

Sistem Deteksi Penggunaan masker (*Face Mask Detection*) Menggunakan Algoritma Deep Learning YOLOv4

Bagus Hardiansyah¹, Aidil Primasetya²

^{1,2}Teknik Informatika, Fakultas Teknik, Universitas 17 Agustus 1945 Surabaya
E-mail: ^{*1}bagushardiansyah@untag-sby.ac.id, ²aidilprimasetya@untag-sby.ac.id

Abstrak – penelitian terkait deteksi wajah adalah salah satu permasalahan yang terjadi di dunia *Computer Vision*, khususnya yang akhir-akhir ini terjadinya perubahan hidup yang signifikan yaitu dengan menggunakan masker pada wajah. Masker pada wajah adalah bagian dari fitur atau sebagai informasi bahwa kita menggunakan masker diwajah. Disisi lain beberapa studi mengindikasikan untuk deteksi wajah dengan memanfaatkan ekstraksi fitur memberika akurasi yang tinggi, khususnya jika bagian masker menutupi bagian bawah mata. Oleh karena itu, penelitian ini melakukan teknik training memanfaatkan ekstraksi fitur dalam metode deep learning. Oleh karena itu, kami menggunakan Primary dataset yang di olah sendiri dengan jumlah 399 subject orang Indonesia. Dataset ini juga berupa video yang telah dirubah dalam frame per second (FPS). Metodologi dari penelitian sebelumnya juga kami terapkan yaitu YOLOv4 untuk metode deteksi wajah, pre-trained menggunakan VGGFace model sebagai pengembangan metode ekstraksi fitur dan artificial neural network sebagai metode klasifikasi. Ketika menggunakan metode teknik training yang kami lakukan maka hasil testing didapat 99.55%. menunjukkan bahwa hasil performa cukup baik ketika menyesuaikan dengan teknik training dalam sistem deteksi penggunaan masker diwajah.

Kata Kunci — sistem deteksi wajah, sistem deteksi penggunaan masker, Deep learning.

1. PENDAHULUAN

Penelitian di bidang *Artificial Intelligent* (AI) sangat berkembang sangat cepat dalam area *Computer Vision*. Oleh karena itu, dalam area *Computer Vision* (CV) adalah salah satu topik yang paling populer yang digeluti oleh peneliti dalam AI. Salah satu topik penelitian CV yaitu *face recognition*, *face expression* dan *face mask*. Maka dengan kemampuan deteksi dari seorang manusia dengan fokus pada wajah untuk deteksi wajah dengan sistem pada tahap awal maka banyak manfaat yang akan dilakukan pada eksperimen berikutnya. Oleh karena itu, banyak contoh yang bisa dilihat dalam aplikasinya yaitu : disekolah, perusahaan dan lain-lain, secara aplikasi bisa diterapkan sistem presensi pada *face verification* contoh dari *face recognition* sistem dapat memverifikasi untuk secara personal untuk *security access*, katakanlah aplikasinya bisa menggunakan *smartphone lock*, *payment*, dan lain-lain

Sistem *face detection* dalam beberapa penelitian menunjukkan hasil performa yang sangat baik. [1] telah membuat sistem pengenalan wajah dengan menggunakan dataset yang sangat banyak dan beragam. Dataset tersebut berjumlah 2.6M image dan lebih dari 2.6K adalah manusia. Untuk membandingkan performa dari sistem tersebut mereka menggunakan LFT dan YTF dataset untuk *benchmarking*. Arsitektur dari VGG face, sistem mereka mendapatkan hasil validasi 98.95% akurasi dari LFW dataset dan 97.03% akurasi dari YTF dataset. [2] menyatakan suatu metode sistem *face detection* dan mendapatkan 95.97% akurasi dalam AR Face dataset dengan 120 orang dan 97.20% akurasi dalam VTU-BEC-DB dengan multi dataset. Oleh karena itu, contoh dalam sistem *face recognition* dapat di implementasikan untuk sistem presensi kehadiran yang telah diterapkan oleh [3]. Dari penjelasan tersebut dimana mengimplementasikan sistem RFID untuk presensi kehadiran karyawan, berdasarkan aplikasi tersebut didapat 95.02% dari semua akurasi dengan dataset yang cukup. Dalam penelitian yang lain [4] sistem *face recognition* untuk sistem presensi kehadiran dalam lingkungan sekolah. Dari hasil tersebut menunjukkan sistem presensi kehadiran dalam mengelola mendapatkan akurasi 97.29%. dalam beberapa penelitian yang lain telah dilakukan survei yang merujuk langsung *face recognition* dalam beberapa tahun terakhir [5], [6], [7], [8], [9] dan [10]. Berdasarkan dari hasil dari survei rujukan penelitian tersebut, dalam pengembangan sistem *face recognition* didapatkan hasil performa yang sangat baik dengan berbagai variasi metode yang dilakukan. Tampaknya dalam sistem *face recognition* menjadi salah satu dari hasil topic tersebut dalam pengembangan AI. Namun, dalam hal ini tidak cukup untuk celah dalam penelitian untuk melakukan penelitian yang terkait sistem *face recognition* dengan kondisi dan kasus yang berbeda-beda.

Dalam COVID-19 mulai muncul dari akhir 2019 menjadi salah satu pandemi paling besar dalam sejarah umat manusia. Sehingga efek dari pandemic sangat viral dan sangat berdampak bagi umat manusia, dampak

ekonomi dunia, sosial, akademik, dan gaya hidup manusia yang sangat berubah. Tujuan penelitian ini adalah untuk mencegah kita untuk terinfeksi COVID-19 dan meminimalkan penyebaran COVID-19. Oleh karena itu, dari dampak tersebut dengan penggunaan masker di wajah dapat meminimalkan penyebaran COVID-19, khususnya sistem *face recognition* sangat bermanfaat dalam situasi pandemi saat ini.

Penelitian terkait [11] sistem penggunaan masker di wajah, sistem tersebut menunjukkan bahwa hasil akurasi 97% dari input image wajah menggunakan masker untuk data training. Namun, hasil dari performa lainnya didapat akurasi 79% tanpa menggunakan masker di wajah. Penelitian lainnya yang terkait dengan menambahkan penggunaan masker di wajah untuk data training dengan manajemen sistem menunjukkan peningkatan berdasarkan validasi *true positif rate* 38% [12]. Dalam penelitian lainnya yang terkait, dengan simulasi augmentasi dari dataset wajah dengan menambahkan penggunaan masker di wajah dari dataset untuk data training model. Dari eksperimen tersebut sistem deteksi penggunaan masker di wajah dapat hasil performa yang maksimal tanpa menggunakan image penggunaan masker di wajah dengan data training model [13].

Dalam citra penggunaan masker di wajah dimana ekstraksi fitur pada wajah dilakukan, yang mana adalah untuk mengenali wajah yang menggunakan masker di wajah dan membedakan dengan wajah yang tidak menggunakan masker. Dari teknik metode tersebut diharapkan akan dapat meningkatkan performa pada saat dilakukan training dan testing waktu melakukan eksperimen. Oleh karena itu, ada beberapa penelitian terkait yaitu eksperimen *deep face recognition using partial faces*, dalam eksperimen tersebut memanfaatkan separuh dari wajah dengan membuat area dibawah mata menjadi berwarna hitam atau dengan kata lain pixel bernilai 0, maka didapat akurasi dengan deteksi yang sangat tinggi yaitu lebih dari 90% [14]. Dengan penelitian terkait juga menerapkan teknik segmentasi yang lebih bervariasi dan tetap menggunakan / membuat hitam di area wajah tepatnya mata kebawah dan mata keatas dibuat hitam atau pixel bernilai 0, maka mendapatkan akurasi yang cukup tinggi [15]. Penelitian terkait lainnya berfokus pada data training dengan teknik membuat hitam pada bagian atas wajah mata citra mendapatkan akurasi 88.09% dengan dataset SMFRD dan deteksi penggunaan masker di wajah didapat akurasi 91.03% dengan dataset RMFRD [16].

Algoritma You only look once (YOLO) adalah merupakan Real Object Detection yang baru-baru ini sangat populer untuk dikembangkan, YOLO menggunakan pendekatan yang sangat berbeda, yakni menerapkan jaringan syaraf tunggal pada keseluruhan gambar. Jaringan ini akan membagi gambar menjadi wilayah-wilayah kemudian memprediksi kotak pembatas dan probabilitas, setiap kotak wilayah pembatas probabilitasnya untuk mengklasifikasikan sebagai objek atau bukan. kotak pembatas dengan nilai yang paling tinggi untuk di jadikan sebagai pemisah objek satu dengan yang lain dan Daerah dengan citra yang diberi skor paling tinggi akan di deteksi sebagai sebuah pendeteksian [17].

Penelitian ini dalam aplikasinya yaitu membuat deteksi penggunaan masker di wajah menggunakan primary dataset yang dapatkan sendiri yaitu orang Indonesia dengan kategori dalam satu orang kami mendapatkan dua citra : 1. wajah tidak menggunakan masker, 2. Wajah menggunakan masker. tujuan sistem deteksi dengan teknik ekstraksi fitur dengan membuat deteksi klasifikasi berdasarkan wajah menggunakan masker dan tidak menggunakan masker. Dalam Section 2, tujuan dari penelitian ini adalah menjelaskan dataset dan metodologi. Dalam Section 3 diskusi hasil eksperimen. Dalam Section 4 hasil dari penelitian.

2. METODE PENELITIAN

Dalam bab ini membahas apa yang kami kerjakan terkait metodologi yaitu sistem deteksi penggunaan masker (*face mask detection*) dengan *primary* dataset orang Indonesia dan metode *deep learning*. Penjelasan secara detail dijelaskan di bab ini.

2.1 You Only Look Once

You only look once (YOLO) menggunakan pendekatan Artificial Neural Network (ANN) untuk mendeteksi objek pada sebuah citra. Jaringan ini membagi citra menjadi beberapa wilayah dan memprediksi setiap kotak pembatas dan probabilitas untuk setiap wilayah pada setiap Kotak-kotak pembatas ini dengan setiap probabilitas yang diprediksi. YOLO memiliki beberapa kelebihan dengan sistem yang berorientasi pada classifier, terlihat dari seluruh citra pada saat dilakukan test dengan prediksi yang diinformasikan pada citra. Sebelum dilakukan proses pelatihan perlu dilakukan proses inialisasi terlebih dahulu untuk membentuk dataset. Oleh karena itu, setiap data memiliki nama kelas pada titik koordinat X objek, titik koordinat Y objek, panjang kotak pembatas, dan lebar kotak pembatas.

2.2 Rekontruksi Dataset

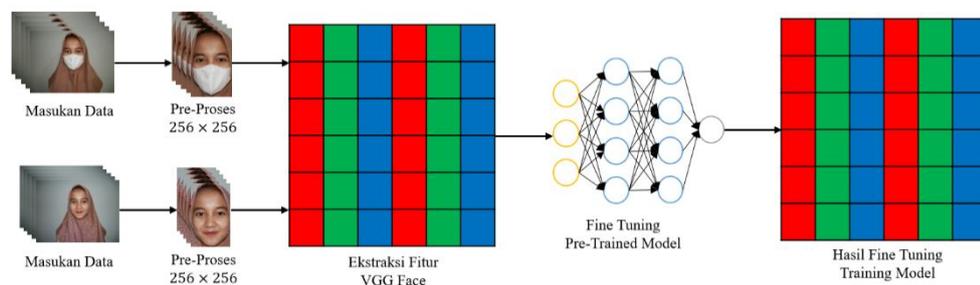
Dataset yang diolah adalah 399 citra orang Indonesia dan 6 video wajah orang Indonesia. tidak ada batasan umur dan kelamin dalam target penelitian, dalam pengambilan dataset untuk 399 citra berfokus pada area wajah menggunakan masker dan tidak menggunakan masker, sedangkan pada dataset video merekam secara acak dalam waktu yang berbeda. Teknik video dirubah dalam 30 *frame per second* (FPS) dengan 5 detik per frame, total citra yang didapat 1.326K. Dataset video yang didalamnya mengandung citra wajah yang menggunakan masker dan tidak menggunakan masker akan digunakan untuk data training sedangkan untuk testing menggunakan 399 citra orang Indonesia. Gambar 1 menunjukkan contoh dataset.



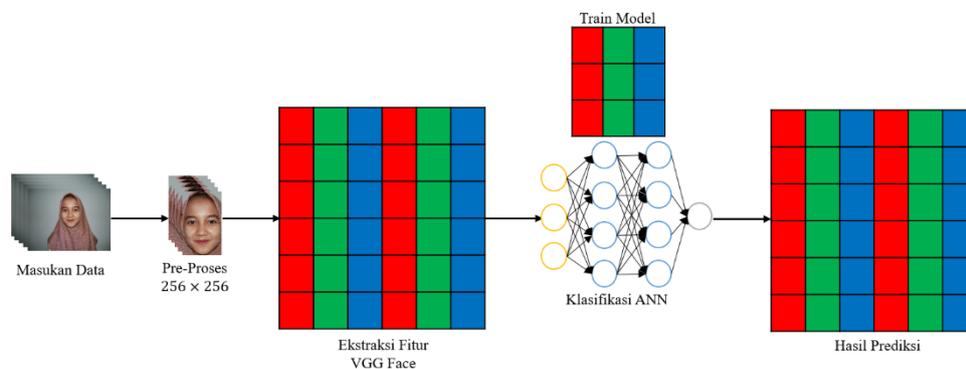
Gambar 1. Contoh Dataset Menggunakan Masker dan Tidak Menggunakan Masker

2.3 Kerangka Sistem

Kerangka sistem dibagi menjadi dua yaitu : Gambar 2 kerangka sistem training dan Gambar 3 kerangka sistem testing.



Gambar 2. Kerangka Sistem Training



Gambar 3. Kerangka Sistem Testing

YOLOv4 menerapkan arsitektur ANN, dimana YOLOv4 hanya menggunakan lapisan konvolusi dan lapisan pooling. Dalam proses tersebut dari hasil ekstraksi fitur dari citra dapat bisa membuat klasifikasi untuk memprediksi hasil yang sudah ditentukan berdasarkan mengenali deteksi wajah yang menggunakan masker dan tidak menggunakan masker hasil testing pada gambar 3. Untuk lapisan konvolusi terakhir disesuaikan dengan jumlah kelas dan jumlah kotak prediksi yang diinginkan. Untuk menghitung ukuran keluaran pada lapisan konvolusi terakhir dapat menggunakan rumus sebagai berikut [18]:

$$Y = S S B \times (5 + C) \dots\dots\dots(1)$$

2.4 Pre-Proses dan Deteksi Wajah

Sangat penting proses dalam sistem deteksi wajah, untuk mendapatkan deteksi wajah secara local pada bagian citra wajah sangat berpengaruh untuk langkah berikutnya yaitu ekstraksi fitur. Oleh karena itu, hasil proses untuk klasifikasi pixel akan sangat mempengaruhi tingkat akurasi baik secara fitur pada saat memotong / mengambil pada bagian wajah untuk identifikasi.

Untuk eksperiment tersebut sangat membantu bagi para peneliti untuk memaksimalkan proses pre-proses dan deteksi wajah terlebih dahulu. Pendekatan kami sangat terbukti untuk performa digunakan You Only Look once (YOLO) [18]. Dibeberapa tahun terakhir, YOLO sangat berkembang dimana ada beberapa meningkatkan implementasinya agar lebih mendapatkan performa yang lebih baik, dari YOLOv2 [19] ke YOLOv3[20].

Dari masukan citra untuk inialisasi dibagi menjadi *grid by grid* dari empat koordinat mengandung untuk setiap pixel. Batas posisi box untuk indikasi *x dan y* dalam posisi *center* untuk deteksi dan koordinat yang lain *w dan h* indikasi dari lebar dan tinggidari batas box. Dalam reflek *score* untuk membuat model, maka *score* untuk menentukan rumus berikut :

$$C(X) = Probability(X).Input Output_{prediction} \dots\dots\dots(2)$$

Equation 1. *X* inialisasi objek, *Probability* adalah per grid pixel dan *Input Output* adalah prediksi box dan *ground truth*.

2.5 Identifikasi Wajah

Beberapa proses terkait sistem deteksi wajah sangat penting untuk identifikasi wajah.identifikasi wajah sangat vital dalam *step by step* untuk sistem deteksi yang dapat memberikan prediksi dari label untuk setiap orang dari citra wajah. Oleh karena itu, ada satu step untuk ekstraksi fitur dimana performa untuk proses ekstraksi dari informasi citra wajah.

VGGFace dibuktikan memberikan hasil performa terbaik, khususnya pada deteksi dengan crop pada bagian bawah mata dan bagian atas mata diberikan nilai pixel 0, referensi dari penelitian sebelumnya [14] dan [15] terkait deteksi wajah. VGGFace untuk Pre-Ttrain model dari referensi [1] dimana train menggunakan dataset yang sangat besar. VVGFace sebagai ekstrasi fitur dapat digunakan untuk tahap selanjutnya untuk klasifikasi dari variasi metode, dengan contoh ANN dan metode klasifikasi konvensional lainnya. Untuk mendapatkan ekstraksi fitur akan lebih digunakan untuk klasifikasi pada tahap selanjutnya untuk digunakan ke metode ANN dengan beberapa modifikasi dan fungsi softmax.

3. HASIL DAN PEMBAHASAN

Penelitian ini yaitu sistem deteksi penggunaan masker dengan dataset dengan orang Indonesia yang telah kami dapatkan dan olah data. Dari penjelasan sebelumnya yaitu 399 citra wajah orang Indonesia dengan umur dan jenis kelamin yang berbeda untuk testing dan dataset video. Dari video FPS didapat hasil citra sebanyak 1.326K digunakan untuk training. Oleh karena itu, rata-rata dari ekperimen adalah sekitar 20% untuk testing. Bisa dilihat dari table 1 performa dari sistem deteksi penggunaan masker.

Tabel 1. Hasil Sistem Deteksi Penggunaan Masker

Eksperimen	Metode Ekstraksi Fitur	Metode Klasifikasi
		ANN
Training	VGGFace	100%
Validasi		100%
Testing		99.55%

Dari table 1 didapat hasil performa dari sistem deteksi penggunaan masker dengan hasil akurasi yang sangat baik. Semenjak digunakannya metode ekstraksi fitur VGGFace dimana pre-trained model dari dataset yang sangat besar mendapatkan hasil performa yang baik.



Gambar 4. Hasil Eksperimen Menggunakan Kamera Handphone

4. SIMPULAN

Dalam penelitian ini, kami menyajikan sistem deteksi penggunaan masker dengan menggunakan metode deep learning, kami memanfaatkan ekstraksi fitur dengan VGGFace pada arsitektur deep learning. Dalam mekanisme yang sama seperti pengelihat manusia, ketika seseorang mencoba untuk mendeteksi pengelihat terhadap orang lain, biasanya masker di wajah tidak selalu menjadi penting jika orang yang kita kenal adalah orang yang kita kenal. Disisi lain, secara aplikasi real time bisa menggunakan webcam, surveillance system. Untuk penelitian selanjutnya, penggunaan dataset dapat ditingkatkan dari pengumpulan subjeknya.

5. SARAN

Saran-saran dan tambahan adalah untuk sistem deteksi penggunaan masker sebaiknya dilakukan re-train agar pre-proses data lebih sempurna. Sehingga pemanfaatan kecerahan, kontras, augmentasi menjadi hasil yang lebih baik agar peningkatan variasi menjadi training model lebih akurat untuk hasil prediksi yang lebih *advance*.

DAFTAR PUSTAKA

- [1] Parkhi, Omkar M and Vedaldi, Andrea and Zisserman, Andrew, Deep face recognition, British Machine Vision Association, 2015.
- [2] Angadi, Shanmukhappa A and Hatture, Sanjeevakumar M, Face recognition through symbolic modeling of face graphs and texture, International Journal of Pattern Recognition and Artificial Intelligence, 33, 12, 1956008, 2019, World Scientific.
- [3] Arsenovic, Marko, et al. "FaceTime—Deep learning based face recognition attendance system." 2017 IEEE 15th International symposium on intelligent systems and informatics (SISY). IEEE, 2017.
- [4] Sunaryono, Dwi, Joko Siswanto, and Radityo Anggoro. "An android based course attendance system using face recognition." Journal of King Saud University-Computer and Information Sciences 33.3 (2021): 304-312.
- [5] Masi, Iacopo, et al. "Deep face recognition: A survey." 2018 31st SIBGRAPI conference on graphics, patterns and images (SIBGRAPI). IEEE, 2018.
- [6] Wang, Mei, and Weihong Deng. "Deep face recognition: A survey." Neurocomputing 429 (2021): 215-244.
- [7] Guo, Guodong, and Na Zhang. "A survey on deep learning based face recognition." Computer vision and image understanding 189 (2019): 102805.
- [8] Ding, Changxing, and Dacheng Tao. "A comprehensive survey on pose-invariant face recognition." ACM Transactions on intelligent systems and technology (TIST) 7.3 (2016): 1-42.
- [9] Li, Pei, et al. "Face recognition in low quality images: A survey." arXiv preprint arXiv:1805.11519 (2018).
- [10] William, Ivan, et al. "Face recognition using facenet (survey, performance test, and comparison)." 2019 fourth international conference on informatics and computing (ICIC). IEEE, 2019.
- [11] Mundial, Imran Qayyum, et al. "Towards facial recognition problem in COVID-19 pandemic." 2020 4rd International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM). IEEE, 2020.

- [12] Anwar, Aqeel, and Arijit Raychowdhury. "Masked face recognition for secure authentication." arXiv preprint arXiv:2008.11104 (2020).
- [13] Li, Yande, et al. "Cropping and attention based approach for masked face recognition." *Applied Intelligence* 51.5 (2021): 3012-3025.
- [14] Elmahmudi, Ali, and Hassan Ugail. "Experiments on deep face recognition using partial faces." 2018 international conference on cyberworlds (CW). IEEE, 2018.
- [15] Elmahmudi, Ali, and Hassan Ugail. "Deep face recognition using imperfect facial data." *Future Generation Computer Systems* 99 (2019): 213-225.
- [16] Hariri, Walid. "Efficient masked face recognition method during the covid-19 pandemic." *Signal, image and video processing* 16.3 (2022): 605-612.
- [17] Jupiyandi, Sisco, et al. "Pengembangan deteksi citra mobil untuk mengetahui jumlah tempat parkir menggunakan CUDA dan modified YOLO." *Jurnal Teknologi Informasi dan Ilmu Komputer* 6.4 (2019): 413-419.
- [18] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [19] Redmon, Joseph, and Ali Farhadi. "YOLO9000: better, faster, stronger." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- [20] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).