

TESIS

**MONOKULAR 3D REKONSTRUKSI WAJAH
MENGUNAKAN 3D MORPHABEL MODEL DAN
ELASTICFACE**

*Monocular 3D Face Reconstruction Using 3D Morphable Model
and ElasticFace*

**ABD SALAM AT TAQWA
D082211003**



**PROGRAM STUDI MAGISTER TEKNIK INFORMATIKA
DEPERTEMEN TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2023**

PENGAJUAN TESIS

**MONOKULAR 3D REKONSTRUKSI WAJAH
MENGUNAKAN 3D MORPHABEL MODEL DAN
ELASTICFACE**

Tesis

Sebagai Salah Satu Syarat Untuk Mencapai Gelar Magister
Program Studi Teknik Informatika

Disusun dan diajukan oleh



**ABD SALAM AT TAQWA
D082211003**

Kepada

**FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2023**

TESIS

MONOKULAR 3D REKONSTRUKSI WAJAH MENGGUNAKAN 3D MORPHABEL MODEL DAN ELASTICFACE

ABD SALAM AT TAQWA
D082211003

Telah dipertahankan di hadapan Panitia Ujian Tesis yang dibentuk dalam rangka penyelesaian studi pada Program Magister Teknik Informatika Fakultas Teknik Universitas Hasanuddin Pada tanggal 25 Oktober 2023 dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

Pembimbing Utama



Dr. Ir. Zahir Zainuddin, M.Sc.
NIP. 19640427 198910 1 002

Pembimbing Pendamping



Dr.Eng. Zulkifli Tahir, S.T., M.Sc.
NIP. 198404032 01012 1 004

Dekan Fakultas Teknik
Universitas Hasanuddin



Prof. Dr.Eng. Ir. Muhammad Isran Ramli, M.T. IPM., ASEAN.Eng.
NIP. 19730926 200012 1 002

Ketua Program Studi
S2 Teknik Informatika



Dr. Ir. Zahir Zainuddin, M.Sc.
NIP. 19640427 198910 1 002

PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA

Yang bertanda tangan di bawah ini

Nama : Abd Salam At Taqwa
Nomor mahasiswa : D082211003
Program studi : S2 Teknik Informatika

Dengan ini menyatakan bahwa, tesis berjudul “Monokular 3D Rekonstruksi Wajah Menggunakan 3D Morphabel Model dan ElasticFace” adalah benar karya saya dengan arahan dari komisi pembimbing Dr. Ir. Zahir Zainuddin, M.Sc. dan Dr.Eng. Zulkifli Tahir, S.T., M.Sc. Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari isi tesis ini telah dipublikasikan di Prosiding 2023 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), halaman 320-325, dan DOI: 10.1109/IAICT59002.2023.10205588 sebagai artikel dengan judul “Monocular 3D Face Reconstruction Using 3D Morphable Model and ElasticFace”.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Gowa, 30 Oktober 2023

Yang menyatakan



Abd Salam At Taqwa

KATA PENGANTAR

Alhamdulillah rabbil'alamin, segala puji bagi Allah Subhanahu Wa Ta'ala Yang Maha Sempurna, yang telah memberikan rahmat, hidayah dan pertolongan-Nya sehingga penulis dapat menyelesaikan tesis dengan judul "Monokular 3D Rekonstruksi Wajah Menggunakan 3D Morphabel Model Dan Elasticface". Tak lupa pula shalawat dan salam kepada Nabi Muhammad Shallahu 'Alaihi Wasallam yang telah menyinari dunia ini dengan keindahan ilmu dan akhlak yang diajarkan kepada seluruh umatnya.

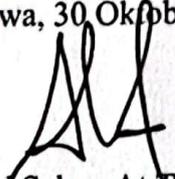
Tesis ini disusun untuk memenuhi persyaratan untuk memperoleh gelar Magister Komputer (M.Kom.) pada Program Pascasarjana Departemen Teknik Informatika Universitas Hasanuddin Makassar. Tentunya penyelesaian tesis ini tidak terlepas dari dukungan dan bantuan dari semua pihak. Untuk itu, dengan penuh kerendahan hati penulis menyampaikan terima kasih setulus-tulusnya dan setinggi-tingginya kepada:

1. Ayahanda penulis Basri Ahmad, S.Pd. dan ibunda tercinta Dra. Muliati, S.Pd. yang telah memberikan dukungan materil, doa dan motivasi yang kuat kepada penulis, hingga penulis dapat menyelesaikan penelitian ini.
2. Bapak Dr. Ir. Zahir Zainuddin, M.Sc. sebagai pembimbing pertama sekaligus sebagai Ketua Program Studi S2 Teknik Informatika dan Bapak Dr.Eng. Zulkifli Tahir, S.T., M.Sc. selaku dosen pembimbing kedua yang telah meluangkan waktunya kepada penulis untuk membimbing, memberikan masukan, memotivasi tiada henti-hentinya hingga tahap penyelesaian tesis ini.
3. Ibu Dr. Ir. Ingrid Nurtanio, M.T., Bapak Dr.Eng. Ady Wahyudi Paundu, S.T., M.T., dan Bapak Prof. Dr.Eng. Ir Syafaruddin., S.T., M.Eng., selaku penguji yang memberikan masukan dan saran yang membangun selama proses penelitian berlangsung.
4. Rekan-rekan Lab *Computer Based System* Teknik Informatika yang selalu saling mendukung dalam suka maupun duka dalam proses penyelesaian tesis ini.

5. Rekan–rekan Mahasiswa S2 Teknik Informatika Angkatan 2021 yang selalu mendukung dalam proses penyelesaian tesis ini.

Penulis menyadari bahwa tesis masih jauh dari kata sempurna dan di dalam penyelesaiannya masih menemui kesulitan dan hambatan, sehingga penulis tetap mengharapkan saran dan kritik untuk pengembangan lebih lanjut, agar dapat memberikan manfaat yang banyak bagi semua pembaca.

Gowa, 30 Oktober 2023



Abd Salam At Taqwa

DAFTAR ISI

HALAMAN SAMPUL	i
PENGAJUAN TESIS	ii
PERSETUJUAN TESIS	iii
PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xi
DAFTAR LAMPIRAN.....	xii
ABSTRAK	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Tujuan Penelitian.....	6
1.4 Manfaat Penelitian.....	6
1.5 Batasan Masalah.....	7
BAB II TINJAUAN PUSTAKA.....	9
2.1 Kajian Pustaka.....	9
2.1.1 3D Morphable Model.....	9
2.1.2 <i>Deep Face Recognition</i>	10
2.1.3 Deteksi Wajah.....	11
2.1.4 Deteksi <i>Landmark</i> Wajah	15
2.1.5 Kullback-Leibler divergence	20

2.1.6 Pytorch 3D	21
2.1.7 Protokol NoW	22
2.2 Metode Penyelesaian Masalah	27
2.2.1 <i>State of The Art</i>	27
2.2.2 Metode Penelitian	38
2.3 Target Hasil Penelitian	38
2.4 Kerangka Pikir.....	40
BAB III METODE PENELITIAN	41
3.1 Jenis Penelitian	41
3.2 Tahapan Penelitian	41
3.3 Sumber Data	42
3.4 Rancangan Sistem	42
3.4.1 Persiapan 3D Morphable Model.....	44
3.4.2 <i>Input Gambar</i>	45
3.4.3 Deteksi Wajah.....	45
3.4.4 Deteksi Landmark.....	45
3.4.5 “ <i>Early Stopping</i> ” Menggunakan Divergensi KL.....	46
3.4.1 <i>Rigid Fitting</i>	49
3.4.2 <i>Non Rigid Fitting</i>	50
3.5 Pengujian Sistem	55
BAB IV HASIL DAN PEMBAHASAN	57
4.1 Pembahasan	57
4.2 “ <i>Early Stopping</i> ” Menggunakan Divergensi KL	59
4.3 <i>Rigid Fitting</i>	60
4.4 <i>Non Rigid Fitting</i>	61
4.5 Hasil Visualisasi Rekonstruksi 3D Wajah.....	63

4.6 Evaluasi Kuantitatif Rekonstruksi 3D Wajah	63
4.7 Evaluasi Bentuk Muka Secara Keseluruhan.....	67
BAB V PENUTUP.....	70
5.1 Kesimpulan.....	70
5.2 Saran	71
DAFTAR PUSTAKA	72
LAMPIRAN.....	81

DAFTAR GAMBAR

Gambar 1 Proses pembuatan 3D Morphabel Model	9
Gambar 2 Gambar <i>input</i> diubah ukurannya ke skala yang berbeda untuk membangun piramida gambar.....	13
Gambar 3 P-Net (dari makalah MTCNN)	13
Gambar 4 R-Net (dari makalah MTCNN).....	14
Gambar 5 O-Net (dari makalah MTCNN)	15
Gambar 6 Database yang menyediakan titik landmark	17
Gambar 7 Point landmark pada wajah.....	19
Gambar 8 Modular <i>rendering pipeline</i> di PyTorch3D	22
Gambar 9 NoW Challenge	23
Gambar 10 <i>vertex</i> Pada 3D wajah	25
Gambar 11 Bentuk dan <i>vertex</i> Pada Hasil 3D Wajah.....	26
Gambar 12 Vertex Pada 3D <i>Ground Truth</i>	26
Gambar 13 Vertex Pada Hasil 3D Wajah dan 3D <i>Ground Truth</i>	27
Gambar 14 Tahapan Penelitian.....	41
Gambar 15 Sistem Penelitian	42
Gambar 16 Rancangan Penelitian.....	43
Gambar 17 Contoh Gambar NoW Dataset yang akan direkonstruksi.....	56
Gambar 18 Grafik <i>loss</i> untuk setiap iterasi (a) <i>landmark loss</i> , dan (b) <i>photometric loss</i>	61
Gambar 19 Grafik <i>cosine</i> similarity pada <i>perception loss</i>	62
Gambar 20 Hasil Rekonstruksi 3D Pada NoW dataset	63
Gambar 21 Gambar Monokular 2D Wajah yang Akan Direkonstruksi.	68
Gambar 22 Hasil Rekonstruksi 3D Wajah Dalam Bentuk 2D	69
Gambar 23 Perbandingan 3D Wajah dan Wajah dengan Sudut yang Berbeda... ..	69

DAFTAR TABEL

Tabel 1 Database wajah yang menyediakan 68 titik landmark	17
Tabel 2 <i>State of the art</i>	27
Tabel 3 Arsitektur <i>Deep Face Recognition</i>	52
Tabel 4 Hasil Evaluasi Setiap Identitas Pada NoW Challenge	65
Tabel 5 Hasil Evaluasi Pada NoW Challenge	67

DAFTAR LAMPIRAN

Lampiran 1 Data identitas 2D wajah yang direkonstruksi menjadi 3D wajah.... 82

ABSTRAK

Abd Salam At Taqwa. *Monokular 3D Rekonstruksi Wajah Menggunakan 3D Morphable Model Dan ElasticFace* (dibimbing oleh **Zahir Zainuddin** dan **Zulkifli Tahir**)

3D Morphable Model, salah satu model yang digunakan untuk merekonstruksi 3D wajah dari gambar monokular 2D wajah, telah mencapai hasil yang memuaskan seiring dengan perkembangan *computer vision* dan grafis. Namun, rekonstruksi 3D wajah menggunakan 3D Morphable Model secara *weakly-supervised* memiliki tantangan tersendiri karena tidak membutuhkan label sebagai *ground truth* dan hanya mengandalkan kemiripan fitur antara citra monokular 2D dan 3D wajah. Penelitian ini menggunakan rekonstruksi 3D wajah secara *weakly-supervised* dengan membandingkan ekstraksi fitur identitas. Dalam hal ini, teknik *deep face recognition* yang digunakan untuk ekstraksi fitur identitas adalah ArcFace, CosFace, dan ElasticFace. Proses rekonstruksi 3D wajah dibagi menjadi 1) *rigid fitting* untuk menyesuaikan landmark 3D wajah dengan landmark wajah dari citra monokular 2D dan 2) *non-rigid fitting feature similarity* dengan *hybrid-level weak supervision* yang mengaplikasikan model *deep face recognition* yang beragam. Hasil rekonstruksi 3D wajah kemudian dievaluasi menggunakan NoW Challenge. Hasil eksperimen pada protokol NoW menunjukkan bahwa memodifikasi model *deep face recognition* dapat mengubah hasil rekonstruksi 3D. ElasticFace-Arc+ dengan hasil error tertinggi dibandingkan dengan ElasticFace-Cos+ yang memiliki nilai error terendah, memiliki perbedaan penurunan error sebesar 0% untuk median dari 1,37 menjadi 1,37, perbedaan penurunan error untuk rata-rata 1,02 % dari 1,96 menjadi 1,94, perbedaan penurunan error standar deviasi sebesar 0,41 % dari 2,43 menjadi 2,42.

Kata Kunci: Rekonstruksi Wajah, *3D Morphable Model*, *Deep Face Recognition*

ABSTRACT

Abd Salam At Taqwa. *Monocular 3D Face Reconstruction Using 3D Morphable Model and ElasticFace* (supervised by **Zahir Zainuddin** dan **Zulkifli Tahir**)

3D Morphable Model, one of the models used to reconstruct 3D face from 2D monocular image of face, has achieved satisfactory results along with computer vision and graphics development. However, reconstructing 3D face using a 3D Morphable Model in a weakly-supervised manner has its challenges because it does not require labels as ground truth and only relies on the similarity of features between 2D monocular image and 3D face. This research uses weakly-supervised 3D face reconstruction by comparing identity feature extraction. In this case, deep face recognition techniques used for identity feature extraction are ArcFace, CosFace, and ElasticFace. The 3D face reconstruction process is divided into 1) rigid fitting to fit the 3D face landmarks into face landmarks of 2D monocular image and 2) non-rigid fitting feature similarity with hybrid-level weak supervision applying diverse deep face recognition models. The results of the reconstruction are subsequently evaluated using the NoW challenge. Experimental results on the NoW protocol show that modifying the deep face recognition model can change the 3D reconstruction results. ElasticFace-Arc+ with the highest error result compared to ElasticFace-Cos+ which has the lowest error value, has an error reduction difference of 0% for the median from 1,37 to 1,37, an error reduction difference for the average of 1,02% from 1,96 to 1,94, standard deviation error reduction difference of 0,41% from 2,43 to 2,42.

Keywords: 3D Face Reconstruction, 3D Morphable Model, Deep Face Recognition

BAB I

PENDAHULUAN

1.1 Latar Belakang

Wajah menjadi salah satu bagian tubuh yang penting untuk setiap individu manusia. Wajah menjadi identitas unik yang menjadi tolak ukur seseorang berbeda dengan yang lainnya. Wajah adalah sumber tanda nonverbal yang paling kaya dan paling bermakna bagi komunikasi sosial manusia. Pengenalan identitas orang lain melalui fitur wajah sangat penting dalam kehidupan. Oleh karenanya wajah menjadi pengukur identitas utama untuk seseorang dibandingkan anggota tubuh lainnya. Wajah manusia memberikan rangsangan visual yang dengannya orang dapat diidentifikasi dan dibedakan satu sama lain. Pada penelitian ini berfokus untuk merekonstruksi 3D wajah bukan seluruh tubuh. Hal ini dikarenakan tubuh manusia bersifat tidak unik meskipun ada yang gemuk dan kurus, namun fitur yang didapatkan terbatas karena biasanya tertutupi baju dan sebagainya.

3D Rekonstruksi wajah dari gambar monokular 2D menjadi topik penting dalam *computer vision* dan *graphic*. Rekonstruksi 3D wajah dapat menyelesaikan beberapa tantangan seperti pengenalan wajah, hiburan, aplikasi medis, forensik, ilmu kognitif, ilmu saraf, dan psikologi (Egger et al., 2020). 3D wajah dapat dijadikan pengenalan wajah karena wajah memiliki fitur-fitur unik yang lebih banyak dibandingkan anggota tubuh lainnya sebagai identitas manusia. Keterbatasan data atau sensor yang dapat ditangkap misalnya, hanya gambar yang ditangkap oleh kamera warna tunggal yang dapat diakses. Pada skenario yang dibatasi seperti itu, model wajah adalah alat yang ampuh untuk membatasi masalah rekonstruksi dengan lebih baik dibandingkan merekonstruksi secara keseluruhan bentuk tubuh.

3D wajah dapat digunakan dan dikonversi menjadi avatar 3D untuk game dan VR atau bahkan karakter yang dirancang oleh seniman dapat dikontrol dalam skenario permainan berdasarkan pelacak padat yang menggunakan model wajah parametrik atau 3D wajah. Kontrol berbasis visi tersebut pertama kali

didemonstrasikan dalam pengaturan off-line (Chai et al., 2003). Pelacakan wajah juga dapat digunakan untuk memungkinkan komunikasi tatap muka (H. Li et al., 2015; Lombardi et al., 2018; Thies et al., 2016) di dalam *virtual reality*. Rekonstruksi 3D wajah berdasarkan model wajah parametrik juga dapat digunakan untuk membangun cermin virtual yang memungkinkan percobaan aksesoris atau make-up. Penggantian wajah juga memungkinkan ketika mendapatkan 3D wajah kemudian penggantian wilayah wajah dalam video target dengan yang ada di video sumber. Untuk tujuan ini, kedua orang direkonstruksi berdasarkan model parametrik yang sama sehingga menghasilkan korespondensi antar-orang. Peragaan Wajah dan Sulih Suara Visual juga dapat dilakukan dengan 3D wajah. Peragaan wajah adalah proses mentransfer ekspresi wajah dari sumber ke video target. Peragaan wajah juga dapat diterapkan pada masalah suara visual, yaitu tugas mengadaptasi gerakan mulut aktor target agar sesuai dengan trek audio baru.

Aplikasi klinis 3DMM mencakup analisis dan sintesis. Aplikasi yang dominan terletak pada analisis, di mana penyakit dapat dikenali dari bentuk wajah. Salah satu contoh dari efek tersebut adalah klasifikasi dan diagnosis dini gangguan spektrum alkohol janin (Suttie et al., 2013) atau epilepsi (Ahmedt Aristizabal, 2019). Demikian pula, (Hammond et al., 2004) menunjukkan visualisasi dan pengenalan gangguan pertumbuhan kraniofasial bawaan. Kedua karya ini menggunakan data 3D. Namun, kemampuan rekonstruksi 3D dari gambar 2D telah dieksplorasi untuk skrining akromegali (Erik Learned-Miller et al., 2006) dan kelainan genetik (Tu et al., 2018). Dalam arah sintesis, model bentuk 3D dieksplorasi untuk melakukan rekonstruksi bagian wajah yang hilang berdasarkan statistik model (Basso et al., 2005; Mueller et al., 2011). Rekonstruksi seperti ini dapat diterapkan untuk desain implan yang dipersonalisasi.

Rekonstruksi 3D wajah secara otomatis sangat mengurangi operasi manual yang memakan waktu dan berulang dalam membuat urutan bentuk 3D untuk setiap karakter baru. Satu Langkah penting dalam model ini adalah memperoleh model 3D wajah yang dipersonalisasi yang idealnya dibuat menggunakan kamera 3D. Namun biaya yang tinggi dan jangkauan yang terbatas

sehingga membatasi penerapannya dalam praktik. Pendekatan alternatif yang dapat dilakukan adalah merekonstruksi 3D wajah menggunakan gambar 2D. Rekonstruksi 3D wajah dari gambar 2D merupakan tantangan karena kemungkinan hilangnya informasi dalam proses akuisisi citra.

3D Morphable Model (Blanz & Vetter, 1999) memberikan "ruang wajah" dimensi yang merepresentasikan wajah umum untuk analisis gambar. Namun, ruang berdimensi rendah membatasi kapasitas pemodelannya seperti yang ditunjukkan pada (Tewari et al., 2018; Tran & Liu, 2018). Metode rekonstruksi 3D wajah tradisional menggunakan *3D morphable model* didasarkan pada algoritma pengoptimalan untuk mendapatkan koefisien *3D morphable model* dan merender 3D wajah yang sesuai dari gambar wajah dengan mengoptimalkan bentuk, tekstur, pose, dan pencahayaan. Namun, metode tersebut biasanya memakan waktu karena kombinasi linear atau nonlinear yang kompleks yang mungkin sulit dipecahkan dalam praktiknya. Penelitian terdahulu telah memanfaatkan *machine learning* dengan cara regresi dari piksel gambar ke koordinat model yang dapat diubah menggunakan *neural network* (Dou et al., 2017; Genova et al., 2018) Masalah utama dengan pendekatan regresi adalah kurangnya data 3D wajah untuk pelatihan (Genova et al., 2018). Pemindaian geometri dan tekstur wajah sulit diperoleh, baik karena pertimbangan biaya maupun privasi.

Beberapa pendekatan lain menggunakan *structure from motion* (SfM) (Lauze et al., 2014; Zhao et al., 2022) untuk memperkirakan bentuk 3D wajah dari kumpulan gambar wajah. Dibandingkan dengan *photometric stereo* yang tidak dikalibrasi, SfM dapat digunakan untuk pencahayaan dan pose yang bervariasi, namun, sulit bagi SfM untuk merekonstruksi bentuk 3D wajah yang padat karena jumlah titik fitur yang terdeteksi pada citra wajah sangat terbatas. Selain itu, SfM tidak dapat menangani transformasi objek yang tidak kaku, seperti variasi ekspresi wajah, yang merupakan komponen penting dari 3D wajah (Dou & Kakadiaris, 2018). Hal lainnya adalah SfM juga membutuhkan komputasi yang banyak karena bergantung pada gambar wajah yang besar. Penelitian (Dou & Kakadiaris, 2018) mengatasi masalah tersebut dengan cara menggunakan *recurrent neural network* untuk memadukan fitur terkait identitas dari *convolutional neural network* untuk

menghasilkan rekonstruksi yang lebih diskriminatif, tetapi batasan geometris *multi-view* tidak dieksploitasi dalam pendekatan mereka.

Saat ini perkembangan 3DMM mengalami kemajuan pesat seiring dengan pertumbuhan *Convolutional Neural Network* (CNN). Metode berbasis CNN (Dou et al., 2017; Genova et al., 2018) telah mencapai kesuksesan luar biasa dalam rekonstruksi 3D wajah dan keselarasan wajah. Berbagai metode berbasis CNN diusulkan untuk meregresi parameter model wajah atau bentuk wajah dari data pelatihan. Metode ini berfokus pada rekonstruksi statis yang akibatnya membutuhkan data pelatihan 3D wajah yang banyak dan gambar RGB yang sesuai yang jumlahnya terbatas dan sulit untuk diperoleh. Akibatnya beberapa penelitian menggunakan data sintetis untuk penyesuaian 3DMM (Dou et al., 2017; Y. Feng et al., 2018; Richardson et al., 2016; Tuan Tran et al., 2017), atau *self supervised learning* (Chinaev et al., 2018; Genova et al., 2018; Richardson et al., 2017). Resiko dari menggunakan data sintesis adalah akurasinya akan rendah jika label pelatihan tidak sempurna dan tidak cocok digunakan untuk gambar wajah 2D yang diambil secara nyata. Seperti pada penelitian (Richardson et al., 2017) yang menunjukkan hasil rekonstruksi 3D wajah kurang memuaskan karena gagal menggeneralisasi fitur wajah yang tersedia di data pelatihan atau ketergantungan terhadap data sintesis yang dilatih.

Untuk mengatasi masalah tersebut beberapa penelitian mengusulkan pelatihan *network* tanpa menggunakan label dengan cara metode *unsupervised* atau *weakly-supervised* dan hasil yang diperoleh memuaskan (Y. Deng et al., 2019; Genova et al., 2018; Sengupta et al., 2018; Shang et al., 2020; Tewari et al., 2017, 2018; Tran & Liu, 2018). Inti dari metode *unsupervised* adalah bagaimana menghasilkan 3D wajah (Genova et al., 2018) yang di render berdasarkan koefisien dari *3D Morphable Model* kemudian mengukur tingkat kesamaan dari 3D wajah dan gambar *2D wajah*. Misalnya pada penelitian (Sengupta et al., 2018; Tewari et al., 2017) menggunakan perbedaan fotometrik pixel antara gambar 2D wajah dan 3D wajah yang dihasilkan. Penelitian (Tewari et al., 2018) menggunakan dua tahap skema rekonstruksi dimana tahap kedua yaitu untuk memprediksi bentuk dan tekstur menggunakan *neural network*. Penelitian (Y. Deng et al., 2019; Genova et al., 2018; Shang et al., 2020; Zhang et al., 2021)

menggunakan jarak antara fitur-fitur yang diekstraksi dari jaringan pengenalan wajah untuk membedakan gambar 2D wajah dan 3D wajah yang telah dirender. Penelitian (Y. Deng et al., 2019; Shang et al., 2020; Zhang et al., 2021) menggunakan fotometrik pixel dan jarak antara fitur-fitur yang diekstraksi dari *deep face recognition* untuk menghasilkan hasil yang lebih akurat.

Rekonstruksi 3D wajah menggunakan *deep face recognition* seperti pada penelitian (Y. Deng et al., 2019; Genova et al., 2018; Shang et al., 2020; Zhang et al., 2021) mendapatkan hasil yang memuaskan. Proses *deep face recognition* pada penelitian tersebut menggunakan Google's FaceNet (Schroff & Philbin, 2015). Pengenalan wajah dilakukan untuk membandingkan fitur-fitur yang didapatkan dari kedua gambar kemudian diukur tingkat kesamaannya indentitas antara kedua wajah. Pengenalan wajah Google's Facenet (Schroff & Philbin, 2015) berbasis klasifikasi menggunakan *softmax loss function* dapat dengan mudah diadopsi untuk melatih model pengenalan wajah karena tidak menimbulkan masalah tersebut. Namun, *softmax loss function* tidak secara langsung mengoptimalkan penyematan fitur yang diperlukan untuk verifikasi wajah. Oleh karenanya beberapa penelitian (Boutros et al., 2022; J. Deng et al., 2019; Liu et al., 2017; Wang et al., 2018) mengembangkan loss function berbasis margin sudut yang dapat mengoptimalkan variasi intra-kelas yang kecil dan variasi antar-kelas yang besar. ElasticFace (Boutros et al., 2022) sebagai *state-of-the-art* proses pengenalan wajah yang memiliki margin sudut yang elastis antar kelas yang memungkinkan fleksibilitas dalam mendorong keterpisahan kelas. Ini bertujuan untuk memberikan peluang *decision boundary* untuk mengekstraksi dan menarik kembali untuk memungkinkan ruang bagi pembelajaran keterpisahan kelas yang fleksibel.

Proses rekonstruksi 3D wajah dari gambar 2D menggunakan 3D *Morphabel Model* melalui *unsupervised* atau *weakly-supervised* memiliki tantangan sendiri karena tidak memerlukan label sebagai *ground truth* dan hanya mengandalkan kesamaan fitur antara gambar 2D wajah dan 3D wajah yang dirender. Sehingga 3D wajah yang dihasilkan bergantung dengan seberapa bagus metode ekstraksi fitur yang digunakan. Kesalahan dalam memilih metode untuk ekstraksi fitur antara gambar 2D wajah dan 3D wajah yang dihasilkan dapat

mengakibatkan 3D wajah yang direkonstruksi mungkin kekurangan detail wajah yang penting, berisi fitur wajah yang salah, dan tidak selaras dengan gambar. Oleh karena itu diperlukan metode ekstraksi fitur yang akurat sebagai hasil dari rekonstruksi 3D wajah yang maksimal.

Berdasarkan permasalahan diatas, maka penelitian ini mengusulkan penelitian terkait metode rekonstruksi 3D wajah dari gambar 2D wajah untuk memperoleh hasil maksimal untuk memprediksi bentuk 3D wajah yang dihasilkan menggunakan gambar monocular 2D wajah. Penelitian ini berfokus untuk membandingkan metode ekstraksi fitur berdasarkan *deep face recognition*. Fitur yang didapatkan diharapkan mampu mendapatkan koefisien render *3D morphable model* yang disesuaikan dengan gambar 2D wajah secara optimal.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah diatas, maka rumusan masalah pada penelitian ini adalah:

1. Bagaimana membandingkan ekstraksi fitur identitas berdasarkan model *deep face recognition* secara *weakly-supervised* untuk rekonstruksi 3D wajah
2. Bagaimana mengevaluasi 3D wajah yang dihasilkan menggunakan protokol *NoW Challenge*?

1.3 Tujuan Penelitian

Tujuan yang ingin dicapai pada penelitian ini adalah sebagai berikut :

1. Merekonstruksi 3D wajah dari gambar monocular 2D wajah dengan menggunakan ekstraksi fitur identitas menggunakan *deep face recognition* yang terbaik.
2. Mengevaluasi 3D wajah yang dihasilkan untuk mendapatkan hasil prediksi bentuk wajah terbaik.

1.4 Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut :

1. Memberi kontribusi berupa model *deep face recognition* terbaik untuk merekonstruksi 3D wajah dari gambar monokular 2D wajah untuk mendapatkan koefisien yang sesuai pada model *3D Morphable Model*.
2. Pengembangan ilmu pengetahuan dan teknologi bagi para peneliti bidang serupa dan akademi, terutama untuk penulis.

1.5 Batasan Masalah

Adapun batasan masalah pada penelitian ini adalah :

1. Penelitian ini hanya difokuskan untuk rekonstruksi wajah berbasis *3D Morphable Model*
2. Data yang direkonstruksi berupa wajah yang diambil menggunakan satu kamera (monokular).

BAB II

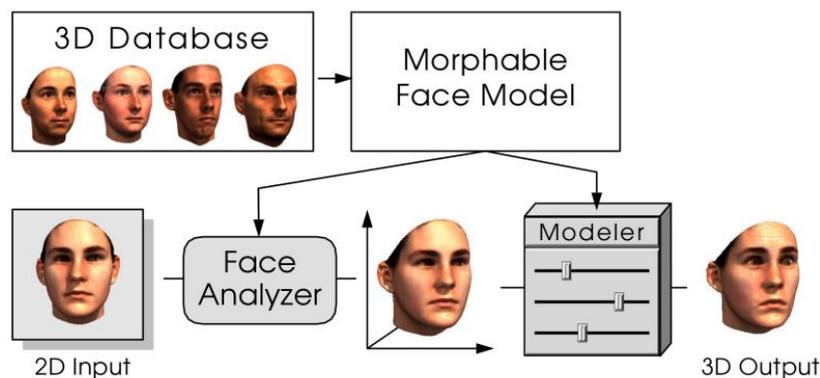
TINJAUAN PUSTAKA

2.1 Kajian Pustaka

2.1.1 3D Morphable Model

3D Morphable Model (3DMM) adalah model generatif untuk bentuk dan penampilan wajah yang didasarkan pada dua gagasan utama: Pertama, semua dataset 3D wajah berada dalam korespondensi point-to-point yang padat yang biasanya ditetapkan pada sekumpulan contoh wajah dalam prosedur pendaftaran. Karena korespondensi ini, kombinasi linier dari wajah dapat didefinisikan dan menghasilkan wajah yang realistis secara morfologis (morf). Ide kedua adalah memisahkan bentuk dan warna wajah dan memisahkannya dari faktor eksternal seperti pencahayaan dan parameter kamera. *Morphable Model* mungkin melibatkan model statistik distribusi wajah, yang merupakan analisis komponen utama dalam karya asli Blanz dan Vetter (Blanz & Vetter, 1999) dan telah memasukkan teknik pembelajaran lainnya dalam penelitian selanjutnya.

3D Morphable Model adalah model objek statistik yang memisahkan bentuk dari variasi tampilan. Biasanya, Model dipelajari dari pemindaian 3D berkualitas tinggi dari beberapa *instance* objek. Ini mengurangi dimensi dan memberikan representasi objek parametrik berdimensi rendah. Model yang dihasilkan adalah generatif, yang berarti bahwa dari satu set parameter sampel acak muncul contoh objek realistis baru. Model seperti itu kemudian dapat digunakan dan menghasilkan 3D wajah berdasarkan koefisien dari 3DMM.



Gambar 1 Proses pembuatan 3D Morphabel Model

Gambar 1 menunjukkan abstrak visual karya Blanz dan Vetter [1999], mengusulkan model statistik untuk wajah untuk melakukan rekonstruksi 3D wajah dari gambar 2D wajah dan ruang wajah parametrik yang memungkinkan manipulasi terkontrol (Egger et al., 2020). Variasi dari 3DMM telah tersedia dari beberapa penelitian (Booth et al., 2018; Cao et al., 2013; Gerig et al., 2017; Huber et al., 2016). Model ini menggunakan representasi dimensi rendah untuk ekspresi wajah, tekstur, dan identitas. Basel Face Model (BFM) adalah salah satu model 3DMM yang tersedia untuk umum. Model dibangun dengan mendaftarkan template mesh yang sesuai dengan wajah yang dipindai dan diperoleh dari Iterative Closest Point (ICP) dan Principal Component Analysis (PCA) (Lin et al., 2020). Menggunakan model *morphable*, wajah manusia 3D dapat diparameterisasi sebagai vektor bobot untuk basis bentuk yang akan di render.

2.1.2 *Deep Face Recognition*

Pengenalan wajah secara mendalam atau *deep face* recognition adalah perangkat lunak yang mengidentifikasi atau mengonfirmasi identitas seseorang menggunakan wajah mereka. Pengenalan wajah berfungsi untuk mengidentifikasi dan mengukur fitur wajah dalam sebuah citra. Pengenalan wajah dapat mengidentifikasi wajah manusia dalam citra atau video, menentukan apakah wajah dalam dua citra milik orang yang sama, atau mencari wajah di antara banyak kumpulan citra yang ada. Sistem keamanan biometrik menggunakan pengenalan wajah untuk mengidentifikasi individu secara unik selama orientasi pengguna atau masuk serta memperkuat aktivitas autentikasi pengguna.

Perkembangan penelitian dalam *deep face recognition* membuat ekstraksi identitas fitur mengenali wajah dengan identitas yang sama. Google's *FaceNet* (Schroff & Philbin, 2015) adalah pengenalan wajah berbasis klasifikasi yang menggunakan fungsi *softmax loss* untuk mengekstrak fitur dan meningkatkan akurasi pengenalan wajah dalam dengan menargetkan peningkatan kekompakan intra-kelas dan perbedaan antar-kelas. Namun, fungsi *softmax loss* tidak secara langsung mengoptimalkan fitur penyematan yang diperlukan untuk verifikasi wajah. Oleh karena itu, beberapa penelitian (Boutros et al., 2022; J. Deng et al., 2019; Wang et al., 2018) mengembangkan fungsi *loss* berbasis margin sudut yang

mengoptimalkan variasi intra-kelas yang kecil dan variasi antar-kelas yang besar. CosFace (Wang et al., 2018) mengusulkan margin kosinus aditif pada sudut kosinus antara fitur-fitur dengan identitas yang sama dan bobot yang sesuai, sedangkan ArcFace (J. Deng et al., 2019) mengusulkan margin sudut aditif dengan menerapkan margin penalti sudut pada sudut antara fitur-fitur dengan identitas yang sama dan bobot yang sesuai. Tidak seperti penelitian sebelumnya (J. Deng et al., 2019; Wang et al., 2018), yang menerapkan nilai margin tetap, ElasticFace (Boutros et al., 2022) mengakomodasi fleksibilitas dengan melonggarkan margin dengan mengambil nilai margin secara acak dari distribusi Gaussian. Karena ketiga konsep deep face recognition ini hampir memiliki konsep yang sama, maka pada penelitian ini, model yang digunakan sebagai pembanding dalam metode rekonstruksi 3D wajah adalah ArcFace (J. Deng et al., 2019), CosFace (Wang et al., 2018) dan ElasticFace (Boutros et al., 2022).

2.1.3 Deteksi Wajah

Face detection alias deteksi wajah merupakan suatu teknologi komputer berbasis Artificial Intelligence (AI) yang digunakan untuk mengekstraksi dan mengidentifikasi wajah seseorang dari gambar digital. Ketika terintegrasi dengan sistem keamanan biometrik (khususnya, pengenalan wajah), maka teknologi semacam inilah yang memungkinkan untuk memantau dan melacak orang secara real-time. Face detection merupakan langkah awal dalam analisis wajah, pelacakan wajah, dan yang terpenting, pengenalan wajah. Teknologi ini diterapkan pada pembukaan kunci perangkat, perbankan, perhotelan, penegakan hukum, keamanan gedung, dan masih banyak lagi lainnya.

Teknologi *face detection* menggunakan pembelajaran mesin dan algoritma untuk mengekstraksi wajah manusia dari gambar yang lebih besar. Gambar seperti itu biasanya akan berisi banyak objek non-wajah, seperti bangunan, lanskap, dan berbagai bagian tubuh lainnya. Algoritma deteksi wajah biasanya akan dimulai dengan mencari mata manusia, yang merupakan salah satu fitur wajah yang paling mudah dideteksi. Kemudian selanjutnya, algoritma mungkin mencoba menemukan mulut, hidung, alis, hingga iris mata. Setelah mengidentifikasi fitur wajah ini, dan algoritma menyimpulkan bahwa ia telah mengekstraksi wajah,

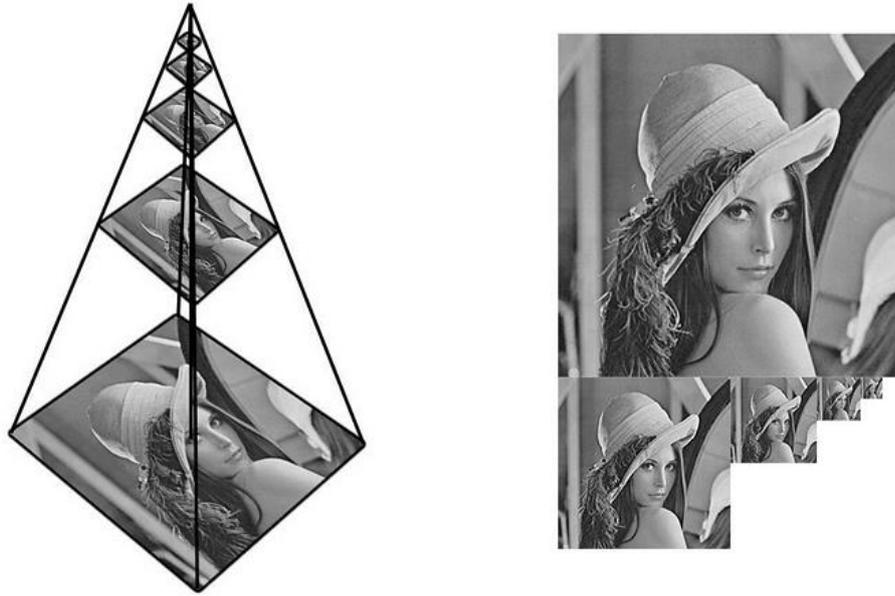
maka ia kemudian menjalani tes tambahan untuk mengonfirmasi bahwa itu memang wajah.

Sistem deteksi wajah atau *face detection* memang menjadi semakin umum di lingkungan online. Misalnya untuk memverifikasi identitas user yang masuk ke situs website dan aplikasi, untuk mengautentikasi user saat mereka melakukan pembelian, hingga untuk mengamankan informasi pribadi seperti kata sandi, dan lainnya. Sistem *face detection* kian berkembang, dan menjadi semakin akurat dari waktu ke waktu. Deteksi dan penyesuaian wajah di lingkungan yang bebas menjadi tantangan karena berbagai pose, iluminasi, dan oklusi. Studi terbaru menunjukkan bahwa pendekatan *deep learning* dapat mencapai kinerja yang mengesankan. Salah satu model pengenalan wajah yang populer digunakan adalah *Multi-task Cascaded Convolutional Networks* (MTCNN). Selain dapat mendeteksi wajah MTCNN juga dapat digunakan untuk Penyesuaian Landmark Wajah.

Multi-task Cascaded Convolutional Networks (MTCNN) adalah sebuah kerangka kerja yang dikembangkan sebagai solusi untuk deteksi wajah dan penyesuaian wajah. Proses ini terdiri dari tiga tahap jaringan konvolusional yang dapat mengenali wajah dan lokasi tengara seperti mata, hidung, dan mulut. Makalah ini mengusulkan MTCNN sebagai cara untuk mengintegrasikan kedua tugas (pengenalan dan penyesuaian) menggunakan pembelajaran *multi-task*. Pada tahap pertama, MTCNN menggunakan *shallow CNN* untuk menghasilkan kandidat *window* dengan cepat. Pada tahap kedua, ia menyempurnakan kandidat *window* yang diusulkan melalui CNN yang lebih kompleks. Dan terakhir, pada tahap ketiga menggunakan CNN ketiga, yang lebih kompleks daripada yang lain, untuk lebih menyempurnakan hasil dan output posisi landmark wajah.

2.1.3.1 3 Tahap MTCNN

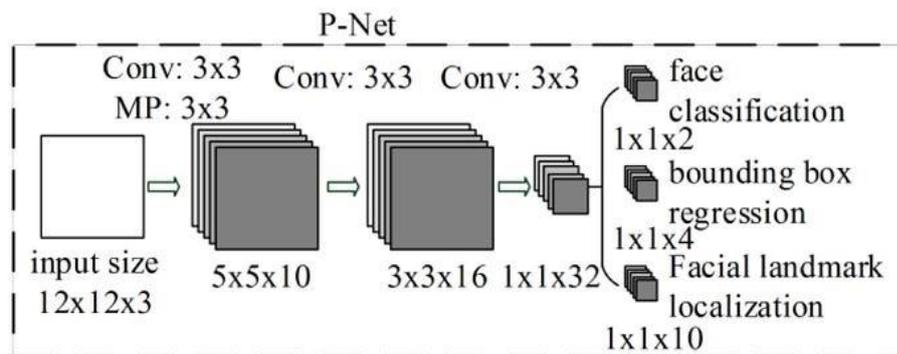
Langkah pertama adalah mengambil gambar dan mengubah ukurannya ke skala yang berbeda untuk membangun piramida gambar, yang merupakan masukan dari jaringan three-staged cascaded berikut ini.



Gambar 2 Gambar *input* diubah ukurannya ke skala yang berbeda untuk membangun piramida gambar

Gambar 2 menunjukkan gambar *input* yang diberikan akan diubah ukurannya menjadi skala yang berbeda sebelum dimasukkan kedalam tahap 1 *Proposal Network*. Hal ini bertujuan untuk membangun piramida gambar untuk mempermudah pendeteksian wajah.

a. Tahap 1: Proposal Network (P-Net)



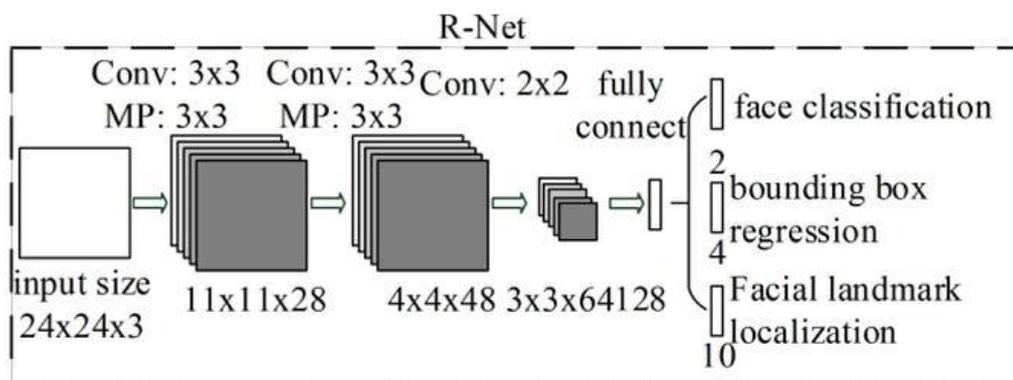
Gambar 3 P-Net (dari makalah MTCNN)

Gambar 3 menunjukkan tahap pertama dari metode pendeteksian wajah. Tahap pertama ini adalah *fully convolutional network* (FCN). Perbedaan antara CNN dan FCN adalah bahwa *fully convolutional network* tidak menggunakan lapisan padat sebagai bagian dari arsitekturnya. Jaringan

Proposal ini digunakan untuk mendapatkan kandidat *window* dan vektor regresi bounding box.

Regresi bounding box adalah teknik yang populer untuk memprediksi lokalisasi kotak ketika tujuannya adalah mendeteksi objek dari beberapa kelas yang telah ditentukan sebelumnya, dalam hal ini wajah. Setelah mendapatkan vektor bounding box, beberapa penyempurnaan dilakukan untuk menggabungkan daerah yang tumpang tindih. Hasil akhir dari tahap ini adalah semua kandidat *window* setelah dilakukan *refinement* untuk mengecilkan volume kandidat.

b. Tahap 2: *Refine Network* (R-Net)

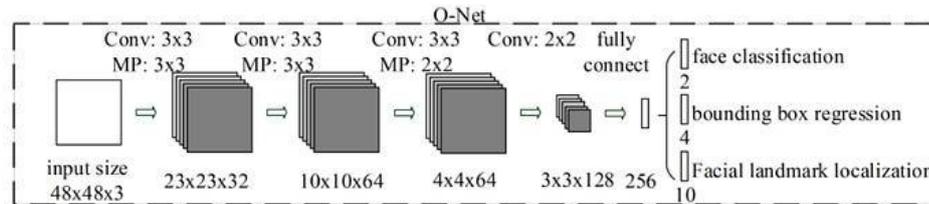


Gambar 4 R-Net (dari makalah MTCNN)

Gambar 4 menunjukkan tahap II metode deteksi wajah, dimana semua kandidat dari P-Net dimasukkan ke dalam *Refine Network*. Jaringan tersebut adalah CNN, bukan FCN seperti yang sebelumnya karena ada lapisan padat pada tahap terakhir dari arsitektur jaringan. R-Net selanjutnya mengurangi jumlah kandidat, melakukan kalibrasi dengan bounding box regression dan menggunakan *non-maximum suppression* (NMS) untuk menggabungkan kandidat yang tumpang tindih.

R-Net mengeluarkan output apakah input adalah wajah atau bukan, vektor 4 elemen yang merupakan kotak pembatas untuk wajah, dan vektor 10 elemen untuk lokalisasi *landmark* wajah.

c. Tahap 3: The Output Network (O-Net)



Gambar 5 O-Net (dari makalah MTCNN)

Gambar 5 menunjukkan tahap III, dimana pada tahap ini mirip dengan R-Net, tetapi Output Network ini bertujuan untuk menggambarkan wajah secara lebih rinci dan menghasilkan lima posisi landmark wajah untuk mata, hidung, dan mulut.

2.1.3.2 3 Tugas MTCNN

Berikut beberapa tugas yang bisa diselesaikan menggunakan metode MTCNN.

- Klasifikasi wajah: ini adalah masalah klasifikasi biner yang menggunakan cross-entropy loss.
- Regresi bounding box: tujuan pembelajaran adalah masalah regresi. Untuk setiap kandidat *window*, offset antara kandidat dan ground truth terdekat dihitung. *Euclidean loss* digunakan untuk tugas ini.
- Lokalisasi Landmark Wajah: lokalisasi landmark wajah diformulasikan sebagai masalah regresi, di mana fungsi kerugiannya adalah jarak *Euclidean*. Ada lima landmark: mata kiri, mata kanan, hidung, sudut mulut kiri, dan sudut mulut kanan.

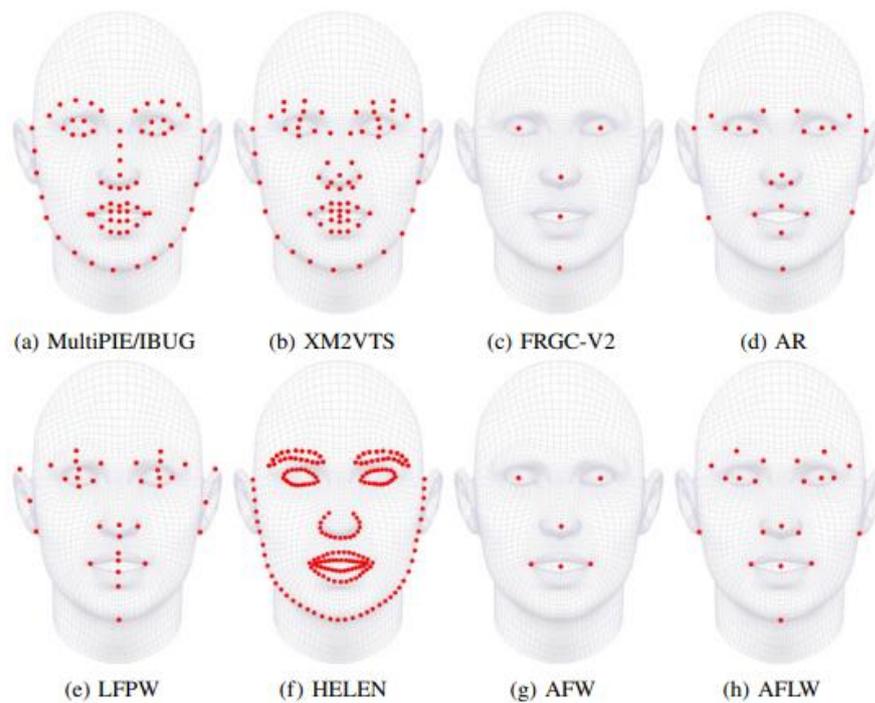
2.1.4 Deteksi *Landmark* Wajah

Peneliti komputer meneliti bagaimana mereka mengidentifikasi wajah manusia dalam sebuah gambar. Untuk itu, pertama-tama harus mengidentifikasi di mana letak wajah manusia dalam keseluruhan gambar. Detektor wajah adalah metode yang menempatkan wajah manusia dalam gambar dan mengembalikannya sebagai nilai kotak pembatas atau kotak persegi panjang. Mengidentifikasi wajah dalam foto atau video memang sangat keren, tetapi informasi ini tidak cukup untuk membuat aplikasi yang kuat, diperlukan lebih banyak informasi tentang

wajah orang tersebut, seperti posisi, apakah mulut terbuka atau tertutup, apakah mata terbuka, tertutup, melihat ke atas dan lain-lain. Fungsi dari deteksi *landmark* adalah mendeteksi wajah dalam gambar tertentu, dan kemudian akan mengembalikan titik *landmark* wajah untuk wajah yang pertama kali ditemukan dalam gambar.

Landmark wajah adalah *key points* pada bentuk wajah yang terdeteksi dan dapat digunakan sebagai fitur wajah untuk melakukan beberapa tugas seperti meningkatkan pengenalan wajah, menyelaraskan gambar wajah, membedakan pria dan wanita, memperkirakan pose kepala, dan sebagainya. Titik-titik penting dari *landmark* jarang digunakan sebagai representasi tugas verifikasi wajah, biasanya menggunakan *facial nodal points* sebagai pengganti. *Facial nodal points* dapat digunakan secara langsung dari *facial landmarks* atau dapat menghitung beberapa titik baru yang dimulai dari *facial landmarks*. Misalnya, mata, hidung, dan mulut merupakan bagian yang sangat mewakili wajah seseorang, sehingga titik-titik yang berhubungan dengan bagian-bagian wajah tersebut dapat relevan untuk mewakili wajah tersebut. Khususnya, misalnya untuk mata, kita dapat menggunakan pusat massa mata dibandingkan menggunakan *facial landmarks* yang membentuk kontur mata (Amato et al., 2018).

Jumlah *key points* pada *landmark* standar dan tetap sesuai dengan algoritma yang diterapkan. Penggunaan *landmark* berfungsi untuk melokalisasi daerah yang menonjol pada wajah untuk penyelarasan wajah, perubahan wajah, penggantian wajah, pengenalan wajah, dll. Hubungan antara rekonstruksi 3D wajah dan penyelarasan wajah pada dasarnya terdiri dari pemetaan dan estimasi geometri 3D wajah dari satu gambar 2D. Tujuan utamanya adalah menghitung visibilitas dan posisi *landmark* 2D (Jabberi et al., 2023). Jumlah keypoints pada *landmark* bergantung pada database yang disediakan, intinya adalah keypoints dapat mewakili wilayah penting pada wajah seperti mata, hidung, mulut, dan garis rahang. Berikut database-database yang tersedia untuk deteksi *landmark*.



Gambar 6 Database yang menyediakan titik landmark

Gambar 6 menunjukkan beragam jumlah dan letak titik *landmark* pada database yang disediakan. Diantara jumlah titik landmark yang dapat merepresentasikan wilayah penting pada wajah, jumlah titik 68 landmark yang disediakan oleh Multi-PIE (Gross et al., 2010) lebih populer dan sering digunakan dibandingkan pada titik-titik landmark yang lainnya. Titik-titik ini ditriangulasi dan gambar dapat lebih fleksibel dengan lengkungan ke dalam koordinat di mana titik-titik kanonik berada di lokasi tetap.

Tabel 1 Database wajah yang menyediakan 68 titik landmark

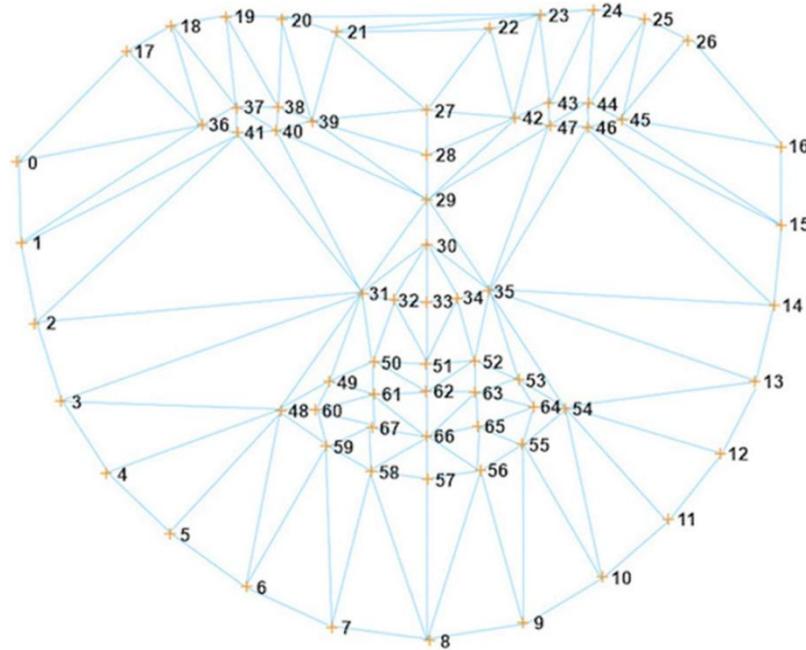
Nama	Author	Sumber	Published	Marks	Sampel
300-W	<u>Imperial College London</u>	https://ibug.doc.ic.ac.uk/resources/300-W/	2013	68	600
300-VW	<u>Imperial College London</u>	https://ibug.doc.ic.ac.uk/resources/300-W/	2015	68	218597

Tabel 1 (lanjutan)

Nama	Author	Sumber	Published	Marks	Sampel
300-W	<u>Imperial College London</u>	https://ibug.doc.ic.ac.uk/resources/300-W/	2013	68	600
300-VW	<u>Imperial College London</u>	https://ibug.doc.ic.ac.uk/resources/300-W/	2015	68	218597
AFW	<u>Imperial College London</u>	https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/	2013	68	337
AFLW2000-3D	<u>Chinese Academy of Sciences</u>	http://www.cbsr.ia.ac.cn/users/xiangyuzhu/projects/3DDFA/main.htm	2015	68	2000
HELEN	<u>Imperial College London</u>	https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/	2013	68	2330
IBUG	<u>Imperial College London</u>	https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/	2013	68	135
LFPW	<u>Imperial College London</u>	https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/	2013	68	1035

Pada **Tabel 1** dapat dilihat daftar database yang menyediakan 68 titik landmark juga lebih banyak dibandingkan dengan jumlah titik-titik landmark yang lainnya. berikut table database yang menyediakan 68 titik-titik landmark pada wajah. Pada penelitian (Jabberi et al., 2023) telah melakukan eksperimen dan tes apakah 68 *landmark* wajah dapat melakukan penyalarsan terhadap 3D wajah dengan baik yang membuktikan ketangguhannya terhadap pose wajah. Hasil yang didapatkan bahwa teknik yang menggunakan 68 *landmark* pada penelitian (Bulat & Tzimiropoulos, 2017) dapat digunakan di hampir semua pose wajah untuk

penyelarasan 3D wajah. Teknik yang digunakan memberikan hasil terbaik untuk anotasi wajah dibandingkan dengan algoritma Chow-Liu (Chow & Liu, 1968), yang banyak digunakan dalam metode pendeteksian *landmark* wajah baru-baru ini meskipun merupakan teknik lama, dan dibandingkan dengan metode Gaussian-Newton (Tzimiropoulos & Pantic, 2014), yang juga banyak digunakan dalam penyelarasan wajah.



Gambar 7 Point landmark pada wajah

Gambar 7 menunjukkan 68 titik landmark yang paling banyak digunakan untuk pendeteksian titik-titik *landmark* pada wajah. Deteksi landmark dilakukan dalam dua langkah:

- Deteksi wajah menemukan wajah manusia dari gambar dan mengembalikan nilai dalam bentuk koordinat persegi panjang pembatas dari wajah yang terdeteksi.
- Deteksi landmark, Setelah mendeteksi wajah, Selanjutnya menemukan titik-titik untuk poin-poin penting.

Bergantung pada penampilan wajah dan pola bentuk wajah, algoritma pendeteksian tengara wajah diklasifikasikan menjadi tiga. Tampilan wajah mengacu pada pola intensitas piksel yang khas di sekitar *landmark* wajah atau di

seluruh wilayah wajah, sedangkan pola bentuk wajah mengacu pada pola bentuk wajah yang ditentukan oleh lokasi tengara dan hubungan spasialnya.

Tiga kategori utama algoritma pendeteksian tengara adalah metode holistic, metode Constrained Local Model (CLM), Metode berbasis regresi.

- Metode holistik secara eksplisit memodelkan tampilan wajah secara menyeluruh dan pola bentuk wajah global.
- CLM bergantung pada tampilan wajah lokal yang eksplisit dan pola bentuk wajah global yang eksplisit.
- Metode berbasis regresi menggunakan informasi tampilan holistik atau lokal dan mereka dapat menanamkan pola bentuk wajah global secara implisit untuk deteksi tengara bersama.

Beberapa metode terbaru menggabungkan model *deep learning* dan model bentuk 3D global untuk deteksi *landmark*. Tantangan dalam pendeteksian wajah dan lokalisasi *landmark* adalah ekspresi wajah dan pose kepala yang berbeda, berbagai kondisi lingkungan seperti pencahayaan, oklusi wajah oleh objek lain atau oklusi diri sendiri oleh pose kepala yang ekstrem

2.1.5 Kullback-Leibler divergence

Dalam statistik matematika, divergensi Kullback–Leibler (juga disebut entropi relatif dan divergensi (Csiszar, 1975) dinotasikan $D_{KL}(P \parallel Q)$ adalah jenis jarak statistik: ukuran bagaimana satu distribusi probabilitas P berbeda dari yang kedua, referensi distribusi probabilitas Q (Kullback & Leibler, 1951). Divergensi Kullback-Leibler (KL) adalah metrik non-simetris yang mengukur entropi relatif atau perbedaan informasi yang diwakili oleh dua distribusi. Hal ini dapat dianggap sebagai pengukuran jarak antara dua distribusi data yang menunjukkan betapa berbedanya kedua distribusi tersebut satu sama lain. Semakin mirip kedua distribusi yang diukur maka hasil pengukuran divergensi KL mendekati nol, dan semakin tidak mirip kedua distribusi maka hasil pengukuran akan semakin besar. Berikut rumus untuk mengukur kedua distribusi menggunakan divergensi KL.

$$D_{KL}(p(x) \parallel q(x)) = \sum_{x \in X} p(x) \ln \frac{p(x)}{q(x)} \quad (1)$$

Di mana:

D_{KL} = Jarak divergensi KL

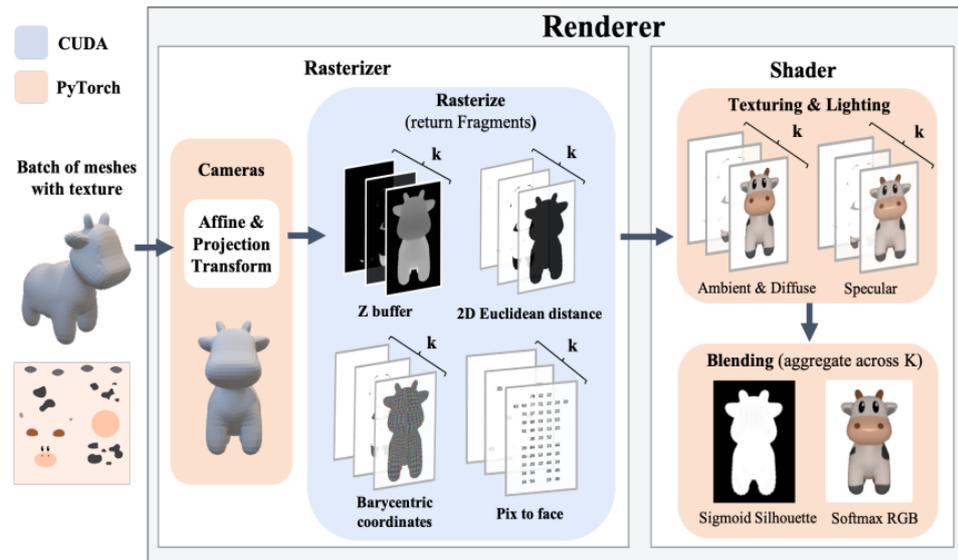
$p(x)$ = Distribusi p

$q(x)$ = Distribusi q

Dalam pemantauan model, divergensi KL digunakan untuk memantau lingkungan produksi, khususnya di sekitar data fitur dan prediksi. KL Divergence digunakan untuk memastikan bahwa data input atau output dalam produksi tidak berubah secara drastis dari data dasar. Garis dasar dapat berupa jendela produksi data pelatihan atau kumpulan data pelatihan atau validasi. Pemantauan penyimpangan dapat sangat berguna bagi tim yang menerima kebenaran lapangan yang tertunda untuk dibandingkan dengan keputusan model produksi. Tim-tim ini dapat mengandalkan perubahan dalam prediksi dan distribusi fitur sebagai proksi kinerja. Divergensi KL biasanya diterapkan pada setiap fitur secara independen; ini tidak dirancang sebagai ukuran fitur kovarian melainkan metrik yang menunjukkan bagaimana setiap fitur menyimpang secara independen dari nilai dasar.

2.1.6 Pytorch 3D

Pytorch 3D adalah sebuah *library* modular, efisien, dan operator *differentiable* untuk 3D *deep learning* (Ravi et al., 2020). Library ini mencakup modular perender terdiferensiasi yang cepat untuk *meshes* dan *point cloud*, memungkinkan pendekatan analisis sintesis. Perender terdiferensiasi atau *differentiable rendering* memproyeksikan data 3D ke gambar 2D, memungkinkan pendekatan *analysis by synthesis* dan *inverse rendering* (Patow & Pueyo, 2003) di mana prediksi 3D dapat dibuat hanya dengan menggunakan image-level supervision (Loper & Black, 2014).



Gambar 8 Modular *rendering pipeline* di PyTorch3D

Gambar 8 menunjukkan differentiable rendering yang digunakan pada pytorch3D. Perender memasukkan informasi *scene* (kamera, geometri, material, pencahayaan, tekstur) dan keluaran gambar. *Differentiable renderer* juga dapat *propagate gradients backward* dari gambar yang dirender ke informasi *scene* (Loper & Black, 2014), memungkinkan rendering untuk di *embedded* ke dalam *deep learning pipelines*. Perender pada pytorch 3D mengikuti tiga prinsip desain inti: diferensiasi, artinya menghitung gradien terhadap semua masukan, dimana hal ini bertujuan untuk menemukan fungsi minimum melalui penurunan sebuah *function loss* untuk mengupdate parameter; efisiensi, artinya berjalan cepat dan berskala hingga *meshes* dan gambar yang besar; dan modularitas, artinya pengguna dapat dengan mudah mengganti komponen untuk menyesuaikan fungsinya dengan kasus bereksperimen dengan formulasi alternatif. Seperti yang ditunjukkan pada Gambar, Pytorch 3D memiliki dua komponen utama: rasterizer memilih *faces* yang memengaruhi setiap piksel, dan shader menghitung warna piksel.

2.1.7 Protokol NoW

Estimasi bentuk wajah 3D dari satu gambar harus kuat terhadap variasi pencahayaan, pose kepala, ekspresi, rambut wajah, riasan, dan oklusi. Kekokohan

membutuhkan set pelatihan yang besar dari gambar-gambar yang bebas, yang secara konstruksi, tidak memiliki bentuk 3D yang benar. Hasil rekonstruksi 3D wajah selanjutnya perlu dievaluasi untuk menentukan kualitas dari metode rekonstruksi yang digunakan.

Salah satu tolak ukur yang bisa digunakan untuk evaluasi metode rekonstruksi wajah adalah tolak ukur NoW untuk tugas rekonstruksi wajah 3D dari gambar bermata tunggal. Tolak ukur ini menggunakan dataset “not quite in-the-wild” (NoW) sebagai data 2D untuk proses rekonstruksi 3D wajah. Tujuan tolak ukur ini adalah memperkenalkan metrik evaluasi standar untuk mengukur akurasi dan ketahanan metode rekonstruksi wajah 3D di bawah variasi sudut pandang, pencahayaan, dan oklusi umum.



Gambar 9 NoW Challenge

Gambar 9 menunjukkan contoh dari dataset NoW dan rekonstruksi 3D wajah dari gambar yang diberikan. Diberikan gambar monokular, tantangannya adalah merekonstruksi wajah 3D. Karena bentuk yang diprediksi terjadi dalam sistem koordinat lokal yang berbeda, bentuk 3D yang direkonstruksi disejajarkan (yaitu, rotasi, translasi, dan penskalaan opsional) ke pemindaian menggunakan serangkaian landmark yang sesuai antara prediksi dan pemindaian. Selanjutnya melakukan penyelarasan berdasarkan jarak *scan-to-mesh* (yang merupakan jarak absolut antara setiap *scan vertex* dan titik terdekat di permukaan *mesh*) antara pindaian *ground truth*, dan *mesh* yang direkonstruksi menggunakan landmarks alignment sebagai inisialisasi.

a. Dataset

Dataset gambar dikategorikan data yang diambil dalam empat tantangan; netral (620 gambar), ekspresi (675 gambar), oklusi (528 gambar), dan selfie (231 gambar). Netral, ekspresi dan oklusi berisi gambar wajah yang netral, ekspresif, dan tertutup sebagian dari semua subjek dalam berbagai tampilan, mulai dari tampilan depan hingga tampilan profil. Ekspresi berisi ekspresi wajah yang berbeda, seperti kebahagiaan, kesedihan, keterkejutan, rasa jijik, dan ketakutan. Oklusi berisi gambar dengan berbagai oklusi, misalnya kacamata, kacamata hitam, rambut wajah, topi atau kerudung. Untuk kategori selfie, para peserta diminta untuk mengambil foto selfie dengan iPhone, tanpa memberikan batasan pada ekspresi wajah yang ditampilkan. Gambar diambil di dalam dan di luar ruangan untuk memberikan variasi cahaya alami dan cahaya buatan.

b. Scan

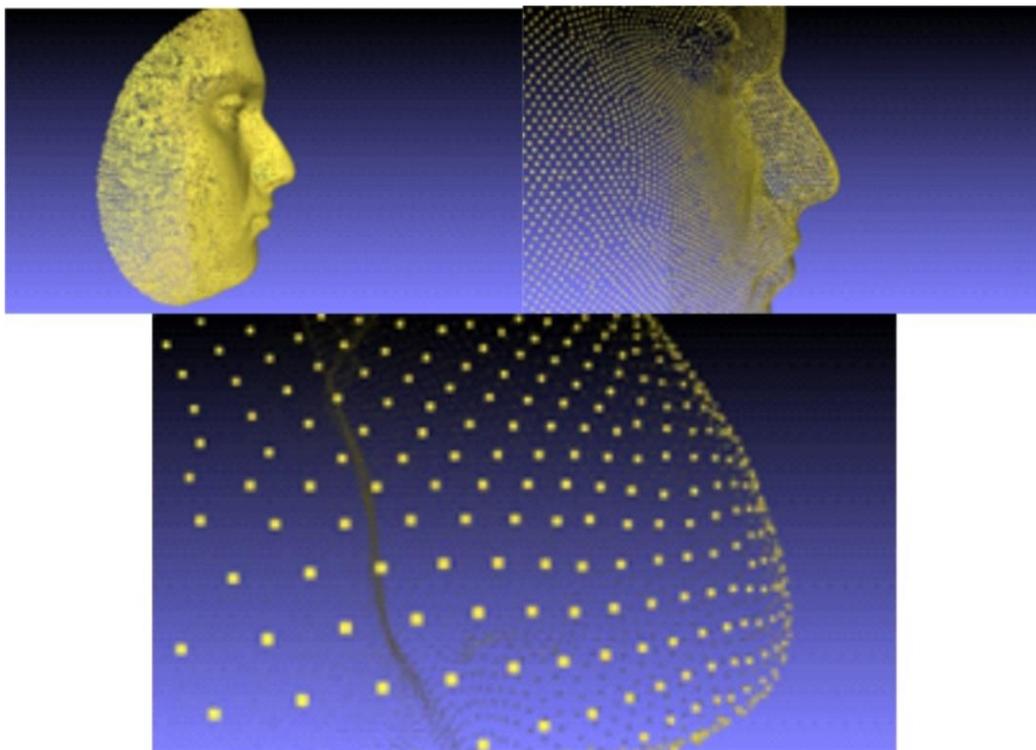
Untuk setiap subjek, dilakukan pemindaian terhadap kepala dalam ekspresi netral dengan sistem stereo aktif (3dMD LLC, Atlanta). Sistem multi-kamera terdiri dari enam pasangan kamera stereo skala abu-abu, enam kamera warna, lima proyektor pola bintang, dan enam panel LED putih. Geometri 3D yang direkonstruksi berisi sekitar 120 ribu *vertices* untuk setiap subjek. Setiap subjek mengenakan penutup rambut selama pemindaian untuk menghindari oklusi dan noise pemindai di daerah wajah atau leher akibat rambut. Tantangan untuk semua kategori adalah merekonstruksi wajah 3D netral yang diberikan gambar monokuler tunggal. Ekspresi wajah hadir dalam beberapa gambar, yang membutuhkan metode untuk memisahkan identitas dan ekspresi untuk mengevaluasi kualitas identitas yang diprediksi.

c. Data Processing

Sebagian besar metode rekonstruksi wajah 3D yang ada saat ini memerlukan *localization* wajah. Untuk mengurangi pengaruh langkah pra-pemrosesan ini, disediakan kotak pembatas untuk setiap gambar, yang menutupi wajah. Untuk mendapatkan kotak pembatas untuk semua gambar, pertama-tama akan menjalankan detektor wajah pada semua gambar, dan kemudian memprediksi titik kunci untuk setiap wajah yang terdeteksi. Secara

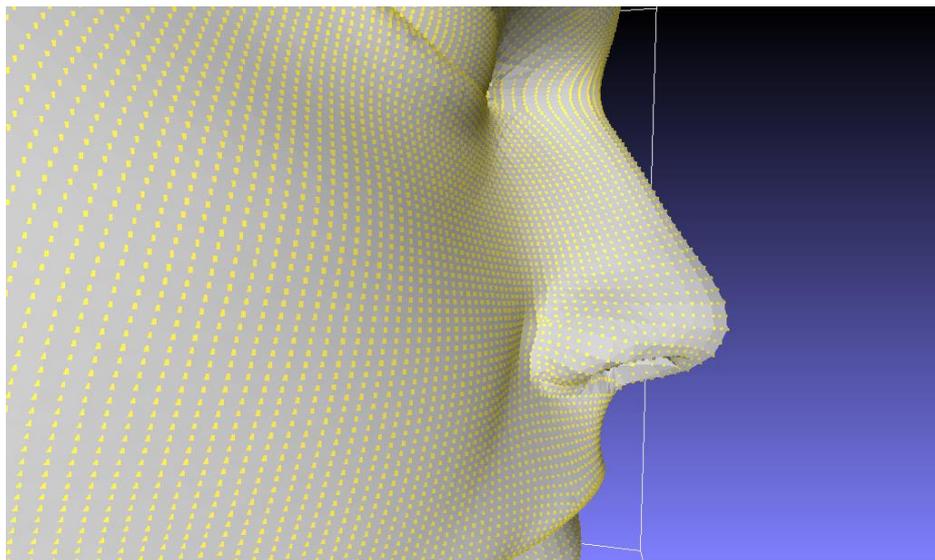
manual memilih *landmark* 2D untuk kasus kegagalan. Kemudian memperluas kotak pembatas *landmark* ke setiap sisi sebesar 5% (bawah), 10% (kiri dan kanan), dan 30% ke atas untuk mendapatkan kotak yang mencakup seluruh wajah termasuk dahi. Untuk pemindaian 3D, kami mengikuti protokol pemrosesan yang serupa dengan makalah ini. Untuk setiap pemindaian, bagian tengah wajah dipilih, dan pemindaian dipangkas dengan menghilangkan semua yang berada di luar radius yang ditentukan. Radius yang dipilih adalah radius khusus yang dihitung sebagai $0,7 \times (\text{jarak mata luar} + \text{jarak hidung})$.

Ground truth pada protokol NoW merupakan 3D wajah yang didapatkan dengan melakukan scan wajah secara langsung. *Ground truth* pada penelitian ini memanfaatkan dataset NoW *Challenge*.

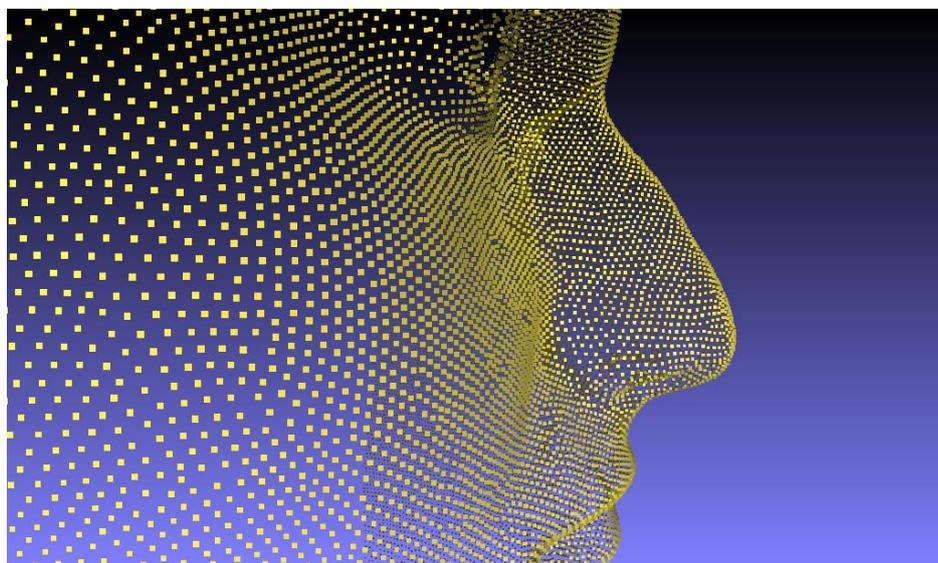


Gambar 10 *vertex* Pada 3D wajah

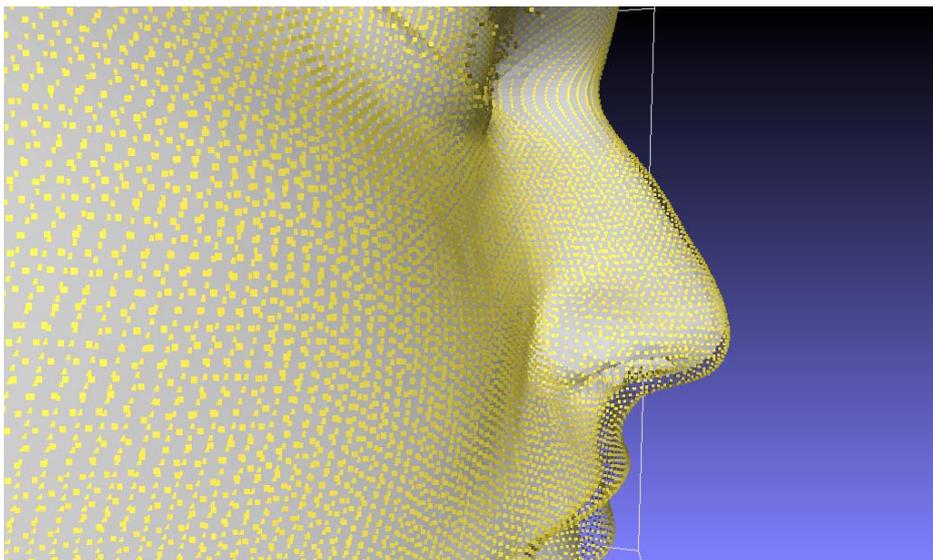
Gambar 10 menunjukkan Vertex adalah titik-titik yang dihubungkan dengan edge, dan merupakan komponen paling dasar dari objek 3 dimensi. Untuk memanipulasi posisi vertex dapat dilakukan dengan mengganti nilai kordinat x, y, dan z pada vertex tersebut.



Gambar 11 Bentuk dan *vertex* Pada Hasil 3D Wajah



Gambar 12 Vertex Pada 3D *Ground Truth*



Gambar 13 Vertex Pada Hasil 3D Wajah dan 3D *Ground Truth*

Gambar 11, 12, dan 13 menunjukkan gambaran pengukuran menggunakan *NoW Challenge*. Pada 3D wajah yang akan dievaluasi setiap vertex dari 3D *ground truth* yang didapatkan dari scan wajah secara langsung akan diukur berdasarkan jarak absolut antara setiap *scan vertex* dan titik terdekat di permukaan *mesh* atau bentuk dari 3D wajah yang direkonstruksi. Pengukuran jarak berbeda kemudian diringkas secara kumulatif dalam median, rata-rata, dan standar deviasi (Z.-H. Feng et al., 2018).

2.2 Metode Penyelesaian Masalah

2.2.1 *State of The Art*

Tabel 2 *State of the art*

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <.,=>
1.	Judul: Rekonstruksi 3D wajah dari 2D menggunakan 3D	Objek: 3D Morphable Model dan Face Recognition	3D Morphable Model dan ElasticFace		

Tabel 2 (lanjutan)

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <.,=,>
	<p><i>Morphabel Model</i> dan <i>ElasticFace</i></p> <p>Penulis : Abd Salam At Taqwa</p>	<p>Permasalahan: Optimasi koefisien 3D Morphable Model untuk <i>3D face Reconstruction</i></p>			
2.	<p>Judul: A Lightweight Monocular 3D Face Reconstruction Method Based on Improved 3D Morphing Models (You et al., 2023)</p> <p>Penulis : You et al.</p> <p>Tahun : 2023</p> <p>Penerbit : Sensors</p>	<p>Objek: 3D Morphing Model (3DMM) dan Lightweight Network</p> <p>Permasalahan: Rekonstruksi 3D wajah menggunakan <i>convolutional neural networks</i> memiliki beban komputasi yang berat dan mengurangi kecepatan komputasi</p>	<p>Rekonstruksi 3D wajah menggunakan modifikasi <i>lightweight network</i></p>	<p>parameter model dan kompleksitas GFLOP masing-masing mencapai pengurangan sebesar 65,7% dan 86,2%, menunjukkan peningkatan yang signifikan</p>	<p><</p>

Tabel 2 (lanjutan)

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <.,=>
3.	<p>Judul: Robust Model-based Face Reconstruction through Weakly-Supervised Outlier Segmentation (C. Li et al., 2023)</p> <p>Penulis : Li, C., et al.</p> <p>Tahun : 2023</p> <p>Penerbit : IEEE/CVF Conference on Computer Vision and Pattern Recognition</p>	<p>Objek: 3D Rekonstruksi Wajah dan Segmentasi</p> <p>Permasalahan: Metode rekonstruksi 3D wajah biasa mengambil bagian yang bukan wajah sebagai data untuk proses rekonstruksi</p>	<p>Rekonstruksi 3D wajah menggunakan Face-autoencoder dengan menghilangkan outlier</p>	<p>CelebA-Unoccluded = 8.38 ± 0.42</p> <p>CelebA-Occluded = 8.71 ± 0.48</p> <p>CelebA-Overall = 8.55 ± 0.48</p> <p>AR-Overall = 8.93 ± 0.35</p>	<p><</p>
4.	<p>Judul: 3D Face Reconstruction with Dense Landmarks (Wood et al., 2022)</p> <p>Penulis : Wood, E., et al.</p> <p>Tahun :</p>	<p>Objek: 3D Morphable Model dan Landmark</p> <p>Permasalahan: Bagaimana merekonstruksi 3D wajah hanya</p>	<p>Menyesuaikan 3D Morphable Model dengan dense landmark yang telah dideteksi</p>	<p>Dapat merekonstruksi 3D wajah dengan kecepatan lebih dari</p>	<p>=</p>

Tabel 2 (lanjutan)

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <.,=>
	2022 Penerbit : European Conference on Computer Vision	bergantung pada deteksi landmark		150FPS pada satu thread CPU	
5.	Judul: Weakly-supervised multi-face 3D reconstruction (Zhang et al., 2021) Penulis : Zhang et al. Tahun : 2021 Penerbit : arXiv preprint arXiv:2101.02000.	Objek: 3D Morphable Model, dan Multi-Face Reconstruction Permasalahan: Sebagian besar rekonstruksi konvensional mengabaikan posisi dan orientasi kepala global untuk setiap individu dalam rekonstruksi 3D	Single shot learning berbasis rekonstruksi weakly supervised multi-face	NME AFLW200-3D Dataset = 3.62	=
6.	Judul: Self-Supervised Monocular 3D Face	Objek: 3D Morphable Model, Multi-	Memanfaatkan konsistensi geometri	RMSE BU-3DFE	=

Tabel 2 (lanjutan)

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <.,=>
	<p>Reconstruction by Occlusion-Aware Multi-view Geometry Consistency(Shang et al., 2020)</p> <p>Penulis : Shang et al.</p> <p>Tahun : 2020</p> <p>Penerbit : European Conference on Computer Vision</p>	<p>view geometry consistency</p> <p>Permasalahan: <i>Monocular Depth Estimation</i> (MDE) hanya memprediksi peta kedalaman dan pose relatif di antara tampilan tanpa menyimpulkan intrinsik kamera.</p>	<p><i>multi-view</i> untuk mengurangi ambiguitas dari estimasi pose wajah dan rekonstruksi kedalaman dalam proses pelatihan.</p>	<p><i>dataset</i> = mean = 1.55</p>	
7.	<p>Judul: Accurate 3D Face Reconstruction with Weakly-Supervised Learning: From Single Image to Image Set (Y. Deng et al., 2019)</p> <p>Penulis : Deng et al.</p> <p>Tahun : 2019</p>	<p>Objek: 3D Morphabel Model</p> <p>Permasalahan: Bagaimana meningkatkan akurasi rekonstruksi wajah secara <i>weakly supervised learning</i></p>	<p>Menggunakan <i>Photometric Loss</i> dan Facenet sebagai error untuk memperbaharui koefisien <i>3D Morphabel Model</i></p>	<p>Mean Root Mean Squared Error (RMSE) <i>Cooperative</i>=1.66 ±0.52 <i>Indor</i> = 1.66±0.46</p>	=

Tabel 2 (lanjutan)

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <.,>
	Penerbit : Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops			<i>Outdoor</i> = 1.69 ± 0.5 3	
8.	Judul: Unsupervised Training for 3D Morphable Model Regression (Genova et al., 2018) Penulis : Genova et al. Tahun : 2018 Penerbit : Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition	Objek: 3D Morphable Model dan Face Recognition Permasalahan: Regresi koefisien 3D Morphable Model secara unsupervised	Menggunakan perhitungan <i>loss function</i> dari <i>face recognition</i> untuk update koefisien 3D Morphable Model	Mean Root Mean Squared Error (RMSE) Cooperati ve= 1.78 ± 0.54 <i>Indor</i> = 1.78 ± 0.5 2 <i>Outdoor</i> = 1.76 ± 0.5 4	>
9.	Judul: Self-supervised	Objek: 3D Morphable	Metode pelatihan <i>end-</i>	NA	>

Tabel 2 (lanjutan)

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <.,=>
	<p>Multi-level Face Model Learning for Monocular Reconstruction at over 250 Hz (Tewari et al., 2018)</p> <p>Penulis : Tewari et al.</p> <p>Tahun : 2018</p> <p>Penerbit : Proceedings of the IEEE conference on computer vision and pattern recognition</p>	<p><i>Model</i> dan Input RGB di atas 250 Hz</p> <p>Permasalahan: Geometri wajah yang direkonstruksi menjadi 3D kurang maksimal</p>	<p><i>to-end</i> untuk regressor yang efisien untuk memperkirakan geometri identitas berkualitas tinggi, ekspresi wajah, dan pantulan kulit berwarna, bersama 2) parameterisasi model wajah <i>multi-level</i> yang lebih baik</p>		
10.	<p>Judul: MobileFace: 3D face reconstruction with efficient CNN regression (Chinaev et al., 2018)</p> <p>Penulis : Chinaev et al.</p>	<p>Objek: 3D Morphable Model</p> <p>Permasalahan: Metode <i>iterative</i> yang membutuhkan waktu lama yang</p>	<p>MobileNet CNN</p>	<p>Area Under the Curve (AUC) 0.854</p>	<p>></p>

Tabel 2 (lanjutan)

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <.,=>
	<p>Tahun : 2018</p> <p>Penerbit : Proceedings of the European Conference on Computer Vision (ECCV) Workshops</p>	<p>dapat mencegah dijalankan secara <i>real time</i></p>			
11.	<p>Judul: Joint 3D Face Reconstruction and Dense Alignment with Position Map Regression Network(Y. Feng et al., 2018)</p> <p>Penulis : Feng et al.</p> <p>Tahun : 2018</p> <p>Penerbit : Proceedings of the European conference on computer vision (ECCV)</p>	<p>Objek: <i>3D Face & UV</i></p> <p>Permasalahan: Representasi dari volumetrik penelitian (Jackson et al., 2017) membuang makna semantik dari titik-titik, sehingga jaringan perlu meregresi seluruh volume untuk mengembalikan bentuk wajah yang hanya merupakan</p>	<p>Teknik rekonstruksi 3D wajah yang disebut <i>Position map Regression Network</i> (PRN) berdasarkan koordinat tekstur peta posisi UV</p>	<p>NME AFLW20 00-3D = 3.62</p>	<p>></p>

Tabel 2 (lanjutan)

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <.,=>
		bagian dari volume			
12.	<p>Judul: SfSNet: Learning Shape, Reflectance and Illuminance of Faces ‘in the wild’(Sengupta et al., 2018)</p> <p>Penulis : Sengupta et al.</p> <p>Tahun : 2018</p> <p>Penerbit : Proceedings of the IEEE conference on computer vision and pattern recognition</p>	<p>Objek: <i>3D Morphable Model</i> dan Deteksi Bayangan</p> <p>Permasalahan: Bagaimana Rekonstruksi 3D wajah berdasarkan bayangan objek</p>	<p>3D Morphable Model dan perhitungan loss bentuk bayangan</p>	NA	>
13	<p>Judul: End-to-end 3D face reconstruction with deep neural networks(Dou et al., 2017)</p> <p>Penulis : Dou et al.</p>	<p>Objek: 3D Morphable Model</p> <p>Permasalahan: Training Regresi koefisien <i>3D Morphable Model</i> dengan akurasi</p>	<p><i>End-to-end</i> dan memprediksi parameter <i>3D Morphable Model</i> yang optimal dengan satu</p>	<p>Mean and standard deviation of RMSE (mm). UHDB31 Dataset = 2.73 ± 0.7</p>	>

Tabel 2 (lanjutan)

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <.,=>
	<p>Tahun : 2017</p> <p>Penerbit : Proceedings of the IEEE conference on computer vision and pattern recognition</p>	lemah	operasi <i>forward.</i>	1 FRGC2 = 3.71±3.0 5 BU- 3DFE = 4.52±1.1 1	
14	<p>Judul: MoFA: Model-based Deep Convolutional Face Autoencoder for Unsupervised Monocular Reconstruction(Tewari et al., 2017)</p> <p>Penulis : Tewari et al.</p> <p>Tahun : 2017</p> <p>Penerbit : Proceedings of the IEEE International Conference on Computer Vision</p>	<p>Objek: Convolutional Autoencoder dan 3D Morphable Model</p> <p>Permasalahan: Bagaimana Rekonstruksi 3D wajah menggunakan <i>convolutional autoencoder</i></p>	Model berbasis Convolutional Autoencoder	Verificati on test = 77% accurasy	>

Tabel 2 (lanjutan)

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <.,=>
	Workshops				
15.	<p>Judul: Learning detailed face reconstruction from a single image (Richardson et al., 2017)</p> <p>Penulis : Richardson et al.</p> <p>Tahun : 2017</p> <p>Penerbit : Proceedings of the IEEE conference on computer vision and pattern recognition</p>	<p>Objek: 3D Morphable Model</p> <p>Permasalahan: Ambiguitas tambahan muncul karena proyeksi wajah ke gambar juga bergantung pada tekstur dan sifat materialnya, kondisi pencahayaan, dan arah pandang.</p>	<p>CNN dan Coarse-to-fne details</p>	<p>Ave. Depth Err. [mm] 3.22</p> <p>90% Depth Err. [mm] 6.69</p>	>
16.	<p>Judul : 3D Face Reconstruction by Learning from Synthetic Data (Richardson et al., 2016)</p> <p>Penulis : Richardson et al.</p>	<p>Objek: 3D Morphable Model</p> <p>Permasalahan: Dataset 3D wajah terbatas</p>	<p>Membuat data sintesis menggunakan geometri wajah, kemudian menggunakan 3D Morphable</p>	NA	>

Tabel 2 (lanjutan)

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <.,=,>
	<p>Tahun : 2016</p> <p>Penerbit : 2016 fourth international conference on 3D vision (3DV)</p>		<p><i>Model dan Shape From Shading</i> dengan landmarks dan deep learning</p>		

Tabel 2 menunjukkan beberapa penelitian-penelitian terkait 3D rekonstruksi wajah yang menjadi dasar dilakukannya penelitian.

2.2.2 Metode Penelitian

Berdasarkan pada tabel State of The Art di atas, dapat kita Tarik kesimpulan bahwa dapat dilakukan penelitian rekonstruksi 3D wajah menggunakan *3D morphable model* secara *weakly supervised*. Masalah yang diangkat adalah bagaimana merekonstruksi 3D wajah dari gambar 2D wajah. Permasalahan ini cukup menarik untuk dibahas karena cukup rumit dan dapat menyelesaikan beberapa masalah seperti pengenalan wajah, hiburan, aplikasi medis, forensik, ilmu kognitif, ilmu saraf, dan psikologi. Penelitian ini berfokus pada peningkatan prediksi koefisien *3D morphable model* agar menghasilkan 3D wajah yang sama dengan gambar 2D wajah.

2.3 Target Hasil Penelitian

Berdasarkan tabel *State of The Art* penelitian yang telah dilakukan, target hasil dalam penelitian ini adalah penelitian ini mampu melakukan rekonstruksi 3D wajah dari gambar input 2D wajah. Gambar 2D wajah digunakan untuk memprediksi koefisien dari *3D morphable model*. Koefisien yang didapatkan

kemudian dirender kedalam 3D wajah. Objek 3D wajah kemudian dilakukan kesamaan terhadap gambar 2D wajah menggunakan ekstraksi fitur. Hasil dari fitur yang didapatkan menggunakan *deep face recognition* kemudian dilakukan perhitungan *cosine similarity*. Hasil dari perhitungannya kemudian digunakan untuk *backproagation* untuk merubah koefisien dari *3D morphable model*. Penelitian ini diharapkan mampu meningkatkan rekonstruksi 3D wajah menjadi lebih baik sesuai dengan gambar 2D wajah yang diberikan.

2.4 Kerangka Pikir

