

DISERTASI

MODEL DETEKSI KEBOHONGAN MELALUI ANALISIS WAJAH DAN SUARA DENGAN PENDEKATAN DEEP LEARNING

*Lie Detection Model Through Face and Voice Analysis with Deep
Learning Approach*

DEWI KUSUMAWATI

D053211004



PROGRAM STUDI DOKTOR ILMU TEKNIK ELEKTRO

DEPARTEMEN TEKNIK ELEKTRO

FAKULTAS TEKNIK

UNIVERSITAS HASANUDDIN

GOWA

2024



Optimized using
trial version
www.balesio.com

DISERTASI

MODEL DETEKSI KEBOHONGAN MELALUI ANALISIS WAJAH DAN SUARA DENGAN PENDEKATAN DEEP LEARNING

DEWI KUSUMAWATI

NIM: D053211004

Telah dipertahankan di hadapan Panitia Ujian Disertasi yang dibentuk dalam
rangka penyelesaian studi pada Program Doktor Teknik Elektro Fakultas
Teknik Universitas Hasanuddin
pada tanggal 23 April 2024
dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui
Promotor



Dr. Amil Ahmad Ilham, ST., M.IT.
NIP.19731010 199802 1 001

Ko-Promotor



Prof. Dr. Ir. H. Andani Achmad, M.T.
NIP.19601231 198703 1 022

Ko-Promotor



Dr. Ir. Ingrid Nurtanio, MT
NIP.19610813 198811 2 001

Dekan Fakultas Teknik
Universitas Hasanuddin



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

Ketua Program Studi S3
Teknik Elektro



Prof. Dr. Ir. H. Andani Achmad, M.T.
NIP.19601231 198703 1 022



Optimized using
trial version
www.balesio.com

ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE
11 Tahun 2008 Pasal 5 Ayat 1
Elektronik dan/atau Dokumen Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"

PERNYATAAN KEASLIAN DISERTASI DAN PELIMPAHAN HAK CIPTA

Yang bertanda tangan di bawah ini

Nama : Dewi Kusumawati

Nomor mahasiswa : D053211004

Program studi : S3 Teknik Elektro

Dengan ini menyatakan bahwa, disertasi ini berjudul “Model Deteksi Kebohongan Melalui Analisis Wajah Dan Suara Dengan Pendekatan Deep Learning” adalah benar karya saya dengan arahan dari komisi pembimbing (Dr. Amil Ahmad Ilham, ST., MIT sebagai Promotor, Prof. Dr. Ir. H. Andani Achmad, M.T sebagai co-promotor-1 dan Dr. Ir. Ingrid Nurtanio, MT sebagai co-promotor-2). Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apapun 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 disertasi ini. Sebagian dari isi disertasi ini telah dipublikasikan di Prosiding (2022 6th International Conference on Information Technology, Information System and Electrical Engineering, DOI. 10.1109/ICITISEE57756.2022.10057748) sebagai artikel dengan judul “VGG-16 And VGG-19 Architecture Models In Lie Detection Using Image Processing”. Kemudian di jurnal (JOIV : International Journal on Informatics Visualization, DOI. 10.30630/joiv.8.1.2062 Vol 8, No 1) .

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

Gowa, Mei 2024

Yang menyatakan



Dewi Kusumawati

KATA PENGANTAR

Dengan memanjatkan puji syukur kehadiran Allah Subhanahu Wata'ala atas karunia dan rahmat-Nya sehingga penulisan penelitian disertasi dengan judul “MODEL DETEKSI KEBOHONGAN MELALUI ANALISIS WAJAH DAN SUARA DENGAN PENDEKATAN DEEP LEARNING” dapat penulis susun dan selesaikan.

Pada penyusunan penelitian disertasi ini, penulis telah berusaha secara maksimal untuk mencari referensi yang ada kaitanya dengan judul yang diangkat, baik dalam bentuk artikel/jurnal penelitian, buku, maupun dari *website* atau internet. Namun semua itu tidak dapat tersusun dengan baik tanpa adanya dukungan dari berbagai pihak yang telah membantu penulis dalam menyusun disertasi yang sesuai dengan kaidah-kaidah ilmiah yang dipersyaratkan. Olehnya itu, melalui kesempatan ini penulis menyampaikan secara khusus apresiasi dan terima kasih banyak secara tertulis kepada:

1. Bapak Dr. Amil Ahmad Ilham, ST., MIT., selaku promotor, Bapak Prof. Dr. Ir. H. Andani Achmad, M.T, selaku co-promotor 1 dan ibu seluruh Dr,Ir.Ingrid Nurtanio,MT., selaku co-Promotor 2 yang telah meluangkan waktu dan pikiran dengan penuh kesabaran dalam membimbing dan memberikan saran sehingga disertasi ini dapat tersusun dengan baik dan terstruktur.
2. Bapak Prof.Dr.Eng.Ir.Muhammad Isran Ramli,S.T.,M.T.,IPM.,ASEAN Eng, Bapak Prof.Dr. Eri Prasetyo Wibowo, S. Si. MMSI. Dr.Ir.Zahir Zainuddin,M.Sc, Bapak Dr.Eng.Wardi,ST,M.Eng, Ibu Dr.Eng.Ir.Dewiani,M.T,IPM, dan Ibu Merna Baharuddin,ST.,M.Tel.Eng.,PhD selaku komisi penguji yang telah berkontribusi dalam memberikan saran dan masukan untuk penyusunan disertasi ini.



ektor Universitas Hasanuddin, Bapak Prof.Dr.Ir.Jamaluddin Jompa,M.Sc., ekan Fakultas Teknik, Bapak Prof.Dr.Eng.Ir.Muhammad Isran amli,S.T.,M.T.,IPM.,ASEAN Eng., Ketua Departemen Teknik Elektro, Ibu

Dr.Eng.Ir.Dewiani,MT.,IPM, Ketua Program Studi S3 Teknik Elektro, Bapak Prof. Dr. Ir. H. Andani Achmad, M.T, dan seluruh dosen serta staf S3 Program studi Teknik Elektro yang telah memberikan dukungan dan layanan selama proses studi penulis.

4. Ketua Yayasan Pendidikan STMIK Bina Mulia Palu Ibu Dr.Samsinar Z Moga, Ketua STMIK Bina Mulia Palu Bapak Ir.Burhanuddin Andi Masse.,M.Kom, Wakil Ketua Bidang Kurikulum Ibu Fitriyanti Andi Masse.,M.Kom.,Wakil Ketua bidang Kemahasiswaan Bapak H.Abdul Razak Ramli.,SE.,MM.,dan seluruh pengelola, dosen dan tenaga kependidikan STMIK Bina Mulia Palu yang telah mendukung penulis selama dalam melaksanakan studi.
5. Kepala Lab Forensik Mabes POLRI beserta tim, Kepala Idensos Densus 88 Sulawesi Tengah dan Kepala Identifikasi POLDA SULTENG yang telah membantu penulis dalam penyediaan dataset, mengolah data dan membantu dalam memvalidasi data yang penulis gunakan
6. Para sahabat, rekan seangkatan S3 Elektro UNHAS dan seluruh mahasiswa Teknik Elektro UNHAS. Terima kasih atas bantuan dan dukungannya selama ini kepada penulis.

Ucapan terima kasih yang sebesar-besarnya kepada suami, anak-anak, dan serta seluruh keluarga yang telah memberikan semangat dan dukungan moril kepada penulis selama menempuh studi.

Penulis telah berusaha secara maksimal dalam penyusunan penelitian disertasi ini, namun tidak menutup kemungkinan di dalamnya masih terdapat banyak kekurangan baik pada aspek penulisan maupun pada aspek konten, oleh karena itu saran, arahan, masukan, petunjuk dan pencerahan, senantiasa penulis harapkan demi perbaikan kearah yang lebih baik lagi.

Akhir kata penulis mengucapkan terima kasih dan mohon maaf atas segala kekurangan yang ada dalam penulisan penelitian ini. Semoga Allah SWT selalu merahmati dan meberikan keberkahan bagi kita semua.

Gowa, Mei 2024

Penulis



ABSTRAK

Dewi Kusumawati. “Model Deteksi Kebohongan Melalui Analisa Wajah dan Suara Menggunakan Pendekatan Deep Learning” (dibimbing oleh **Amil Ahmad Ilham, Andani Achmad, Ingrid Nurtanio**)

Penelitian ini mengembangkan suatu model deteksi kebohongan yang menggabungkan analisis wajah dan suara menggunakan pendekatan *deep learning*. Kebohongan merupakan perilaku kompleks yang dapat tercermin dalam ekspresi wajah dan karakteristik suara seseorang. Model ini dirancang untuk meningkatkan akurasi dan ketepatan dalam mengidentifikasi indikasi kebohongan dengan memanfaatkan kekuatan representasi fitur dari metode *deep learning*. Dalam proses analisis wajah, kami menggunakan arsitektur *Convolutional Neural Network* (CNN) dengan menambahkan layer *multihead attention*, dan menambahkan algoritma untuk menghitung perubahan ukuran pupil manusia dalam ekstraksi pola fitur dari ekspresi wajah berdasarkan rekaman video. Sebagai pendukung, analisis suara dilakukan menggunakan jaringan rekuren *Long Short Term Memory* (LSTM), dengan ekstraksi fitur MFCC untuk menangkap pola temporal dalam suara yang dapat mengisyaratkan ketidakjujuran. Model yang dihasilkan dari setiap fitur yaitu dari wajah dan suara digabungkan dengan model fusi untuk meningkatkan performa model dalam deteksi kebohongan. Model diperkuat melalui pelatihan menggunakan dataset luas yang mencakup berbagai situasi dan konteks kebohongan dan menggunakan dataset yang dibangun sendiri dan telah tervalidasi. Pengujian lintas dataset dilakukan untuk mengevaluasi generalitas dan kinerja model.

Hasil eksperimen menunjukkan bahwa model yang diusulkan berhasil mencapai tingkat akurasi deteksi kebohongan yang signifikan, dengan memanfaatkan keunggulan integrasi output model dari analisis wajah dan suara dimana hasil rata-rata yang didapatkan dari akurasi dan presisi sebesar 100%. Penelitian ini memberikan kontribusi penting dalam pengembangan model deteksi kebohongan dengan mengkombinasikan model CNN untuk wajah dan model LSTM untuk suara.

Kata kunci: CNN, *deep learning*, deteksi kebohongan, LSTM, MFCC, pupil, suara, *spatial attention*.



ABSTRACT

Dewi Kusumawati. "Lie Detection Model Through Face and Voice Analysis with Deep Learning Approach" (supervised by **Amil Ahmad Ilham, Andani Achmad, Ingrid Nurtanio**)

This research develops a lie detection model that combines face and voice analysis using a deep learning approach. Deception is a complex behavior that can be reflected in facial expressions and vocal characteristics. The model is designed to enhance accuracy and precision in identifying indications of deception by leveraging the feature representation strength of deep learning methods. In the face analysis process, we employ a Convolutional Neural Network (CNN) architecture with the addition of spatial attention layers, optical flow algorithms, and Gaussian Mixture Model to extract feature patterns from facial expressions recorded in videos. As a complement, voice analysis is conducted using recurrent neural networks, such as Long Short Term Memory (LSTM), employing MFCC feature extraction to capture temporal patterns in the voice that may indicate dishonesty. Furthermore, both facial and vocal features are combined using ensemble methods to improve the model's performance in lie detection. The model is strengthened through training using a comprehensive dataset covering various deception situations and contexts, including a self-built and validated dataset. Cross dataset testing is performed to evaluate the model's generalization and performance.

The experimental results show that the proposed model successfully achieves a significant level of accuracy in lie detection, utilizing the advantages of integrating face and voice analysis, with an average accuracy and precision of 100%. This research contributes significantly to the development of lie detection models by combining facial and vocal features.

Keywords: CNN, Deep learning, Lie Detection, LSTM, MFCC, Pupil, Spatial attention, Voice.



DAFTAR ISI

HALAMAN JUDUL	
PERSETUJUAN DISERTASI	
PERNYATAAN KEASLIAN DISERTASI DAN PELIMPAHAN HAK CIPTA	
KATA PENGANTAR.....	i
ABSTRAK	iii
ABSTRACT	iv
DAFTAR ISI.....	v
DAFTAR TABEL	viii
DAFTAR GAMBAR.....	ix
DAFTAR LAMPIRAN	xi
BAB I PENDAHULUAN.....	1
1.1 Latar belakang.....	1
1.2 Rumusan masalah.....	4
1.3 Tujuan penelitian.....	4
1.4 Manfaat penelitian.....	5
1.5 Batasan masalah penelitian	5
1.6 Ruang lingkup penelitian	5
BAB II KERANGKA KONSEPTUAL DAN HIPOTESIS PENELITIAN	8
2.1 Kerangka konsep penelitian	8
2.1.1 State of the art	13
2.1.2 Posisi penelitian	16
2.1.3 Tahapan penelitian	18
2.1.4 Jenis penelitian.....	23
2.1.6 Instrumen penelitian.....	24
2.1.7 Hipotesis penelitian.....	25
2.1.8 Gambaran Umum Penelitian	26
BAB III MODEL ARSITEKTUR VGG 16 DAN VGG 19 DALAM DETEKSI	
KEBOHONGAN MENGGUNAKAN IMAGE PROCESSING	30
Pendahuluan	30
Related Work	32
Materials and Method	33
Deep Learning.....	33



3.3.2 CNN	33
3.3.3. VGG Net	34
3.3.4. Augmentasi citra	35
3.3.5 Dropout	35
3.3.6 Fungsi Optimasi	36
3.3.7 Model Evaluasi.....	37
3.4. Model Yang dikembangkan	37
3.5 .Pembahasan.....	45
3.6 Kesimpulan	49
BAB IV PERFORMANCE ANALYSIS OF FITUR MEL FREQUENCY CEPSTRAL COEFFICIENT (MFCC) AND SHORT TIME FOURIER TRANSFORM (STFT) INPUT FOR LIE DETECTION USING CNN	51
4.1 Pendahuluan	52
4.2 Related Work	54
4.3 Material dan Metode	54
4.3.1 Konsep Dasar Kebohongan.....	55
4.3.2 Metode deteksi kebohongan yang sudah ada	55
4.3.3 MFCC dan STFT	57
4.3.4 Prosodic.....	58
4.3.5 Optimizer.....	58
4.3.6 Dropout	59
4.3.7 CNN	59
4.4 Gambaran umum desain model.....	60
4.5 PEMBAHASAN UMUM	65
4.5.1 Deskripsi Data.....	65
4.6 Implikasi penelitian.....	75
4.7 Kesimpulan	75
BAB V PENELITIAN LANJUTAN	77
5.1 Pengaruh Penambahan Layer Attention Pada Convolutional Neural Network (CNN) Dalam Deteksi Kebohongan Melalui Fitur Wajah.....	77
5.1.1 Pendahuluan	77
5.1.2. Material and Methods.	80
5.1.3. Desain arsitektur.....	85
4. Model validasi	94
5. Result and Discussion	96
5 Analisis Performa.....	98



5.1.7 Kesimpulan	102
5.2 Peningkatan Kinerja Model <i>Mel Frequency Cepstral Coefficients</i> (MFCC) dan <i>Long-Short Term Memory</i> (LSTM) dalam Deteksi Kebohongan Melalui Suara.....	102
5.2.1. Pendahuluan	103
5.2.2. Materials and Methods	106
5.2.3 Pembahasan Umum.....	110
5.2.4 Model Evaluasi.....	110
5.2.5 Hasil	110
5.2.6 Kesimpulan	115
5.3 Deteksi Kebohongan Melalui Wajah Menggunakan <i>Transfer Learning</i> Pada <i>Convolutional Neural Network</i>	115
5.3.1 Pendahuluan	116
5.3.2 Related work	119
5.3.3 Material and methods	121
5.3.4 Hasil dan Pembahasan.....	133
5.3.5 Kesimpulan	135
5.4 Pendekatan Metode Fusion Dalam Deteksi Kebohongan Dengan Multimodal.....	135
5.4.1 Latar Belakang	136
5.4.2 Related work	138
5.4.3 Model dan Pembahasan Umum	140
5.4.4 Kesimpulan	147
BAB VI PEMBAHASAN UMUM	148
6.1 Model Arsitektur VGG 16 Dan VGG 19 Dalam Deteksi Kebohongan Menggunakan Image Processing.....	148
6.2 Performance Analysis of Fitur Mel Frequency Cepstral Coefficient (MFCC) And Short Time Fourier Transform (STFT) Input For Lie Detection Using CNN.	161
6.3 Pengaruh Penambahan <i>Layer Attention</i> pada <i>Convolutional Neural Network</i> (CNN) Dalam Deteksi Kebohongan Melalui Fitur Wajah.....	172
6.4 Pendekatan Metode Fusion Dalam Deteksi Kebohongan Dengan Multimodal.....	181
BAB VII KESIMPULAN DAN SARAN	197
7.1 Kesimpulan.....	197
7.2 Saran.....	198
DAFTAR PUSTAKA.....	199
LAMPIRAN.....	210



DAFTAR TABEL

Tabel 1 <i>State Of the art</i> penelitian terkait	13
Tabel 2 Rumusan hiotesa dan penelitian yang terkait.....	26
Tabel 3 Skenario eksperimen	42
Tabel 4 Performance untuk tiap algoritma.....	42
Tabel 5 Performance MFCC-CNN	66
Tabel 6 <i>Performance</i> STFT-CNN	67
Tabel 7 <i>Performance Prosodic</i> -CNN for Lie.....	68
Tabel 8 <i>Proposed model</i>	71
Tabel 9 Nilai interpretasi AUC	96
Tabel 10 Nilai performa setiap model.....	98
Tabel 11 Perbandingan performa model yang dikembangkan dengan <i>state of the art</i>	101
Tabel 12 Parameter yang digunakan	111
Tabel 13 Performance improve model untuk deteksi kebohongan melalui suara	112
Tabel 14 Parameter yang digunakan	130
Tabel 15 Hasil beberapa skenario eksperimen.....	131
Tabel 16 Dataset yang telah digunakan dalam deteksi kebohongan(Gupta <i>et al.</i> , 2019a).....	138
Tabel 17 Perbandingan performa model yang dikembangkan dengan <i>state of the art</i>	146
Tabel 18 Spesifikasi yang akan digunakan pada pengujian model.....	188
Tabel 19 Perbandingan performa dalam deteksi kebohongan dengan model yang telah ada	192



DAFTAR GAMBAR

Gambar 1 Representasi <i>deep learning</i>	10
Gambar 2 Kerangka konsep penelitian	12
Gambar 3 Posisi penelitian.....	16
Gambar 4 Diagram fishbone penelitian	19
Gambar 5 Model hipotesis dari penelitian	25
Gambar 6 Model yang dikembangkan	27
Gambar 7 Convolutional Neural Network (CNN)	34
Gambar 8 Ilustrasi Very Depth Convolutional Network (VGG)	34
Gambar 9 Model Yang dikembangkan	38
Gambar 10 Contoh hasil augmentasi citra	40
Gambar 11 Grafik Nilai Loss Algoritma Adam.....	44
Gambar 12 Grafik Nilai Akurasi Algoritma Adam.....	44
Gambar 13 Kesamaan antar label prediksi dan label aktual	45
Gambar 14 Arsitektur VGG-16.....	46
Gambar 15 Ilustrasi jaringan syaraf biasa(a), Jaringan yang dengan layer dropout(b).....	47
Gambar 16 Investigasi menggunakan polygraph.....	55
Gambar 17 Sample spectrogram images of audio.....	57
Gambar 18 Sample MFCC of audio	58
Gambar 19 Arsitektur CNN	60
Gambar 20 Gambaran umum prancangan model.....	60
Gambar 21 Perbandingan Akurasi MFCC, STFT dan <i>Prosodic</i>	69
Gambar 22 Perbandingan AUC dari MFCC-CNN, STFT-CNN dan <i>Prosodic</i> -CNN	69
Gambar 23 Confusion matrix dari proses MFCC- CNN	71
Gambar 24 Nilai AUC Proses MFCC-CNN	71
Gambar 25 Architectur model CNN yang di kembangkan.....	74
Gambar 26 Count dataset.....	81
Gambar 27 Struktur desain sistem deteksi kebohongan	86
Gambar 28 Landmark wajah.....	87
Gambar 29. Nilai landmark setiap fitur.....	87
Gambar 30. Penggalan source untuk ekstraksi fitur.....	88
Gambar 31 Proses perhitungan perbedaan ukuran pupil	89
Gambar 32 Contoh Hasil dari ekstraksi setiap fitur.....	90
Gambar 33 Arsitektur Base CNN	92
Gambar 34 Arsitektur Base CNN Spatial attention layer	92
Gambar 35 Arsitektur Base CNN Multihead attention layer	93
Gambar 36 Ilustrasi Model arsitektur attention	93
Gambar 37 Performa berdasarkan precision	100
Gambar 38 Desain arsitekturnya	108
Gambar 39 Grafik AUC dari model MFCC dan LSTM dengan epoch 200	113
Gambar 40 Wave plot data lie.....	113



Gambar 41 Spectrogram data lie.....	114
Gambar 42 Wave plot data truth	114
Gambar 43 Wave plot data truth	114
Gambar 44 Spectrogram data truth.....	115
Gambar 45 Total dataset publik yang digunakan.....	122
Gambar 46 Arsitektur umum model deep CNN berbasis <i>transfer learning</i> untuk deteksi kebohongan.....	123
Gambar 47 Arsitektur VGG-16.....	124
Gambar 48 Arsitektur Resnet-50	125
Gambar 49 Arsitektur MobileNet	126
Gambar 50 Gambar ROC Komparasi dari beberapa arsitektur model	133
Gambar 51 Blok diagram model fusion untuk deteksi kebohongan melalui fitur wajah dan suara	140
Gambar 52 Ilustrasi hasil augmentasi	151
Gambar 53 Ilustrasi hasil resize	151
Gambar 54 Nilai pixel RGB dataset wajah lie	152
Gambar 55 Kernel 3x3	153
Gambar 56 Ilustrasi perhitungan layer posisi 1 dan 2 (konvolusi <i>channel Red</i>)	153
Gambar 57 Ilustrasi perhitungan layer posisi 3 sampai 10 (konvolusi <i>channel Red</i>)	154
Gambar 58 Ilustrasi perhitungan layer posisi 11 sampai 16 (konvolusi <i>channel Red</i>)	155
Gambar 59 Hasil tahapan perhitungan konvolusi <i>channel red</i>	155
Gambar 60 Hasil perhitungan layer konvolusi <i>channel red, green, blue</i>	156
Gambar 61 Total Perhitungan <i>Convolutional</i>	156
Gambar 62 Hasil <i>Convolutional</i> 1.....	156
Gambar 63 Hasil <i>convolutional</i>	157
Gambar 64 Gambar Ilustrasi perhitungan layer (<i>maxpooling</i>)	157
Gambar 65 Ilustrasi proses <i>faltten</i>	158
Gambar 66 Ilustrasi proses <i>dense + softmax</i>	158
Gambar 67 Normalisasi audio waveform untuk suara <i>lie</i>	162
Gambar 68 Normalisasi audio waveform untuk suara <i>truth</i>	163
Gambar 69 Normalisasi audio spektogram untuk suara lie	163
Gambar 70 Normalisasi audio spektogram untuk suara truth	163
Gambar 71 Pre emphasize suara lie	164
Gambar 72 Pre emphasize suara lie	164
Gambar 73 Windowing dari suara lie	165
Gambar 74 Windowing dari suara truth	165
Gambar 75 Ilustrasi visualisasi fourier transform suara lie	166
Gambar 76 Ilustrasi visualisasi fourier transform suara truth	167
Gambar 77 Contoh mel spektogram dari suara lie.....	168
Gambar 78 Contoh mel spektogram dari suara truth	168
Gambar 79 Source ekstraksi fitur suara	169
Gambar 80 Citra dalam bentuk BGR	173
Gambar 81 Keypoint dari wajah	173
Gambar 82 Detect keypoint 468 landmark wajah.....	173
Gambar 83 Bentuk pupil	177
Gambar 84 Posisi setiap fitur	177



Gambar 85 Ilustrasi salah satu video lie	178
Gambar 86 Dataset yang melalui proses ekstraksi fitur.....	178
Gambar 87 Performa dengan split data 80:20	180
Gambar 88 Performa dengan split data 60:40	180
Gambar 89 Performa dengan split data 50:40	180
Gambar 90 Blok diagram teknik <i>fusion</i> deteksi kebohongan melalui fitur wajah dan suara.....	182
Gambar 91 Konversi citra dari BGR ke RGB.....	184
Gambar 92 time stretching audio	185
Gambar 93 Ilustrasi Landmark mata.....	185
Gambar 94 Nilai yang di dapatkan dari ekstraksi fitur pupil.....	187
Gambar 95 Nilai yang di dapatkan dari esktraksi suara.....	188
Gambar 96 Performa model CNN dengan penambahan multihead attention.....	189
Gambar 97 Performa model LSTM	189
Gambar 98 Performa model CNN-multihead attention	190
Gambar 99 Performa model MFCC-LSTM.....	191
Gambar 100 Hasil performa dari mean fusion kedua model.	191
Gambar 101 <i>Use case diagram</i> dari desain sistem deteksi kebohongan	194
Gambar 102 Sequence diagram	195
Gambar 103 GUI dari sistem Deteksi kebohongan.....	195
Gambar 104 layout dari analays system	195
Gambar 105 Video hasil analys	196

DAFTAR LAMPIRAN

Lampiran 1 Source code model untuk data image.....	210
Lampiran 2 Source code model untuk data numeric.....	211
Lampiran 3 source code data suara.....	212
Lampiran 4 Source code gabungan tiga model.....	214



BAB I

PENDAHULUAN

1.1 Latar belakang

Deteksi kebohongan adalah salah satu cara atau mekanisme yang digunakan untuk mengevaluasi nilai-nilai etika dan moral seseorang. Tindakan berbohong mengacu pada menyembunyikan atau menyimpan informasi yang sangat penting. Itu terjadi pada saat seseorang diinterogasi tentang suatu peristiwa atau kejadian. Pentingnya untuk mendeteksi kebohongan ini jika informasi yang disembunyikan oleh orang tersebut sangat penting. Kebohongan dapat dideteksi dari perubahan biologis yang tampak pada diri seseorang seperti konduktivitas kulit, detak jantung, tekanan darah, pernapasan dan denyut nadi. Jenis metode deteksi kebohongan ini sekarang dianggap sebagai sistem deteksi kebohongan tradisional yang biasa disebut sebagai tes poligraf. Jika orang tersebut cukup pandai maka akan mampu memanipulasi poligraf tersebut, maka dari itu metode poligraf salah satu deteksi kebohongan yang tingkat akurasi kurang dan hampir tidak akurat sehingga tidak dibolehkan dalam persidangan. (Cakmak and Zeki, 2016). Deteksi kebohongan dapat dilakukan dengan menggunakan indeks psikologis, karena memperlihatkan perilaku mental seseorang. Dewasa ini banyak peneliti tertarik melakukan penelitian tentang bagaimana untuk mendeteksi kebohongan secara akurat.

Dalam beberapa tahun terakhir deteksi kebohongan dan aplikasinya mendapatkan banyak perhatian khususnya bidang keamanan, dikarenakan dapat membantu mengungkapkan kejahatan dan merupakan suatu sarana dan teknik dalam penyidikan yang memiliki kegunaan untuk mengetahui keterangan yang disampaikan oleh saksi maupun tersangka yang merupakan kejadian yang sebenar benarnya atau berbohong. Analisa tentang kebohongan merupakan hal yang sangat penting dan bermanfaat untuk semua bidang dan implementasinya secara riil digunakan dalam *screening* pegawai baru, menganalisa calon nasabah bank, bahkan terutama bagi pihak penegak hukum, kepolisian, kejaksaan dan kehakiman.

pada saat ini berbagai kasus yang menyangkut masalah pidana, tindak korupsi, teroris yang menjadi isu penting, dan membutuhkan instrumen



untuk penyelidikan kasus. Dari sisi hukum, penggunaan alat deteksi kebohongan dalam proses penyelidikan dan penyidikan dibenarkan (Gunadi and Hartati, 2013). Para peneliti sebelumnya banyak yang tertarik untuk melakukan riset tentang bagaimana mendeteksi kebohongan secara akurat dan *realtime*. Penelitian yang mayoritas dilakukan oleh beberapa peneliti yaitu pada aspek non verbal seperti pencitraan otak, tatapan mata, ekspresi mikro wajah, gestur tubuh serta poligraf, dan pada aspek verbal (analisa suara), untuk aspek verbal beberapa diantaranya sudah menggunakan algoritma otomatis dan dipasarkan secara komersial seperti VSA (*Voice stress analyzer*) dan LVA (*Layered Voice Analysis*). Namun baik VSA dan LVA telah mengalami uji validitas dan reliabilitas, hasilnya tidak cukup handal dan kurang memuaskan. Selain itu deteksi kebohongan berdasarkan fitur fonetik akustik diantaranya intonasi, dan *pitch* dengan menganalisa secara statistik (Kusnadi, Widyantara and Linawati, 2021). Terdapat perbedaan keberhasilan dari penelitian tersebut yaitu 63.66% dan 33.33%. Salah satu metode deteksi kebohongan non verbal yaitu dengan pencitraan otak, menggunakan gelombang infrared yang dipancarkan langsung ke otak biasa disebut dengan *functional near-infrared sensor* (FNIR) dimana metode ini dilakukan dengan cara memasang alat di kepala untuk memantau jumlah oksigen dalam darah di berbagai bagian otak untuk menentukan kapan subjek berbohong, dimana alat ini tidak tersedia untuk umum (Owayjan, Kashour, Al Haddad, *et al.*, 2012). Deteksi kebohongan nonverbal lainnya melalui mikro ekspresi wajah, dimana tingkat akurasi yang didapatkan 85%. Penelitian deteksi kebohongan melalui *image processing* dengan pola kedipan mata menggunakan algoritma HAAR *cascade* untuk mendeteksi (Singh, Rajiv and Chandra, 2015), di mana tingkat kedipan peserta diukur dalam periode target (periode deteksi kebohongan) dan dibandingkan dengan nilai ambang yang menunjukkan tingkat kedipan normal seseorang. Identifikasi kebenaran melalui sinyal *Electroencephalography* (EGG) (Baghel *et al.*, 2020), dimana menggunakan CNN dengan melakukan *cross validation* untuk melihat *performance* dari metode tersebut, dengan membandingkan dataset yang digunakan maksimal tingkat akurasi



ipatkan 84%.

ional neural network (CNN) merupakan salah satu metode *Deep Learning* yang dapat digunakan untuk mendeteksi dan mengenali sebuah objek pada

sebuah citra digital. *Deep learning* merupakan salah satu sub bidang dari *machine learning*. Pada dasarnya *deep learning* adalah implementasi konsep dasar dari *machine learning* yang menerapkan algoritma ANN dengan lapisan yang lebih banyak. Banyaknya lapisan tersembunyi yang digunakan antara lapisan masukan dan keluaran, maka jaringan ini dapat dikatakan *deep neural network*. *Deep learning* telah menunjukkan performa yang luar biasa. Hal ini Sebagian besar dipengaruhi faktor komputasi yang lebih kuat, dataset yang besar dan teknik untuk melatih jaringan yang lebih dalam. Kemampuan CNN diklaim sebagai model terbaik untuk memecahkan permasalahan *object detection* dan *object recognition*. Dari beberapa penelitian yang dilakukan dengan pengolahan citra digital menghasilkan akurasi yang menyaingi manusia pada dataset tertentu. Namun dalam CNN seperti model *deep learning* lainnya, memiliki kelemahan yaitu proses pelatihan model yang cukup lama. Tetapi dengan perkembangan *hardware* yang semakin pesat, hal tersebut dapat diatasi menggunakan teknologi *Graphical Processing Unit* (GPU) dan PC yang memiliki spesifikasi tinggi. Selain CNN penelitian ini juga akan menggunakan LSTM dimana bagian dari *Recurrent Neural Network* (RNN). LSTM mampu menyimpan memori atau ingatan (*feedback loop*) yang memungkinkan untuk mengenali pola data dengan baik. Pengenalan ucapan memungkinkan suatu perangkat lunak untuk mengenali dan memahami kata-kata yang diucapkan oleh manusia. Dari beberapa penelitian yang dikemukakan sebelumnya, metode LSTM mampu memberikan hasil yang baik untuk metode pengenalan ucapan.(Hamzah,2022)

Berdasarkan latar belakang yang diuraikan sebelumnya maka dari itu penulis menggunakan kombinasi antara fitur wajah dan fitur suara dalam mendeteksi kebohongan. Disini, penulis mencoba sebuah metode *deep learning* untuk mengoptimalkan akurasi deteksi kebohongan, dimana untuk wajah dengan melihat perubahan pergerakan setiap bagian wajah misalnya mata, dahi serta mulut, menggunakan metode *convolutional neural networks*(CNN) sementara untuk analisis suara menggunakan metode LSTM dimana *Mel-Frequency Cepstral*

nt (MFCC) sebagai metode ekstraksi ciri dan LSTM sebagai pengenalnya. dapat meniru tingkah laku pendengaran manusia. MFCC mempunyai an dapat menganalisis kelompok suara berdasarkan pola sinyal suara



masuk, memberikan fitur dengan ukuran data yang kecil berupa vektor dengan tidak meniadakan karakteristik dari suara asli, dan mengadaptasi organ pendengaran manusia dalam melakukan persepsi terhadap sinyal suara. Sinyal suara yang diekstraksi menggunakan MFCC akan menjadi sebuah fitur vektor (Yohannes and Wijaya, 2021).

1.2 Rumusan masalah

Berdasarkan uraian dari latar belakang, maka adapun rumusan masalah yang diusulkan dalam penelitian ini secara umum yaitu:

1. Bagaimana memodelkan deteksi kebohongan melalui proses *deep learning* berdasarkan data video?.
2. Bagaimana kemampuan library mediapipe dan CNN dalam mencari fitur wajah yang memiliki performa yang baik diantara beberapa fitur dalam mendeteksi kebohongan?.
3. Bagaimana melakukan pengenalan fitur pitch suara dari rekaman audio sehingga mampu mendeteksi?.
4. Bagaimana melakukan fusion antara wajah dan suara sehingga mampu menghasilkan performa yang baik dalam mendeteksi kebohongan?

1.3 Tujuan penelitian

Tujuan penelitian yang diusulkan secara umum yaitu:

1. Menghasilkan model deteksi kebohongan dengan proses *deep learning*.
2. Mengevaluasi akurasi penggunaan library MediaPipe dan arsitektur Convolutional Neural Networks (CNN) dalam mengidentifikasi dan memproses fitur wajah yang krusial untuk deteksi kebohongan..
3. Mengevaluasi keakuratan dari metode yang dikembangkan dalam mengidentifikasi pitch suara, dengan menguji berbagai parameter dan kondisi akustik untuk deteksi kebohongan.
4. Mengoptimalkan algoritma fusion untuk mengevaluasi hasil dari dua modalitas berbeda visual dan audio dapat memberikan akurasi yang baik untuk deteksi kebohongan.



1.4 Manfaat penelitian

Adapun manfaat yang di dapatkan dari penelitian ini adalah:

1. Menambah wawasan pada bidang penelitian tentang deteksi kebohongan melalui analisa wajah dan suara menggunakan *deep learning*.
2. Berkontribusi pada bidang penyelidikan untuk mengetahui apakah tersangka berkata jujur atau berbohong melalui model yang dikembangkan.
3. Membantu penyidik dalam mendapatkan alat bukti dan untuk membantu menguatkan BAP yang telah dibuat oleh penyidik karena keterangan tersangka yang selalu berubah-ubah.
4. Dapat digunakan sebagai langkah awal untuk membantu tim penyidik dalam melakukan investigasi.
5. Sebagai bahan evaluasi terhadap berbagai pendekatan lainnya dalam mendeteksi kebohongan.

1.5 Batasan masalah penelitian

Agar pembahasan disertasi lebih fokus maka adapun batasan masalah dalam disertasi ini sebagai berikut:

1. Data penelitian utama berasal dari rekaman video yang diambil selama proses investigasi. Dalam skenario yang dibuat, kamera video biasa digunakan untuk mengambil rekaman video. Batasan ini mengasumsikan bahwa kualitas dan fitur rekaman video sesuai dengan standar kamera video
2. Proses deteksi wajah dilakukan pada rekaman video di mana tidak ada objek yang menghalangi wajah subjek, sehingga fokus analisis implementasi metode lebih pada peningkatan akurasi deteksi kebohongan.
3. Selain itu, deteksi suara fokus pada situasi normal tanpa gangguan atau perubahan latar belakang suara yang signifikan.

1.6 Ruang lingkup penelitian



Berdasarkan hasil studi literatur yang dilakukan terkait dengan berbagai penelitian fokus pada deteksi kebohongan menggunakan pemanfaatan teknologi dan informasi, maka dapat disimpulkan bahwa saat ini berbagai pendekatan telah dilakukan dalam upaya mendukung deteksi kebohongan secara akurat terhadap

manusia. Dari berbagai pendekatan tersebut, metode dengan penggunaan fitur wajah atau suara banyak diusulkan untuk deteksi kebohongan. Selain metode dengan menggunakan fitur wajah atau suara tersebut, berbagai penelitian terbaru juga berusaha mengeksplorasi berbagai pendekatan lain yang dapat digunakan dalam deteksi kebohongan melalui EEG. Penelitian terkait deteksi kebohongan melalui suara dan wajah saat ini masih dalam tahap awal sehingga masih memiliki ruang yang luas untuk di eksplorasi lebih jauh terutama dengan ketersediaan dukungan dari metode kecerdasan buatan seperti *machine learning* dan *deep learning*.

Ruang lingkup dari penelitian ini adalah pemanfaatan berbagai fitur wajah dan suara dalam mendeteksi kebohongan pada manusia. Penggunaan suara dan wajah sebagai parameter utama dalam deteksi kebohongan memiliki banyak kelebihan dibandingkan pendekatan lain, diantaranya adalah suara dan wajah dapat dengan mudah diakuisisi dan bersifat non invasif dalam proses akuisisinya sehingga potensi intervensi terhadap subjek juga akan minimal dan tidak berpengaruh terhadap kondisi yang mungkin dialami oleh subjek. Metode ini sangat sesuai untuk diaplikasikan dalam bidang salah satunya keamanan. Seperti yang telah diuraikan sebelumnya bahwa deteksi kebohongan merupakan suatu permasalahan yang kompleks dan membutuhkan penanganan khusus karena memberikan keterangan yang tidak benar sangatlah berakibat fatal. Kemampuan mereka menyembunyikan kebenaran tidak bisa disembunyikan dengan ekspresi wajah ataupun intonasi suara, sehingga tidak jarang pihak-pihak tertentu menggunakan bantuan psikologi untuk mendeteksi apakah subjek tersebut berbohong atau tidak. Berdasarkan berbagai masalah dan juga potensi yang dihadapi, maka penelitian ini berfokus dalam mengeksplorasi metode deteksi kebohongan dengan memanfaatkan metode *deep learning* dengan fokus pada fitur wajah dan fitur suara.

Novelty atau keterbaruan yang diharapkan dalam penelitian ini, sehingga memberikan kontribusi bagi ilmu pengetahuan dan teknologi khususnya terkait pemanfaatan teknologi dalam kecerdasan buatan untuk mendukung bidang

1 khususnya, sebagai berikut:

eteksi kebohongan melalui wajah dengan poin landmark pada wajah,
 enggunaan poin landmark wajah yang spesifik dan terdefinisi dengan baik,



sistem deteksi kebohongan dapat meningkatkan sensitifitas terhadap perubahan ekspresi wajah yang halus, yang mungkin menunjukkan ketidakjujuran. Ini termasuk perubahan kecil pada mata khususnya area pupil.

2. Mengetahui pola dari karakteristik suara secara sensitif terhadap perubahan kecil dalam intonasi, pitch, dan dinamika suara yang sering terkait pada saat subjek berbohong dan pada saat jujur dengan penambahan fitur zero crossing rate, energy entropy dan audio energy.
3. Menggabungkan hasil dari setiap model dari analisa wajah dan suara untuk mendapatkan model yang optimal untuk deteksi kebohongan.

