

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Pustaka

Tinjauan ini mencakup pengenalan topik secara umum, alasan mengapa topik tersebut penting, gambaran penelitian sebelumnya dan kesenjangan yang ada, serta struktur yang akan diikuti dalam tinjauan literatur. Hal ini memberikan kerangka kerja yang jelas dan terfokus bagi pembaca untuk memahami konteks dan tujuan tinjauan literatur.

Menurut (Dong et al., 2023) yang melakakukan penelitian tentang *Contrastive learning-based general Deepfake detection with multi-scale RGB frequency clues Multi-Scale RGB Frequency Clues*. Khususnya, pada kumpulan data Celeb-DF yang menunjukkan domain target, metode yang diteliti menunjukkan peningkatan metrik ACC sebesar 5,3% dan metrik AUC sebesar 12,8%, serta penurunan metrik ERR sebesar 7,2 persen dibandingkan dengan metode canggih.

Menurut (Warke et al., 2023), melakukan penelitian tentang *DEEPEFAKE DETECTION THROUGH DEEP LEARNING USING RESNEXT CNN AND LSTM*. Peneliti menawarkan alat baru yang mengukur kepercayaan yang luas terhadap prediksi model dengan mengklasifikasikan video menjadi *deepfakes* atau asli dengan menggunakan arsitektur jaringan saraf. Pendekatan ini memanfaatkan efisiensinya, karena mencapai hasil yang akurat hanya dengan menganalisis satu detik rekaman video, yang setara dengan kecepatan bingkai 10 bingkai per detik. Peneliti

memanfaatkan kekuatan ResNext CNN yang telah dilatih sebelumnya untuk membuat model ini, yang memiliki kemampuan yang lebih baik untuk mengekstraksi fitur detail pada tingkat frame. Mengidentifikasi perubahan temporal antara frame yang berurutan dengan menggabungkan komponen LSTM yang melakukan analisis sekvensial, yang dirancang untuk menangani rangkaian video dengan berbagai panjang frame, seperti 10, 20, 40, 60, 80, dan 100 bingkai. Dengan mempertimbangkan berbagai urutan frame untuk melayani berbagai konteks video dan meningkatkan kemampuan untuk mendeteksi deepfake.

Menurut penelitian (Yadav et al., 2021) melakukan penelitian mengenai *DeepFake Detection using InceptionResNetV2 and LSTM* model yang dirancang diuji selama 20 epoch dan 40 epoch, dan mencapai akurasi 84,75% dan 91,48%, masing-masing. Grafik yang dihasilkan setelah implementasi mengklaim kebenaran bahwa akurasi validasi dan pengujian meningkat seiring dengan jumlah *epoch*. *Confusion matrix* yang dibuat membantu dalam menentukan keakuratan pengujian sistem.

Menurut (Abidin et al., 2022), yang melakukan penelitian mengenai *Deepfake detection in videos using Long Short-Term Memory and CNN ResNext*, dengan hasil klasifikasi menggunakan *ResNext CNN-LSTM* menunjukkan tingkat akurasi tertinggi sebesar 90%, presisi sebesar 100%, dan *recall* sebesar 97% untuk data sebanyak 60 *frame*. Di sisi lain, data dari 10 *frame* video memiliki tingkat akurasi sebesar 52%, presisi sebesar 52%, dan *recall* sebesar 50%. Selain itu, *ResNext CNN-LSTM* memiliki kinerja yang baik pada data video dengan 20 frame.

Menurut (Chang et al., 2020), melakukan penelitian tentang *DeepFake Face Image Detection based on Improved VGG Convolutional Neural Network*. Hasil dari percobaan menunjukkan bahwa filter SRM memiliki kemampuan untuk meningkatkan noise gambar sebesar 16,8% dibandingkan dengan jaringan VGG16 dan peningkatan gambar sebesar 12,5%, yang menunjukkan bahwa filter SRM memiliki kemampuan untuk menonjolkan fitur *noise* gambar dan peningkatan gambar yang efektif untuk meningkatkan akurasi deteksi.

Menurut (Dinata et al., 2023), yang mengambil penelitian tentang IMPLEMENTASI DEEP LEARNING DALAM KLASIFIKASI CITRA GAMBAR DENGAN MENGGUNAKAN METODE CNN. Dari hasil penelitian ini tingkat akurasi prediksi yang paling tertinggi yaitu untuk sepatu sneakers dengan tingkat akurasi prediksi sebesar 99% sementara untuk tingkat akurasi prediksi terendah yaitu untuk sepatu Brogue sebesar 27%. Data gambar sepatu dengan model CNN arsitektur *Inception V-3* berhasil untuk dibuat dengan hasil akurasi sebesar 68.75% dan untuk validasi *loss* sebesar 27.73%.

Menurut (Saikia et al., 2022), yang membahas penelitian tentang *A Hybrid CNN-LS*. Model hibrid bekerja dengan baik pada kumpulan data sumber terbuka seperti DFDC, FF++, dan *Celeb-DF*. Hasilnya menunjukkan akurasi sebesar 66,26%, 91,21%, dan 79,49% untuk DFDC, FF++, dan Celeb-DF, masing-masing, dengan jumlah sampel yang sangat rendah (*frame*).

Menurut (Thing, 2023), yang membahas tentang *Deepfake Detection with Deep Learning: Convolutional Neural Networks versus Transformers*. Hasil mendapatkan

tingkat akurasi 88,74%, 99,53%, 97,68%, 92,02%, dan AUC 99,95%, 100%, 99,88%, 99,99%, dan 97,61% untuk pendekripsi deepfake FF++ 2020, Google DFD, CelebDF, dan Deeper Forensics, dan menemukan bahwa model Transformers lebih baik dalam menilai lintas kumpulan data daripada model CNN dalam evaluasi kumpulan data pelatihan-ke-pengujian yang sama.

Menurut (Jiwtoode et al., 2022), melakukan penelitian tentang *Deepfake Video Detection Using Recurrent Neural Networks* mereka menggunakan sub-urutan dengan panjang $N = 20, 40$, dan 80 frame untuk menunjukkan akurasi deteksi sistem. Urutan frame ini diekstraksi dari setiap video secara berurutan, tanpa lompatan frame. Seluruh pipeline dilatih secara *end-to-end* sampai mencapai 10 epoch tertinggi dalam set validasi. Hasil menunjukkan bahwa sistem dapat memprediksi dengan akurasi lebih dari 97 persen apakah fragmen yang dianalisis berasal dari video deepfake dengan video berdurasi kurang dari 2 detik (40 frame untuk video yang diambil sampelnya pada 24 frame per detik).

Menurut (Afchar et al., 2018), yang meneliti tentang *MesoNet: a Compact Facial Video Forgery Detection Network*. Eksperimen menunjukkan bahwa metode ini memiliki tingkat deteksi rata-rata 98% untuk video Deepfake dan 95% untuk video Face2Face.

Menurut (Xia et al., 2022) dalam penelitian mereka mengemukakan bahwa penambahan modul prapemrosesan pada MesoNet secara signifikan meningkatkan deteksi video deepfake. Dengan menggabungkan fitur spatio-temporal, pendekatan ini berhasil meningkatkan akurasi deteksi dan mengurangi tingkat kesalahan.

2.1.1 State Of The Art Penelitian

Tabel 2.1 Menunjukkan perbandingan penelitian sebelumnya yang berhubungan dengan focus penelitian deteksi *deepfake*. Terdapat beberapa kesamaan serta perbedaan dari masing-masing penelitian. Hal ini dapat dilihat dari penggunaan metode serta algoritmanya.

Tabel 2.1 *State of The Art* Penelitian Terkait

No	Peneliti / Tahun	Judul	Metode	Hasil penelitian
1	(Dong et al., 2023)	<i>Contrastive learning-based general Deepfake detection with multi-scale RGB frequency clues</i>	<i>Multi-Scale RGB Frequency Clues</i>	Hasil menunjukkan bahwa, dibandingkan dengan metode deteksi lainnya, model mereka lebih dapat digeneralisasikan. Khususnya, pada kumpulan data Celeb-DF yang menunjukkan domain target, metode mereka menunjukkan peningkatan metrik ACC sebesar 5,3% dan metrik AUC sebesar 12,8%, serta penurunan

No	Peneliti / Tahun	Judul	Metode	Hasil penelitian
				metrik ERR sebesar 7,2 persen dibandingkan dengan metode canggih.
2	(Warke et al., 2023)	<i>DEEPCODETION THROUGH DEEP LEARNING USING RESNEXT CNN AND LSTM</i>	<i>RESNEXT CNN AND LSTM</i>	Mereka menawarkan alat baru yang mengukur kepercayaan yang luas terhadap prediksi model dengan mengklasifikasikan video menjadi deepfakes atau asli dengan menggunakan arsitektur jaringan saraf. Pendekatan ini memanfaatkan efisiensinya, karena mencapai hasil yang akurat hanya dengan menganalisis satu detik rekaman

No	Peneliti / Tahun	Judul	Metode	Hasil penelitian
				video, yang setara dengan kecepatan bingkai 10 bingkai per detik. Mereka memanfaatkan kekuatan ResNext CNN yang telah dilatih sebelumnya untuk membuat model ini, yang memiliki kemampuan yang lebih baik untuk mengekstraksi fitur detail pada tingkat frame. Mereka juga memungkinkan identifikasi perubahan temporal antara frame yang berurutan dengan menggabungkan komponen LSTM yang melakukan analisis sekuensial.

No	Peneliti / Tahun	Judul	Metode	Hasil penelitian
				Khususnya, dicancang untuk menangani rangkaian video dengan berbagai panjang <i>frame</i> , seperti 10, 20, 40, 60, 80, dan 100 bingkai. Dengan mempertimbangkan berbagai urutan <i>frame</i> , solusi mereka melayani berbagai konteks video dan meningkatkan kemampuan untuk mendeteksi deepfake.
3	(Yadav et al., 2021)	<i>DeepFake Detection using InceptionResNetV2 and LSTM</i>	<i>InceptionResNetV2 and LSTM</i>	Karena keterbatasan waktu pengoperasian, model yang dirancang diuji selama 20 epoch dan 40 epoch, dan mencapai akurasi 84,75% dan

No	Peneliti / Tahun	Judul	Metode	Hasil penelitian
				91,48%, masing-masing. Grafik yang dihasilkan setelah implementasi mengklaim kebenaran bahwa akurasi validasi dan pengujian meningkat seiring dengan jumlah epoch. Matriks kebingungan yang dibuat membantu dalam menentukan keakuratan pengujian sistem.
4	(Abidin et al., 2022)	<i>Deepfake detection in videos using Long Short-Term Memory and CNN ResNext</i>	<i>Long Short-Term Memory and CNN ResNext</i>	Hasil klasifikasi menggunakan ResNext CNN-LSTM menunjukkan tingkat akurasi tertinggi sebesar 90%, presisi sebesar 100%, dan recall sebesar 97% untuk data sebanyak 60

No	Peneliti / Tahun	Judul	Metode	Hasil penelitian
				frame. Di sisi lain, data dari 10 frame video memiliki tingkat akurasi sebesar 52%, presisi sebesar 52%, dan recall sebesar 50%. Selain itu, ResNext CNN-LSTM memiliki kinerja yang baik pada data video dengan 20 frame.
5	(Chang et al., 2020)	<i>DeepFake Face Image Detection based on Improved VGG Convolutional Neural Network</i>	<i>VGG Convolutional Neural Network</i>	Hasil dari percobaan menunjukkan bahwa filter SRM memiliki kemampuan untuk meningkatkan noise gambar sebesar 16,8% dibandingkan dengan jaringan VGG16 dan peningkatan gambar

No	Peneliti / Tahun	Judul	Metode	Hasil penelitian
				sebesar 12,5%, yang menunjukkan bahwa filter SRM memiliki kemampuan untuk menonjolkan fitur <i>noise</i> gambar dan peningkatan gambar yang efektif untuk meningkatkan akurasi deteksi.
6	(Dinata et al., 2023)	IMPLEMENTASI <i>DEEP LEARNING</i> DALAM KLASIFIKASI CITRA GAMBAR DENGAN MENGGUNAKAN METODE CNN	CNN	Dari hasil penilitian ini tingkat akurasi prediksi yang paling tertinggi yaitu untuk sepatu <i>sneakers</i> dengan tingkat akurasi prediksi sebesar 99% sementara untuk tingkat akurasi prediksi terendah yaitu untuk sepatu <i>Brogue</i> sebesar 27%. Data gambar

No	Peneliti / Tahun	Judul	Metode	Hasil penelitian
				sepatu dengan model CNN arsitektur <i>Inception V-3</i> berhasil untuk dibuat dengan hasil akurasi sebesar 68.75% dan untuk validasi <i>loss</i> sebesar 27.73%.
7	(Saikia et al., 2022)	<i>A Hybrid CNN-LSTM model for Video Deepfake Detection by Leveraging Optical Flow Features</i>	<i>CNN-LSTM</i>	Model hibrid bekerja dengan baik pada kumpulan data sumber terbuka seperti DFDC, FF++, dan Celeb-DF. Hasilnya menunjukkan akurasi sebesar 66,26%, 91,21%, dan 79,49% untuk DFDC, FF++, dan Celeb-DF, masing-masing, dengan jumlah sampel yang sangat rendah (<i>frame</i>).

No	Peneliti / Tahun	Judul	Metode	Hasil penelitian
8	(Thing, 2023)	<i>Deepfake Detection with Deep Learning: Convolutional Neural Networks versus Transformers</i>	<i>Convolutional Neural Networks versus Transformers</i>	Hasil mendapatkan tingkat akurasi 88,74%, 99,53%, 97,68%, 92,02%, dan AUC 99,95%, 100%, 99,88%, 99,99%, dan 97,61% untuk pendekripsi deepfake FF++ 2020, Google DFD, Celeb-DF, dan Deeper Forensics, dan menemukan bahwa model Transformers lebih baik dalam menilai lintas kumpulan data daripada model CNN dalam evaluasi kumpulan data pelatihan-ke-pengujian yang sama.

No	Peneliti / Tahun	Judul	Metode	Hasil penelitian
9	(Jiwtoode et al., 2022)	<i>Deepfake Video Detection Using Recurrent Neural Networks</i>	<i>Recurrent Neural Networks (RNN)</i>	Mereka menggunakan sub-urutan dengan panjang $N = 20, 40,$ dan 80 frame untuk menunjukkan akurasi deteksi sistem. Urutan frame ini diekstraksi dari setiap video secara berurutan, tanpa lompatan frame. Seluruh pipeline dilatih secara end-to-end sampai mencapai 10 epoch tertinggi dalam set validasi. Hasil menunjukkan bahwa sistem dapat memprediksi dengan akurasi lebih dari 97 persen apakah fragmen yang dianalisis berasal dari video <i>deepfake</i>

No	Peneliti / Tahun	Judul	Metode	Hasil penelitian
				dengan <i>video</i> berdurasi kurang dari 2 detik (40 <i>frame</i> untuk <i>video</i> yang diambil sampelnya pada 24 <i>frame</i> per detik).
10	(Afchar et al., 2018)	<i>MesoNet: a Compact Facial Video Forgery Detection Network</i>	<i>MesoNet</i>	Eksperimen menunjukkan bahwa metode ini memiliki tingkat deteksi rata-rata 98% untuk <i>video Deepfake</i> dan 95% untuk <i>video Face2Face</i> .
11	(Xia et al., 2022)	<i>Deepfake Video Detection Based on MesoNet with Preprocessing Module</i>	MesoNet dan <i>Preprocessing module</i>	Kinerja deteksi metode yang diusulkan diverifikasi pada dua set data; AUC pada FaceForensics++ dapat mencapai 0,974, dan dapat mencapai 0,943 pada Celeb-DF yang

No	Peneliti / Tahun	Judul	Metode	Hasil penelitian
				lebih baik daripada metode saat ini. Yang lebih penting, bahkan dalam kasus kompresi berat, tingkat deteksi masih dapat lebih dari 88%.

2.1.2 Matriks Penelitian

Matrix penelitian menunjukkan perbedaan dan persamaan antara penelitian yang sudah dilakukan dengan penelitian yang akan dilakukan. Matrix penelitian dapat dilihat dari table berikut

Tabel 2.2 Matriks Penelitian

No	Peneliti / Tahun	Judul	Ruang Lingkup Penelitian											
			Metode					Dataset				Optimizer		
			CNN	LSTM	RNN	MesoNet	Inception ResNetV2	FF++	F2F	Celeb-DF	DFDC	Adam	Nadam	Other method
1.	(Dong et al., 2023)	<i>Contrastive learning-based general Deepfake detection with multi-scale</i>	✓	-	-	-	-	✓	✓	✓	✓	✓	-	-

		<i>RGB frequency clues</i>														
2	(Warke et al., 2023)	<i>DEEPMFAKE DETECTION THROUGH DEEP LEARNING USING RESNEXT CNN AND LSTM</i>	✓	✓	✓	-	-	-	-	-	✓	✓	-	-		
3	(Yadav et al., 2021)	<i>DeepFake Detection using InceptionResNet V2 and LSTM</i>	✓	✓	✓	-	✓	✓	-	✓	-	-	-	-	✓	
4	(Abidin et al., 2022)	<i>Deepfake detection in videos using Long Short-Term Memory and CNN ResNext</i>	✓	✓	✓	-	-	-	-	-	✓	-	-	-	✓	
5	(Chang et al., 2020)	<i>DeepFake Face Image Detection based on Improved VGG Convolutional Neural Network</i>	✓	-	-	-	-	✓	-	✓	-	-	-	-	✓	

6	(Dinata et al., 2023)	IMPLEMENTASI DEEP LEARNING DALAM KLASIFIKASI CITRA GAMBAR DENGAN MENGGUNAKAN METODE CNN	✓	-	-	-	-	-	-	-	-	-	-	✓	-	-	
7	(Saikia et al., 2022)	<i>A Hybrid CNN-LSTM model for Video Deepfake Detection by Leveraging Optical Flow Features</i>	✓	✓	✓	-	-	✓	-	✓	✓	✓	✓	-	-	-	
8	(Thing, 2023)	<i>Deepfake Detection with Deep Learning: Convolutional Neural Networks versus Transformers</i>	✓	-	-	-	-	✓	-	✓	✓	✓	-	-	-	✓	
9	(Jiwttode et al., 2022)	<i>Deepfake Video Detection Using</i>	-	✓	✓	-	-	-	-	-	-	-	✓	-	-		

		<i>Recurrent Neural Networks</i>														
10	(Afchar et al., 2018)	<i>MesoNet: a Compact Facial Video Forgery Detection Network</i>	-	-	-	✓	-	-	✓	-	-	✓	-	-	-	-
11	(Xia et al., 2022)	<i>Deepfake Video Detection Based on MesoNet with Preprocessing Module</i>	-	-	-	✓	-	✓	-	✓	-	-	-	-	-	✓
12	Penelitian yang dilakukan saat ini	STUDI ANALISIS KEASLIAN IMAGE MENGGUNAKAN ARSITEKTUR MESONET	✓	-	-	✓	-	-	-	✓	-	✓	✓	✓	-	-

2.1.3 Relevansi Penelitian

Tabel 2. 3 Relevansi Perbandingan

Peneliti	(Afchar et al., 2018)	Penelitian yang dilakukan saat ini
Judul	<i>MesoNet: a Compact Facial Video Forgery Detection Network</i>	STUDI ANALISIS DAN KOMPARASI METODE OPTIMASI ADAM DAN NADAM MENGGUNAKAN ARSITEKTUR MESONET
Masalah Penelitian	Masalah utama yang dihadapi adalah deteksi <i>deepfake</i> atau pemalsuan video wajah dengan menggunakan pendekatan jaringan neural yang kompak dan efisien	Bagaimana tingkat akurasi hasil perbandingan optimizer adam dan nadam dalam mendekripsi <i>deepfake</i>
Objek Penelitian	Deteksi Deepfake	Deteksi Deepfake
Algoritma	MesoNet	MesoNet
Optimasi	Adam	Adam dan Nadam

2.1.4 Kebaruan Penelitian

Berdasarkan penelitian sebelumnya yang telah diuraikan, penelitian ini menyumbangkan kontribusi baru dengan membandingkan langsung antara dua metode optimasi yang umum digunakan dalam deteksi deepfake menggunakan arsitektur MesoNet, yaitu optimisasi Adam dan Nadam. Penelitian sebelumnya sering kali fokus pada implementasi arsitektur atau teknik evaluasi, namun penelitian ini menghadirkan perspektif yang lebih mendalam dengan menganalisis perbandingan kinerja antara kedua optimisasi.

2.2 Landasan Teori

2.2.1 Deepfake

Deepfake adalah metode yang memungkinkan seseorang untuk mengganti wajah orang lain ke dalam gambar atau *video*. Meskipun teknik pemalsuan bukan hal baru, *deepfake* menggunakan pembelajaran mesin dan kecerdasan buatan untuk memanipulasi dan menghasilkan konten visual dan *audio* yang digunakan untuk menyebarkan informasi palsu (Pishori et al., 2020).

2.2.2 CNN

Algoritma *deep learning*, *Convolutional Neural Network* (CNN/ConvNet), dikembangkan dari Multilayer Perceptron (MLP) dan dimaksudkan untuk mengolah data dua dimensi, seperti suara atau gambar. CNN digunakan untuk mengklasifikasi data yang terlabel dengan metode supervised learning. Metode ini bekerja karena ada data yang dilatih dan variabel yang ditargetkan, sehingga tujuan dari metode ini adalah

untuk mengelompokkan data ke dalam data sebelumnya (Ilahiyah S & Nilogiri A, 2018).

Convolutional Neural Networks (CNN) adalah kelas jaringan saraf dalam, yang paling umum diterapkan untuk menganalisis citra visual. Mereka terinspirasi oleh sistem visual manusia dan telah terbukti sangat efektif dalam tugas-tugas yang melibatkan pengenalan dan klasifikasi gambar, deteksi objek, dan jenis pemrosesan data visual lainnya.

Adapun beberapa fungsi CNN, yaitu :

1) Operasi inti dalam CNN

Operasi inti dalam CNN adalah konvolusi, yang melibatkan penerapan serangkaian filter (kernel) pada gambar masukan untuk menghasilkan peta fitur. Filter ini membantu mendeteksi fitur seperti tepi, tekstur, dan pola di berbagai lokasi spasial dalam gambar. Tujuannya adalah untuk mengekstrak fitur tingkat tinggi dari citra masukan, yang kemudian digunakan untuk pemrosesan dan klasifikasi lebih lanjut.

2) *Pooling layers*

Pooling layers lapisan pengumpulan, seperti pengumpulan maksimum dan pengumpulan rata-rata, digunakan untuk mengurangi dimensi spasial peta fitur, sehingga mengurangi jumlah parameter dan beban komputasi dalam jaringan. Untuk mencapai invariansi spasial dengan merangkum keberadaan fitur-fitur di patch peta fitur, yang membantu membuat jaringan lebih kuat terhadap variasi pada gambar masukan. Fungsi aktivasi non-linier seperti ReLU (*Rectified Linear*

Unit) diterapkan pada peta fitur untuk memperkenalkan non-linier ke dalam jaringan, sehingga memungkinkannya mempelajari pola dan representasi yang kompleks. Tujuannya untuk jaringan memodelkan hubungan yang kompleks dan membuat prediksi yang lebih baik.

3) *After several convolutional and pooling layers*

After several convolutional and pooling layers, setelah beberapa lapisan konvolusional dan penggabungan, jaringan biasanya mencakup satu atau lebih lapisan yang terhubung sepenuhnya (padat). Lapisan ini mengambil fitur tingkat tinggi yang diekstraksi oleh lapisan konvolusional dan menggabungkannya untuk membuat prediksi akhir.

4) *Dropout*

Dropout adalah teknik regularisasi di mana sebagian kecil neuron disetel secara acak ke nol selama pelatihan. Ini membantu mencegah *overfitting*. Tujuannya untuk meningkatkan kemampuan generalisasi jaringan dengan mengurangi *overfitting*.

Tujuan CNN adalah untuk image *classification, object detection* :

- 1) *Image Classification*, CNN banyak digunakan dalam tugas klasifikasi gambar yang tujuannya adalah untuk mengkategorikan gambar ke dalam salah satu dari beberapa kelas yang telah ditentukan sebelumnya. Contohnya seperti mengklasifikasikan gambar binatang ke dalam kategori seperti kucing, anjing, burung, dll.

2) *Object Detection*, dalam deteksi objek, CNN digunakan untuk mengidentifikasi dan menemukan lokasi objek dalam gambar, menggambar kotak pembatas di sekitar objek yang terdeteksi. Contoh: Mendeteksi mobil, pejalan kaki, dan rambu lalu lintas pada aplikasi mengemudi otomotif.

2.2.3 MesoNet

MesoNet adalah arsitektur jaringan saraf yang dirancang khusus untuk deteksi deepfake, yang melibatkan identifikasi media yang dimanipulasi atau sintetis. Seiring dengan semakin maju dan meluasnya teknologi *deepfake*, kemampuan untuk mendeteksi pemalsuan ini menjadi sangat penting dalam menjaga integritas konten digital. MesoNet berfokus pada sifat mesoskopik (fitur tingkat menengah) gambar untuk mengidentifikasi inkonsistensi halus yang muncul selama pembuatan *deepfake*.

MesoNet terdiri dari beberapa lapisan konvolusi yang diikuti oleh lapisan *pooling* dan *fully connected*. Terdapat dua varian utama dari MesoNet:

1. Meso-4.

Arsitektur dasar yang terdiri dari empat lapisan konvolusi (*convolutional layers*) diikuti oleh lapisan sepenuhnya terhubung (*fully connected layers*). Model ini dirancang untuk menangkap fitur-fitur halus pada gambar yang dapat menunjukkan adanya manipulasi.

2. MesoInception-4

Arsitektur yang lebih kompleks yang menggabungkan blok *Inception* untuk menangkap fitur dari berbagai skala. Blok *Inception* memungkinkan model ini untuk

memproses informasi dari berbagai ukuran filter secara simultan, sehingga lebih efektif dalam mendeteksi variasi manipulasi pada gambar.

Menurut penelitian MesoNet adalah jenis CNN yang dirancang khusus untuk mendeteksi deepfake dalam konten multimedia. MesoNet terdiri dari jaringan saraf dengan sedikit lapisan saraf dan terutama berkonsentrasi pada fitur mesoskopik gambar, yang mencakup fitur tingkat rendah dan tingkat tinggi (Afchar et al., 2018).

2.2.4 Deep Learning

Deep learning adalah salah satu jenis pembelajaran mesin (*machine learning*) yang menggunakan algoritma pemodelan abstraksi data yang sangat kompleks, yang memiliki fungsi transformasi non-linear yang mendalam yang dapat mencapai ratusan lapisan. Pembangunan dan pelatihan model jaringan saraf tiruan adalah inti dari *deep learning*. Jaringan saraf tiruan terdiri dari banyak lapisan, biasanya lebih dari tiga, dan setiap lapisan memiliki neuron atau unit pemrosesan yang memiliki kemampuan untuk berinteraksi satu sama lain. Model membuat prediksi dan keputusan setelah menerima data dari jaringan dan mengekstrak elemen penting dari data melalui berbagai lapisan (Kumar & L., 2018).

2.2.5 Python

Python adalah bahasa pemrograman tujuan umum yang digunakan untuk membangun situs *web*, *software*, atau aplikasi, mengotomatiskan tugas, dan melakukan analisis data. Artinya, python dapat digunakan untuk membuat berbagai program, bukan hanya untuk menyelesaikan masalah tertentu.

Bahasa pemrograman Python merupakan bahasa pemrograman populer yang memiliki keunggulan sebagai berikut (Kadarina & Ibnu Fajar, 2019) :

1. Mudah untuk digunakan dalam mengembangkan sebuah produk perangkat lunak, perangkat keras, *Internet of Things*, aplikasi *web*, maupun *video game*.
2. Selain memiliki keterbacaan kode yang tinggi, sehingga kode mudah dipahami, bahasa pemrograman ini memiliki *library* yang sangat banyak dan luas.
3. Merupakan bahasa yang mendukung ekosistem *Internet of Things* dengan sangat baik.

2.2.6 TensorFlow

TensorFlow, *framework open-source* milik Google, digunakan untuk mengembangkan dan melatih model *machine learning* dan *deep learning* yang sudah ada, serta proyek terkait analisis statistik lainnya. Banyak *library* dan *Application Program Interface* (API) saat ini telah membantu kemajuan bidang *Deep Learning*. Tensorflow adalah *library* yang digunakan untuk mengekspresikan algoritma pembelajaran mesin dan mengeksekusi perintah dengan menggunakan informasi tentang objek atau target yang dikenal. Ini juga memiliki kemampuan untuk membedakan objek satu sama lain. Tensorflow dapat menggunakan CPU dan GPU untuk menjalankan pelatihan model (Nurfita & Ariyanto, 2018).

2.2.7 Optimizer

Optimizer adalah algoritma yang digunakan dalam pelatihan model pembelajaran mesin untuk memperbarui bobot model dengan tujuan meminimalkan fungsi *loss*, yang mengukur seberapa baik model melakukan prediksi. Dalam proses

pelatihan, *optimizer* mengatur perubahan bobot berdasarkan gradien dari fungsi *loss* yang dihitung melalui *backpropagation*. Salah satu *optimizer* yang sangat populer adalah Adam (*Adaptive Moment Estimation*) (PRATIWI et al., 2021).

Adam menggabungkan keuntungan dari Adagrad dan RMSprop dengan menggunakan rata-rata momentum gradien pertama (*mean*) dan kedua (*variance*) untuk menyesuaikan learning rate secara adaptif untuk setiap parameter. Adam mengelola dua momenta, yaitu *mean* dan *variance* dari *gradien*, yang memungkinkan konvergensi lebih cepat dan stabil dibandingkan dengan metode lain seperti *Stochastic Gradient Descent* (SGD). Adam memiliki parameter utama seperti beta1 dan beta2 yang mengontrol eksponensial *moving average* dari gradien dan kuadrat gradien, sehingga memberikan keseimbangan antara kecepatan konvergensi dan stabilitas (Asy Syifa & Amelia Dewi, 2022).

NAdam (*Nesterov-accelerated Adaptive Moment Estimation*) adalah varian dari Adam yang mengintegrasikan Nesterov momentum. Teknik Nesterov momentum memperkirakan posisi parameter di masa depan sebelum menghitung gradien, yang membantu mengurangi osilasi dan meningkatkan stabilitas serta kecepatan konvergensi. NAdam menghitung gradien berdasarkan posisi parameter yang diproyeksikan, bukan posisi saat ini, sehingga seringkali memberikan kinerja yang lebih baik dibandingkan Adam dalam berbagai situasi (PRATIWI et al., 2021).

2.2.8 *Confusion Matrix*

Confusion Matrix adalah alat pengukuran kinerja untuk masalah klasifikasi pembelajaran mesin. Hal ini sangat berguna untuk mengevaluasi performa model klasifikasi, memberikan wawasan tidak hanya tentang akurasi tetapi juga jenis kesalahan yang dibuat oleh model (Qadrini et al., 2021). Matriks konfusi membantu memahami kekuatan dan kelemahan model dengan memberikan wawasan tentang jenis kesalahan tertentu yang dibuatnya. Contoh: Jumlah negatif palsu yang tinggi mungkin menunjukkan bahwa model tersebut terlalu konservatif dalam memprediksi kelas positif. Adapun struktur *confusion matrix* biasanya berupa matriks persegi yang membandingkan nilai target aktual dengan nilai prediksi. Ini disusun sebagai berikut untuk masalah klasifikasi biner:

<i>Actual \ Predicted</i>	<i>Positive (Predicted)</i>	<i>Negative (Predicted)</i>
<i>Positive (Actual)</i>	<i>True Positive (TP)</i>	<i>False Negative (FN)</i>
<i>Negative (Actual)</i>	<i>False Positive (FP)</i>	<i>True Negative (TN)</i>

Komponen Matriks Kebingungan

1) *True Positive (TP)*:

Kasus ketika model memprediksi kelas positif dengan tepat. Contoh: Seorang model dengan tepat mengidentifikasi gambar seekor anjing sebagai seekor anjing.

2) *True Negative (TN)*:

Kasus ketika model memprediksi kelas negatif dengan tepat. Contoh: Seorang model dengan tepat mengidentifikasi gambar kucing sebagai bukan anjing.

3) *False Negative* (FP):

Kasus ketika model salah memprediksi kelas positif. Contoh: Model salah mengidentifikasi gambar kucing sebagai anjing (kesalahan Tipe I).

4) *False Negative* (FN):

Kasus ketika model salah memprediksi kelas negatif. Contoh: Model salah mengidentifikasi gambar anjing sebagai bukan anjing (kesalahan Tipe II).

Confusion Matrix memiliki beberapa fungsi, diantaranya yaitu :

- 1) Untuk evaluasi kinerja, *confusion matrix* memberikan perincian mendetail tentang performa model, yang menunjukkan berapa banyak instance yang diklasifikasikan dengan benar atau salah. Tujuannya untuk memahami keakuratan dan keandalan model klasifikasi.
- 2) Analisis kesalahan, dengan memeriksa positif palsu dan negatif palsu, seseorang dapat mengidentifikasi area tertentu di mana model tersebut melakukan kesalahan. Tujuannya untuk menyempurnakan model dengan mengatasi jenis kesalahan yang rentan terjadi.
- 3) Perhitungan Metrik, beberapa metrik kinerja penting dapat diperoleh dari matriks konfusi, termasuk:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$\text{Presisi} = \frac{TP}{TP+FP} \times 100\%$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\%$$

$$F-1\ score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \times 100\%$$

2.2.9 Google Colab

Google Colaboratory, juga dikenal sebagai *Google Colab*, adalah dokumen yang dapat dijalankan yang memungkinkan pengguna menyimpan, menulis, dan membagikan *program* yang telah mereka tulis melalui *Google Drive*. Untuk menggunakan alat ini, hanya perlu login dengan akun *Google* dan menyimpan *file Python* ke *Google Drive* (Carneiro et al., 2018).