakili titik lubang hidung. Kerangka Kerja Sistem Ekspresi Wajah Otomatis dapat di lihat pada gambar 2.



Gambar 1. Feature Point

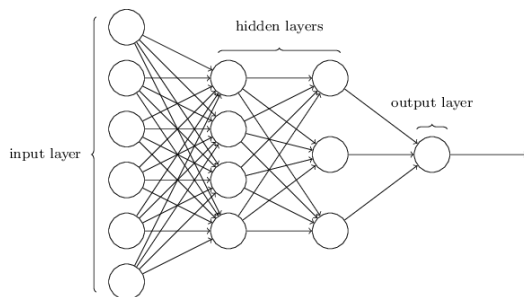


Gambar 2. Kerangka Kerja Sistem Ekspresi Wajah Otomatis

**3.2. Pemilihan Metode**

Dalam pemilihan *metode* terbaik peneliti menggunakan tool Matlab tool ini menyediakan perangkat set algoritma untuk pembelajaran mesin dalam penambahan data, peneliti menggunakan 5 Metode perbandingan perhitungan. Antara lain : *Multilayer Perceptron*, *K-NN*, *Naïve bayes*, *Super vector machine*, *Logistic Regression*.

- a. *Multilayer Perceptron* adalah salah satu permodelan dalam teknologi jaringan saraf tiruan (JST) dengan karakteristik memiliki nilai bobot yang lebih baik dari pada pemodelan yang lain, sehingga menghasilkan klasifikasi yang lebih akurat pula. Multi layer perceptron diperkenalkan pertama kali pada tahun 1969 oleh S. Papert dan M. Minsky, peserta dapat di lihat pada gambar 4.



Gambar 4. Multi Layer Perceptron

- b. *K-NN* adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Data pembelajaran

diprojeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi data pembelajaran. Sebuah titik pada ruang ini ditandai kelas *c* jika kelas *c* merupakan klasifikasi yang paling banyak ditemui pada *k* buah tetangga terdekat titik tersebut. Dekat atau jauhnya tetangga biasanya seperti rumus (1), dihitung berdasarkan jarak Euclidean.

$$= \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2}$$

$$= \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \dots \dots \dots (1)$$

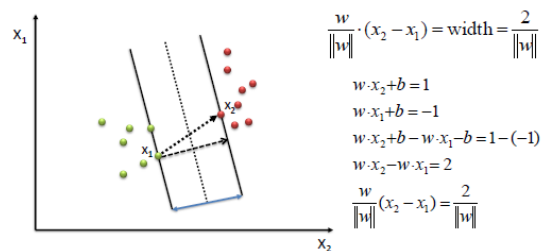
- c. *Naïve bayes* adalah merupakan sebuah metoda klasifikasi menggunakan metode probabilitas dan statistik yg dikemukakan oleh ilmuwan Inggris Thomas Bayes. Algoritma Naive Bayes memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dr *Naïve Bayes Classifier* ini adalah asumsi yg sangat kuat (naif) akan independensi dari masing-masing kondisi / kejadian, seperti rumus (2),

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \dots \dots \dots (2)$$

Likelihood
Class Prior Probability  
Posterior Probability
Predictor Prior Probability

$$P(c | X) = P(x_1 | c) \times P(x_2 | c) \times \dots \times P(x_n | c) \times P(c)$$

- d. *Super Vector Machine (SVM)* adalah suatu teknik untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi. SVM berada dalam satu kelas dengan Artificial Neural Network (ANN) dalam hal fungsi dan kondisi permasalahan yang bisa diselesaikan.



- e. *Logistic Regression* merupakan alternative uji jika asumsi multivariate normal distribution pada variabel bebasnya tidak bisa terpenuhi ketika akan dilakukan analisis diskriminan.

Tidak terpenuhinya asumsi ini dikarenakan variabel bebas merupakan campuran antara variabel kontinyu (metric) dan kategorial (non metric), seperti rumus (3),

$$P = \frac{e^{a+bX}}{1 + e^{a+bX}} \dots\dots\dots (3)$$

**3.3. Dataset Learning**

DATABASE	CHANNELS	METHODE	SIZE	LABELLING	ACCEBILITY
RTCO4 [8]	Facial expressions and audio	Natural subject were interviewed to describe their childhood experience.	60 adults, each interview lasts 30-60 minutes	FACS (33 AU(s))	FREE
RU-FACT [9]	Facial expressions and audio	experience. Natural subjects were tried to convince the interviewers they were telling the truth	100 adults	FACS (33 AU(s))	FREE

Gambar 4. Dataset Multimodal Affect

**3.4. Tool Penelitian**

Dalam melakukan research ini menggunakan MATLAB sebagai alat bantu untuk proses feature ekstrasi dan proses klasifikasi, tool ini sudah dilengkapi dengan beberapa library algoritma yang di butuhkan untuk bisa melakukan pemrosesan sesuai dengan kebutuhan. Untuk pemisahan gambar frame per frame menggunakan windows movie maker dan untuk memisahkan suara menggunakan aplikasi Adobe Audition.

**3.5. Tahapan Penelitian Deteksi Ekspresi dan Dialog**

- a. Proses Pengambilan Gambar dan dialog  
 Penelitian deteksi ekspresi wajah dan dialog atau speech dengan melakukan proses pengambilan gambar yang dapat di unduh di internet lewat web penyedia dataset seperti di [www.uci.org](http://www.uci.org), [www.kaggle.com](http://www.kaggle.com) dan yang lainnya.
- b. Preprocessing Gambar dan suara  
 Setelah mendapatkan proses pengambilan gambar langkah berikutnya adalah Malakukan pemotongan, merisize gambar dan memisahkan gambar dengan suara frame per frame dengan menggunakan Aplikasi Adobe Audition

**3.6. Klasifikasi Gambar dan Suara**

Berikutnya hasil dari pemisahan selanjutnya melakukan proses feature ekstraksi yang akan di olah menjadi metodi Multimodal expression, untuk

selanjutnya dilakukan pengujian di computer untuk mendapatkan evaluasi

**5. KESIMPULAN**

Bahwa pada hasil dari review beberapa jurnal yang terkait maka peneliti menentukan penambahan metode yang sebelumnya belum pernah di implementasikan pada penelitian Model Multimodal Ekspresi.

**6. DAFTAR PUSTAKA**

- [1] C. J. Beedie, P. C. Terry, and A. M. Lane, "Distinctions between emotion and mood," *Cogn. Emot.*, vol. 19, no. 6, pp. 847–878, 2005.
- [2] K. Kaulard, D. W. Cunningham, H. H. Bülhoff, and C. Wallraven, "The MPI facial expression database - a validated database of emotional and conversational facial expressions," *PLoS One*, vol. 7, no. 3, 2012.
- [3] E. Setyati, Y. K. Suprpto, and M. H. Purnomo, "Facial emotional expressions recognition based on active shape model and radial basis function network," *CIMSA 2012 - 2012 IEEE Int. Conf. Comput. Intell. Meas. Syst. Appl. Proc.*, no. July 2012, pp. 41–46, 2012.
- [4] R. K. Gowda, V. Nimbalkar, R. Lavanya, S. Lalitha, and S. Tripathi, "Affective computing using speech processing for call centre applications," *2017 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2017*, vol. 2017–Janua, pp. 766–771, 2017.
- [5] T. I. Hupont, *A ective Computing: emotional facial sensing and multimodal fusion*. 2010.
- [6] M. Pantic and L. J. M. Rothkrantz, "Facial gesture recognition in face image sequences: a study on facial gestures typical for speech articulation," no. November 2002, p. 6, 2003.
- [7] D. Jaglan, "A Literature Review : Various Facial Expression," vol. 7, no. 5, pp. 72–77, 2017.
- [8] G. I. Roisman, J. L. Tsai, and K. H. S. Chiang, "The emotional integration of childhood experience: Physiological, facial expressive, and self-reported emotional response during the adult attachment interview," *Dev. Psychol.*, vol. 40, no. 5, pp. 776–789, 2004.
- [9] M. S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, "Computer Vision and Pattern Recognition 2005 Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behavior," 2005.