

PERBANDINGAN PERFORMA MODEL *CONVOLUTIONAL NEURAL NETWORK* ARSITEKTUR CONVNEXT, MOBILENET V3, DAN EFFICIENTNET PADA PENGLASIFIKASIAN PENYAKIT MULUT

Skripsi

**Oleh
FERLI MALKAN AMIEN
2117051050**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2025**

PERBANDINGAN PERFORMA MODEL *CONVOLUTIONAL NEURAL NETWORK* ARSITEKTUR CONVNEXT, MOBILENET V3, DAN EFFICIENTNET PADA PENGKLASIFIKASIAN PENYAKIT MULUT

Oleh

FERLI MALKAN AMIEN

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar

SARJANA KOMPUTER

Pada

Jurusan Ilmu Komputer

Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Lampung



FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS LAMPUNG

BANDAR LAMPUNG

2025

ABSTRAK

PERBANDINGAN PERFORMA MODEL *CONVOLUTIONAL NEURAL NETWORK* ARSITEKTUR CONVNEXT, MOBILENET V3, DAN EFFICIENTNET PADA PENGKLASIFIKASIAN PENYAKIT MULUT

Oleh

FERLI MALKAN AMIEN

Penyakit mulut merupakan masalah kesehatan global yang memengaruhi miliaran orang di dunia. Deteksi dini berbasis teknologi kecerdasan buatan, khususnya dengan Convolutional Neural Network (CNN), dapat meningkatkan akurasi diagnosis penyakit mulut. Penelitian ini membandingkan performa tiga arsitektur CNN, yaitu ConvNeXt Tiny, MobileNet V3 Large, dan EfficientNet B0, dalam klasifikasi penyakit mulut berdasarkan dataset citra medis. Evaluasi dilakukan dengan mempertimbangkan akurasi, efisiensi komputasi, dan waktu pemrosesan. Hasil penelitian menunjukkan bahwa EfficientNet B0 mencapai akurasi tertinggi (98,21%) dengan jumlah parameter lebih kecil dibandingkan ConvNeXt Tiny. Namun, ConvNeXt Tiny memiliki performa terbaik dalam akurasi absolut (96,96%), meskipun dengan konsumsi sumber daya lebih besar. MobileNet V3 Large menunjukkan efisiensi tinggi dalam penggunaan parameter dengan akurasi kompetitif (96,44%). Dengan demikian, pemilihan arsitektur CNN bergantung pada kebutuhan antara akurasi tinggi atau efisiensi sumber daya komputasi dalam implementasi klinis.

Kata kunci: Convolutional Neural Network (CNN), ConvNeXt, MobileNetV3, EfficientNet, Penyakit Mulut

ABSTRACT

PERFORMANCE COMPARISON OF CONVOLUTIONAL NEURAL NETWORK ARCHITECTURES CONVNEXT, MOBILENET V3, AND EFFICIENTNET FOR ORAL DISEASE CLASSIFICATION

By

FERLI MALKAN AMIEN

Oral diseases are a global health issue affecting billions of people worldwide. Early detection using artificial intelligence technology, particularly Convolutional Neural Networks (CNN), can enhance the *accuracy* of oral disease diagnosis. This study compares the performance of three CNN architectures—ConvNeXt Tiny, MobileNet V3 Large, and EfficientNet B0—in classifying oral diseases based on a medical image dataset. The evaluation considers *accuracy*, computational efficiency, and processing time. The results show that EfficientNet B0 achieved the highest *accuracy* (98.21%) with fewer parameters than ConvNeXt Tiny. However, ConvNeXt Tiny demonstrated the best absolute *accuracy* (96.96%) despite higher computational resource consumption. MobileNet V3 Large exhibited high efficiency in parameter usage with competitive *accuracy* (96.44%). Thus, the choice of CNN architecture depends on the trade-off between high *accuracy* and computational efficiency in clinical implementation.

Keywords: Convolutional Neural Network (CNN), ConvNeXt, MobileNetV3, EfficientNet, Oral Disease

Judul Skripsi

**: PERBANDINGAN PERFORMA MODEL
CONVOLUTIONAL NEURAL NETWORK
ARSITEKTUR CONVNEXT, MOBILENET
V3, DAN EFFICIENTNET PADA
PENGKLASIFIKASIAN PENYAKIT MULUT**

Nama Mahasiswa

: Ferli Malkan Amien

Nomor Pokok Mahasiswa

: 2117051050

Program Studi

: S1 Ilmu Komputer

Jurusan

: Ilmu Komputer

Fakultas

: Matematika dan Ilmu Pengetahuan Alam



MENYETUJUI

1. Komisi Pembimbing

Didik Kurniawan, S.Si., M.T.

NIP 19800419 200501 1 004

2. Ketua Jurusan Ilmu Komputer

Dwi Sakethi, S.Si., M.Kom.

NIP 19680611 199802 1 001

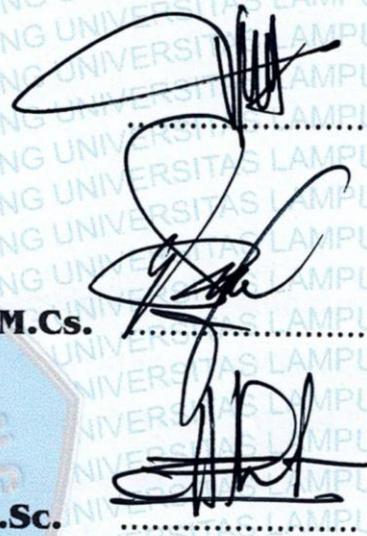
MENGESAHKAN

1. Tim Penguji

Ketua : Didik Kurniawan, S.Si., M.T.

Sekretaris : Bambang Hermanto, S.Kom., M.Cs.

Penguji Utama : Dr. rer. nat. Akmal Junaidi, M.Sc.



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Heri Satria, S.Si., M.Si.
NIP 19711001 200501 1 002

Tanggal Lulus Ujian Skripsi : 28 April 2025

PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Ferli Malkan Amien

NPM : 2117051050

Dengan ini menyatakan bahwa skripsi saya yang berjudul “**Perbandingan Performa Model *Convolutional Neural Network* Arsitektur ConvNeXt, MobileNet V3, dan EfficientNet Pada Pengklasifikasian Penyakit Mulut**” merupakan karya saya sendiri dan bukan karya orang lain. Semua tulisan yang tertuang di skripsi ini telah mengikuti kaidah penulisan karya tulis ilmiah Universitas Lampung. Apabila di kemudian hari terbukti skripsi saya merupakan hasil jiplakan atau dibuat orang lain, maka saya bersedia menerima sanksi berupa pencabutan gelar yang saya terima.

Bandar Lampung, 28 April 2025

Penulis,



Ferli Malkan Amien

NPM. 2117051050

RIWAYAT HIDUP



Lahir di Liwa pada tanggal 28 September 2003 sebagai anak pertama dari pasangan Bapak Feri Shaputra, S.T. dan Ibu Lilis Nurhayati, S.Pd.SD. Penulis telah menyelesaikan Pendidikan formal di SD N 3 Liwa pada tahun 2015. Kemudian SMP N Sekuting Terpadu pada tahun 2018 dan SMAN 1 Liwa pada tahun 2021. Penulis terdaftar sebagai mahasiswa program studi S1 Ilmu Komputer di Universitas Lampung melalui jalur tes atau SBMPTN.

Selama menjadi mahasiswa di Ilmu Komputer, penulis aktif di dalam berbagai kegiatan baik di dalam maupun di luar Universitas Lampung. Kegiatan yang dilakukan adalah sebagai berikut,

1. Menjadi Koordinator Asisten Dosen untuk mata kuliah Dasar – Dasar Pemrograman di Jurusan Ilmu Komputer pada tahun 2022.
2. Menjadi Koordinator Asisten Dosen untuk mata kuliah Sistem Informasi di Jurusan Ilmu Komputer pada tahun 2023.
3. Menjadi Asisten Dosen untuk mata kuliah Pemrograman Interpreter di Jurusan Ilmu Komputer pada tahun 2023.
4. Menjadi Asisten Dosen untuk mata kuliah Sistem Pakar di Jurusan Ilmu Komputer pada tahun 2024.
5. Menjadi Anggota Bidang Keilmuan Himpunan Mahasiswa Jurusan Ilmu Komputer pada tahun 2022.
6. Menjadi Ketua Himpunan Mahasiswa Jurusan Ilmu Komputer pada tahun 2023.
7. Mengikuti Magang di SDGs Center Universitas Lampung pada tahun 2022 sebagai *Graphic Designer*.

8. Mengikuti Magang di PT Bank CIMB Niaga Tbk. pada tahun 2024 sebagai *Data Governance Staff*.
9. Menjadi Peserta Kegiatan Latihan Keterampilan Manajemen Mahasiswa – Tingkat Menengah dan Lanjut yang diselenggarakan Kemendikbudristek pada tahun 2023.
10. Menjadi Delegasi untuk Indonesia pada kegiatan *South East Asia-Australia Youth Leaders Week 2021, Asia Pacific Youth SDGs Summit 2022, dan 3rd Global Youth Conference on SDGs 2022*.
11. Menjadi *volunteer* untuk 30+ kegiatan di berbagai bidang seperti pendidikan, kesehatan, teknologi, lingkungan hidup, dll.
12. Menjadi Pemateri pada kegiatan Sosialisasi Kampus Merdeka bertema “Membuka Gerbang Peluang: Menggali Potensi Diri dengan Kampus Merdeka” pada tahun 2024.
13. Menjadi Pemateri pada kegiatan Upgrading Pengurus Himakom FMIPA Unila Periode 2024.
14. Menjadi Pemateri pada kegiatan LKMMIK-TP Himakom FMIPA Unila Periode 2024.
15. Menjadi Pemateri pada kegiatan LKMM-TD Himakom FMIPA Unila Periode 2025.

PERSEMBAHAN

Segala puji saya panjatkan kepada Allah Subhanahu wa Ta'ala atas segala rahmat-Nya. Shalawat dan salam semoga senantiasa tercurah kepada Nabi Muhammad Shallallahu 'Alaihi wa Sallam, sehingga saya dapat menyelesaikan skripsi ini.

Skripsi ini saya persembahkan kepada:

Kedua orang tua saya dan seluruh rekan yang telah kebersamai serta membantu dalam proses penyelesaian skripsi ini. Juga untuk diri saya sendiri, atas keberhasilan menuntaskan pendidikan di Program Studi S1 Ilmu Komputer, Universitas Lampung.

MOTTO

“Faber est suae quisque fortunae”

-Appius Claudius Caecus

SANCAWACANA

Puji syukur kehadiran Allah Subhanahu wa Ta'ala, yang telah memberikan rahmat dan hidayat-Nya kepada penulis, sehingga penulis dapat menyelesaikan skripsi ini tepat waktu. Shalawat serta salam penulis sanjungan kepada Nabi Muhammad Shallallahu 'alaihi wa sallam yang penulis harapkan syafaatnya di hari akhir kelak.

Skripsi yang berjudul "**Perbandingan Performa Model *Convolutional Neural Network* Arsitektur ConvNeXt, MobileNet V3, dan EfficientNet pada Pengklasifikasian Penyakit Mulut**" ini disusun sebagai salah satu syarat untuk memperoleh gelar sarjana pada Program Studi S1 Ilmu Komputer di Universitas Lampung.

Pada kesempatan ini penulis mengucapkan terima kasih kepada semua pihak yang telah berkontribusi dalam penyusunan skripsi ini, yaitu:

1. Kedua orang tua penulis, Ayah Feri Shaputra dan Bunda Lilis Nurhayati, yang tidak henti-hentinya memberikan doa, semangat, serta kepercayaan kepada penulis dalam menyelesaikan skripsi ini. Juga kepada kedua adik penulis, Ashifa Meilia Putri dan Aurellia Arvia Zahra, yang selalu memberikan dukungan dan doa.
2. Bapak Dwi Sakethi, S.Si., M.Kom., selaku Ketua Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung sekaligus Dosen Pembimbing Akademik penulis.
3. Bapak Didik Kurniawan, S.Si., M.T. selaku pembimbing utama yang telah membimbing dan memberikan banyak arahan kepada penulis sehingga penelitian yang penulis lakukan dapat menjadi lebih baik.

4. Bapak Dr. rer. nat. Akmal Junaidi, M.Sc. selaku pembahas satu yang telah memberikan kritik dan saran yang bermanfaat yang sangat membantu sehingga penulisan skripsi ini dapat terselesaikan.
5. Bapak Bambang Hermanto, S.Kom., M.Cs. sebagai pembahas dua yang telah memberikan masukan dan saran yang bermanfaat bagi penulis dalam menyusun serta menyelesaikan skripsi.
6. Bapak/Ibu Dosen dan seluruh Staf Jurusan Ilmu Komputer FMIPA Universitas Lampung yang telah memberikan ilmu dan membantu dalam berbagai urusan akademik maupun administratif.
7. Nur Farisa Aqila yang telah memberikan bantuan, saran, dan dukungan, terutama terkait sumber daya dalam bidang kedokteran gigi yang mendukung penyusunan skripsi ini.
8. Seluruh teman-teman Ilmu Komputer 2021 yang terlibat dalam proses penelitian skripsi ini yang tidak dapat penulis sebutkan satu persatu.
9. Untuk diri penulis, Ferli Malkan Amien, terima kasih telah bertahan dan berjuang dalam setiap proses, hingga berhasil menyelesaikan pendidikan di Program Studi S1 Ilmu Komputer tepat waktu.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna, oleh karena itu kritik dan saran yang membangun sangat penulis harapkan demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi semua pihak.

Bandar Lampung, 28 April 2025

Penulis,

Ferli Malkan Amien

NPM. 2117051050

DAFTAR ISI

	Halaman
DAFTAR ISI	iii
DAFTAR TABEL	v
DAFTAR GAMBAR	vi
I. PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah	5
1.4. Tujuan	5
1.5. Manfaat	6
II. TINJAUAN PUSTAKA	7
2.1. Penelitian Terdahulu	7
2.2. Penyakit Mulut	10
2.3. Convolutional Neural Network (CNN)	14
2.4. ConvNeXt	18
2.5. MobileNet	20
2.6. EfficientNet	24
2.7. Augmentasi Data	27
2.8. Hyperparameter	28
2.9. Overfitting	31
2.10. Confusion matrix	31
2.11. Framework OSEMN	34

III. METODOLOGI PENELITIAN.....	37
3.1. Waktu dan Tempat.....	37
3.2. Perangkat Penelitian.....	38
3.3. Dataset.....	38
3.4. Tahap Penelitian.....	39
V. SIMPULAN DAN SARAN.....	48
5.1. Simpulan.....	48
5.2. Saran.....	49
DAFTAR PUSTAKA.....	50

DAFTAR TABEL

Tabel	Halaman
1. Penelitian terdahulu.....	7
2. Spesifikasi detail arsitektur ConvNeXt Tiny.....	19
3. Spesifikasi MobileNetV3-Large	22
4. Spesifikasi MobileNetV3-Small	22
5. EfficientNet-B0 baseline network.....	25
6. Confusion matrix (Maxwell et al., 2021).....	32
7. Rencana penelitian	37
8. Spesifikasi perangkat keras	38
9. Hyperparameter training model	45

DAFTAR GAMBAR

Gambar	Halaman
1. Kanker mulut (Rashid et al., 2023).	11
2. Oral potentially malignant disorders (Piyarathne et al., 2024).	13
3. Ilustrasi convolutional neural network (MathWorks, 2021).	14
4. Ilustrasi convolutional layer.	16
5. Ilustrasi pooling layer.	16
6. Ilustrasi sederhana fully-connected layer (NVIDIA Corporation & affiliates, 2023).	17
7. Desain block ConvNeXt (Liu et al., 2022).	19
8. Ilustrasi overfitting.	31
9. Framework OSEMN (Data Science PM, 2024).	34
10. Komposisi dataset	39
11. Tahap penelitian.	40
12. Diagram alir proses <i>training</i> dan evaluasi.	46

I. PENDAHULUAN

1.1. Latar Belakang

Penyakit mulut, atau *oral diseases*, merupakan masalah kesehatan global yang memengaruhi lebih dari 3,5 miliar orang di seluruh dunia pada tahun 2019, termasuk karies gigi, penyakit periodontal, kehilangan gigi, kanker rongga mulut, trauma gigi, noma, dan kelainan bawaan seperti celah langit-langit dan bibir sumbing (Institute for Health Metrics and Evaluation (IHME), 2020; World Health Organization, 2022). Menurut laporan WHO yang berjudul “*Global Oral Health Status Report: Towards Universal Health Coverage for Oral Health by 2030*”, penyakit mulut seperti karies gigi dan penyakit periodontal adalah masalah utama yang sering dijumpai, terutama di negara berpenghasilan rendah dan menengah yang menghadapi keterbatasan akses layanan kesehatan gigi. Dampak dari penyakit ini sangat signifikan, mencakup gangguan berbicara, kesulitan mengunyah, peningkatan risiko infeksi, serta potensi berkembang menjadi kondisi yang lebih serius, seperti kanker mulut. Selain itu, WHO menekankan bahwa penyakit mulut terkait dengan faktor risiko penyakit tidak menular lainnya, seperti pola makan buruk, merokok, dan konsumsi alkohol, sehingga penting untuk mengembangkan strategi kesehatan yang terpadu guna mencegah penyakit ini secara efektif.

Kesadaran akan dampak besar dari penyakit mulut dan keterbatasan layanan kesehatan yang ada menggarisbawahi kebutuhan mendesak untuk mengadopsi pendekatan yang lebih maju dalam penanganan masalah ini. Banyak strategi kesehatan tradisional belum mampu mengatasi tantangan tersebut secara optimal, sehingga memerlukan solusi inovatif. Penerapan teknologi berbasis AI dan perkembangan pesat dalam *deep learning* menawarkan peluang besar dengan menganalisis data medis secara cepat dan akurat, mempercepat deteksi dini, serta

meningkatkan kualitas perawatan secara global (Fawaz et al., 2023). Teknologi ini memungkinkan pemrosesan data medis dalam jumlah besar yang sulit dilakukan dengan metode tradisional, mendukung dokter dalam membuat keputusan klinis yang lebih baik dan berbasis data. Selain mempercepat proses diagnosis dan perawatan, AI juga memungkinkan pengembangan perawatan kesehatan yang dipersonalisasi, berpotensi menghasilkan hasil yang lebih baik dan akurat (Vodanović et al., 2023).

Kemajuan pesat dalam *deep learning* telah memberikan kontribusi besar dalam dunia kesehatan, terutama dalam hal deteksi dini dan prediksi penyakit yang lebih akurat (Kasnanda Bintang & Imaduddin, 2024). Salah satu penerapan *deep learning* yang paling menonjol adalah pada bidang *computer vision*, khususnya dalam pengenalan pola dan klasifikasi gambar medis. Model-model *deep learning* seperti *convolutional neural networks* (CNN) telah menjadi standar emas dalam tugas-tugas ini karena kemampuannya yang secara efisien dalam mengekstraksi fitur penting dari data gambar (Kasnanda Bintang & Imaduddin, 2024). CNN dapat belajar secara otomatis dari data berlabel, menjadikannya alat yang sangat efektif untuk mengenali gambar medis dan mendiagnosis penyakit dengan akurasi yang tinggi. Penggunaan CNN dalam diagnosis berbasis gambar telah menunjukkan hasil yang menjanjikan dalam berbagai aplikasi medis, mulai dari deteksi tumor hingga analisis retina.

Convolutional Neural Network (CNN) adalah salah satu jenis arsitektur jaringan saraf tiruan yang dirancang khusus untuk mengenali serta pembuatan gambar (Ding et al., 2023). CNN terdiri dari beberapa lapisan konvolusi yang bertugas mengekstraksi fitur dari input gambar, diikuti oleh lapisan pooling untuk mengurangi dimensi data tanpa kehilangan informasi penting. Kombinasi lapisan ini memungkinkan CNN untuk belajar dari fitur rendah hingga fitur tinggi, seperti tepi objek hingga pola yang lebih kompleks. Kekuatan CNN terletak pada kemampuannya untuk menangani variasi dalam gambar dan mengidentifikasi fitur penting yang dapat membedakan antara kategori yang berbeda, membuatnya sangat berguna untuk tugas klasifikasi (Carrillo-Perez et al., 2022).

Di tengah berbagai arsitektur CNN yang telah dikembangkan, muncul kebutuhan untuk mengidentifikasi model yang tidak hanya akurat, tetapi juga efisien dalam penggunaan sumber daya komputasi, seperti memori dan kecepatan proses. Hal ini penting terutama dalam aplikasi medis yang memerlukan real-time processing dengan sumber daya terbatas. Oleh karena itu, penelitian ini akan membandingkan tiga arsitektur CNN yang terkenal dengan efisiensinya, yaitu ConvNeXt, MobileNet V3, dan EfficientNet. Setiap arsitektur memiliki pendekatan unik untuk mengoptimalkan penggunaan sumber daya sambil menjaga kualitas hasil klasifikasi.

ConvNeXt adalah varian yang dirancang dengan mengoptimalkan proses konvolusi melalui pendekatan modern seperti depthwise convolution dan normalisasi lapisan yang lebih baik (Liu et al., 2022). MobileNet V3, di sisi lain, memfokuskan pada penggunaan efisiensi melalui blok inverted residual dan squeeze-and-excitation yang mengurangi beban komputasi tanpa mengorbankan performa (Howard et al., 2019). Sementara itu, EfficientNet memanfaatkan scaling yang optimal antara depth, width, dan resolution untuk mencapai keseimbangan antara akurasi dan efisiensi (Tan & Le, 2019).

Dataset yang digunakan dalam penelitian ini bersumber dari riset yang dilakukan oleh (Piyarathne et al., 2024). yang dipublikasikan pada artikel yang berjudul "*A comprehensive dataset of annotated oral cavity images for diagnosis of oral cancer and oral potentially malignant disorders*" yang mana dataset tersebut dapat diakses di *platform Zenodo data repository*. Dataset ini dipilih karena kualitas dan keberagamannya yang mencakup berbagai kondisi di mulut, memungkinkan model untuk belajar dan menggeneralisasi dengan baik terhadap variasi gambar yang mungkin ditemui dalam praktik klinis. Dengan menggunakan dataset yang kaya dan terstruktur ini, penelitian ini bertujuan untuk mengeksplorasi seberapa baik model ConvNeXt, MobileNet V3, dan EfficientNet dapat mengklasifikasikan berbagai penyakit gigi berdasarkan citra digital, dengan fokus pada efisiensi penggunaan sumber daya komputasi dan tetap menjaga tingkat akurasi yang tinggi.

Deteksi penyakit pada mulut menggunakan teknologi berbasis gambar telah banyak diteliti dalam beberapa tahun terakhir. Beberapa penelitian sebelumnya

menunjukkan bahwa model CNN dapat mencapai akurasi tinggi dalam mendeteksi kondisi seperti karies dan kalkulus dengan menggunakan foto intraoral (Oroh & Lubis, 2023). Penelitian ini berupaya untuk mengevaluasi efektivitas dari arsitektur CNN yang lebih efisien, dengan harapan dapat menemukan solusi yang optimal untuk diterapkan dalam skenario klinis yang nyata, terutama yang memerlukan pemrosesan cepat dengan perangkat keras yang terbatas.

Berdasarkan pemaparan tersebut, karya tulis ilmiah berjudul “**Perbandingan Performa Model *Convolutional Neural Network* Arsitektur ConvNeXt, MobileNet V3, dan EfficientNet pada Pengklasifikasian Penyakit Mulut**” ini akan membahas pengembangan dan evaluasi model CNN yang menggunakan arsitektur-arsitektur tersebut. Penelitian ini bertujuan untuk mengidentifikasi model yang paling efisien dalam hal penggunaan sumber daya, namun tetap memberikan akurasi yang tinggi dalam diagnosis penyakit mulut, yang pada akhirnya dapat berkontribusi pada peningkatan kualitas pelayanan kesehatan gigi secara keseluruhan.

1.2. Rumusan Masalah

Adapun rumusan masalah dari penelitian ini adalah sebagai berikut:

1. Bagaimana mengimplementasikan dan mengoptimalkan model CNN dengan arsitektur ConvNeXt Tiny, MobileNet V3 Large, dan EfficientNet B0 untuk mendeteksi penyakit mulut?
2. Bagaimana perbandingan tingkat akurasi dan efisiensi dari arsitektur ConvNeXt Tiny, MobileNet V3 Large, dan EfficientNet B0 dalam mendeteksi penyakit mulut?
3. Model arsitektur CNN mana yang paling efisien dalam penggunaan sumber daya komputasi namun tetap menjaga tingkat akurasi tinggi untuk klasifikasi penyakit mulut?

1.3. Batasan Masalah

Adapun batasan masalah dari penelitian ini adalah sebagai berikut:

1. Penelitian ini hanya menggunakan dataset gambar mulut dari riset yang dilakukan oleh Piyrathne et al. (2024), yang mungkin tidak mencakup seluruh variasi kondisi mulut yang ada dalam populasi global.
2. Fokus pengklasifikasian terbatas pada empat kondisi mulut yaitu *Healthy*, *Benign*, *OPMD*, dan *OCA*, yang mungkin bukan representasi lengkap dari semua penyakit mulut.
3. Evaluasi performa model hanya mencakup arsitektur ConvNeXt Tiny, MobileNet V3 Large, dan EfficientNet B0, tanpa perbandingan langsung dengan model deep learning lainnya di luar arsitektur CNN ini.
4. Dataset *pretraining* yang digunakan oleh model berasal dari ImageNet-1K, yang terdiri dari 1,28 juta gambar dengan 1000 kelas umum.
5. Penelitian ini tidak mencakup implementasi klinis atau uji coba model dalam lingkungan praktik kesehatan yang nyata, sehingga hasil penelitian lebih bersifat eksploratif dan memerlukan validasi lebih lanjut sebelum dapat diterapkan secara klinis.

1.4. Tujuan

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Mengimplementasikan dan mengevaluasi performa arsitektur ConvNeXt Tiny, MobileNet V3 Large, dan EfficientNet B0 dalam pengklasifikasian penyakit mulut menggunakan dataset yang tersedia.
2. Membandingkan hasil klasifikasi penyakit mulut menggunakan arsitektur ConvNeXt Tiny, MobileNet V3 Large, dan EfficientNet B0, untuk menentukan model yang paling efisien dalam penggunaan sumber daya namun tetap menjaga akurasi tinggi.

1.5. Manfaat

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Memberikan kontribusi pada pengembangan teknologi diagnostik di bidang kedokteran gigi, khususnya dalam hal akurasi dan kecepatan diagnosis dengan memanfaatkan model deep learning yang efisien.
2. Meningkatkan pemahaman tentang penggunaan arsitektur ConvNeXt Tiny, MobileNet V3 Large, dan EfficientNet B0 dalam klasifikasi gambar medis dan potensi penerapannya dalam aplikasi klinis yang nyata.
3. Menyediakan informasi yang bermanfaat bagi peneliti dan praktisi dalam memilih model CNN yang paling sesuai untuk aplikasi medis berbasis gambar yang efisien dan akurat.

II. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian yang pernah dilakukan dan berkaitan dengan pembuatan dan implementasi CNN guna kebutuhan *computer vision* khususnya dalam konteks klasifikasi penyakit mulut dapat dilihat pada Tabel 1.

Tabel 1. Penelitian terdahulu

Penulis	Judul	Metode	Hasil
(Devindi et al., 2024)	<i>Multimodal Deep Convolutional Neural Network Pipeline for AI-Assisted Early Detection of Oral Cancer</i>	<i>Convolutional Neural Network (CNN)</i>	Penelitian tersebut berhasil mendapatkan nilai akurasi tertinggi sebesar 81% menggunakan arsitektur MobileNetV3-Large untuk klasifikasi kanker mulut.
(Jeyaraj & Samuel Nadar, 2019)	<i>Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm</i>	<i>Convolutional Neural Network (CNN)</i>	Penelitian tersebut berhasil mendapatkan nilai akurasi sebesar 91,4% untuk klasifikasi tumor kanker dengan tumor jinak dan 94,5% untuk klasifikasi tumor kanker dengan jaringan normal.

(Karacan & Yucebas, 2022)	<i>A Deep Learning Convolutional Model with Neural Attention Mechanism for Dental Image Segmentation</i>	<i>Convolutional Neural Network (CNN)</i>	Penelitian tersebut berhasil mendapatkan nilai IoU dari arsitektur Vision Transformer (ViT), Segmenter, dan ConvNeXt sebesar 88,67%, 91.86%, dan 90.77% untuk segmentasi gigi dan 77,51%, 96,32%, dan 92,01% untuk segmentasi maxillomandibular.
(Xu et al., 2022)	<i>High-Accuracy Oral Squamous Cell Carcinoma Auxiliary Diagnosis System Based on EfficientNet</i>	<i>Convolutional Neural Network (CNN)</i>	Penelitian tersebut berhasil mendapatkan nilai akurasi terbaik sebesar 93,1% dengan arsitektur EfficientNet.

Penelitian terkait penerapan teknologi computer vision untuk pengklasifikasian penyakit mulut sekaligus juga implementasi arsitektur MobileNetV3 pernah dilakukan oleh (Devindi et al., 2024) dengan judul “*Multimodal Deep Convolutional Neural Network Pipeline for AI-Assisted Early Detection of Oral Cancer*”. Penelitian tersebut menggunakan dataset yang sama dengan yang Penulis gunakan pada penelitian ini yaitu dataset yang dipublikasikan melalui paper yang dilakukan oleh (Piyarathne et al., 2024). Penelitian ini menggunakan 6 arsitektur CNN yaitu DenseNet-121, Inception-v3, HRNet-W18-C, MixNet_s, ResNet50, dan MobileNetV3-Large, dengan hasil terbaik diberikan oleh arsitektur MobileNetV3-Large dengan akurasi sebesar 81%.

Penelitian terkait penerapan teknologi *computer vision* untuk pengklasifikasian penyakit mulut juga pernah dilakukan oleh (Jeyaraj & Samuel Nadar, 2019) dengan judul “*Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm*”. Penelitian ini menggunakan model CNN *custom* yang diusulkan oleh penulis dan kemudian dibandingkan kinerjanya dengan model *support vector machine* (SVM) dan *deep belief networks* (DBN). Data yang digunakan pada penelitian ini berjumlah 100 gambar tumor kanker dengan tumor jinak dan 500 gambar tumor kanker dengan jaringan normal. Hasil yang didapatkan dari pengujian yang dilakukan adalah akurasi sebesar 91,4% untuk klasifikasi tumor kanker dengan tumor jinak dan 94,5% untuk klasifikasi tumor kanker dengan jaringan normal, yang mana lebih baik dari dua metode yang menjadi pembandingnya.

Penelitian terkait implementasi arsitektur ConvNeXt telah dilakukan oleh (Karacan & Yucebas, 2022) dengan judul “*A Deep Learning Model with Attention Mechanism for Dental Image Segmentation*”. Dataset yang digunakan pada penelitian ini didapatkan dari *TUFTS Dental Dataset* yang dipublikasikan oleh (Panetta et al., 2022). Penelitian ini menggunakan tiga arsitektur CNN yaitu Vision Transformer (ViT), Segmenter, dan ConvNeXt. Hasil menunjukkan performa IoU dari masing masing arsitektur sebesar 88,67%, 91.86%, dan 90.77% untuk segmentasi gigi dan 77,51%, 96,32%, dan 92,01% untuk segmentasi maxillomandibular.

Penelitian terkait implementasi arsitektur EfficientNet telah dilakukan oleh (Xu et al., 2022) dengan judul “*High-Accuracy Oral Squamous Cell Carcinoma Auxiliary Diagnosis System Based on EfficientNet*”. Dataset yang digunakan pada penelitian tersebut mencakup dataset gambar TCGA OSCC dan dataset gambar TMA independen yang dimiliki oleh tim penulis. Dataset OSCC berjumlah 757 gambar slide penuh yang didapat dari situs web resmi TCGA pada tahun 2019 sebagai dataset asli TCGA. Dataset TCGA digunakan untuk pelatihan dan pengujian model, sedangkan dataset TMA digunakan untuk verifikasi eksternal. Riset ini menggunakan tiga arsitektur CNN yaitu EfficientNet b0, ShuffleNetV2, dan ResNeXt_18. Hasil menunjukkan bahwa arsitektur EfficientNet menghasilkan performa yang paling baik dengan akurasi sebesar 93,1% pada gambar dengan resolusi 20x.

2.2. Penyakit Mulut

Penyakit mulut adalah istilah yang digunakan untuk menggambarkan beberapa masalah gigi yang lebih umum, mencakup berbagai kondisi klinis kronis yang memengaruhi gigi dan mulut, termasuk karies gigi, penyakit periodontal, dan kanker mulut (Peres et al., 2019). Penyakit ini menyerang area mulut dan jaringan sekitarnya, seperti gusi, bibir, lidah, dan lapisan mukosa. Penyebab penyakit mulut dapat beragam, mulai dari infeksi bakteri, virus, atau jamur, kebiasaan buruk seperti merokok dan konsumsi alkohol, hingga gangguan sistem kekebalan tubuh. Penyakit mulut umumnya menimbulkan rasa nyeri, peradangan, serta kerusakan jaringan yang bisa berdampak pada kesehatan secara keseluruhan jika tidak ditangani dengan baik.

Penyakit mulut adalah masalah kesehatan yang dapat menyerang siapa saja, tanpa memandang usia. Beberapa jenis penyakit mulut, seperti sariawan dan gingivitis, relatif umum terjadi dan bisa sembuh dengan sendirinya atau dengan perawatan sederhana. Namun, ada pula jenis penyakit mulut yang lebih serius, seperti kanker mulut dan lichen planus, yang memerlukan perhatian medis khusus. Penyakit-penyakit ini tidak hanya memengaruhi fungsi makan dan berbicara, tetapi juga bisa menurunkan kualitas hidup penderitanya.

2.2.1. *Oral Cancer (OCA)*

Oral cancer atau kanker mulut merupakan salah satu bentuk kanker yang telah diamati dan dideskripsikan oleh dunia kedokteran sejak zaman kuno (Inchingolo et al., 2020). Penyakit ini telah disebutkan dalam berbagai literatur medis awal oleh dokter dan ahli bedah dari peradaban masa lalu yang menggambarkan neoplasma atau tumor di area mulut. Hingga saat ini, kanker mulut tetap menjadi tantangan besar dalam bidang onkologi, dengan berbagai pendekatan baru yang terus dikembangkan untuk mendiagnosis dan mengobatinya. Kanker mulut mencakup kanker yang terjadi pada bibir, seluruh bagian rongga mulut, serta orofaring (Ferlay et al., 2024). Jenis kanker ini menjadi penyebab utama kematian ke-15 di dunia dan termasuk dalam urutan ke-16 sebagai kanker paling umum secara global.



Gambar 1. Kanker mulut (Rashid et al., 2023).

Tingkat kejadian kanker mulut bervariasi secara signifikan di seluruh dunia, dengan angka kejadian yang disesuaikan dengan usia mencapai empat kasus per 100.000 orang. Variasi ini tergantung pada berbagai faktor, seperti jenis kelamin, kelompok usia, negara, ras, kelompok etnis, serta kondisi sosial-ekonomi (Ferlay et al., 2019, 2024). Perbedaan utama antara negara-negara berkembang dan negara-negara Barat dalam hal prevalensi kanker mulut berkaitan erat dengan kebiasaan populasi, harapan hidup, pendidikan pencegahan, serta kualitas catatan medis di masing-masing negara. Negara berkembang sering kali menghadapi tantangan seperti kemiskinan, buta huruf, presentasi kanker pada stadium lanjut, kurangnya akses terhadap layanan kesehatan, dan infrastruktur pengobatan yang terbatas, yang semuanya berkontribusi pada tingginya angka kejadian dan tingkat kematian akibat kanker mulut.

Sejumlah faktor risiko yang berkaitan dengan kondisi fisik, lingkungan, dan genetika telah diidentifikasi sebagai penyebab utama kanker rongga mulut (Greenspan et al., 2020). Penggunaan tembakau, baik dalam bentuk merokok maupun mengunyah tembakau, merupakan salah satu faktor risiko yang paling signifikan. Konsumsi alkohol yang berlebihan juga diketahui meningkatkan risiko terjadinya kanker mulut, terutama bila dikombinasikan dengan penggunaan tembakau. Selain itu, infeksi virus HPV (*Human Papillomavirus*), khususnya tipe HPV-16, menjadi penyebab signifikan kanker orofaring, yang merupakan bagian dari rongga mulut dan tenggorokan. Paparan sinar ultraviolet yang berlebihan, terutama pada orang-orang yang sering terpapar sinar matahari tanpa perlindungan, juga dapat menyebabkan kanker bibir.

Di sisi lain, faktor-faktor genetik juga memainkan peran penting dalam meningkatkan risiko seseorang terhadap kanker mulut. Beberapa studi menunjukkan adanya predisposisi genetik tertentu yang membuat individu lebih rentan terhadap kanker ini. Faktor sosial-ekonomi seperti pola makan yang buruk, kurangnya asupan vitamin A dan C, serta kebersihan mulut yang tidak terjaga, juga turut menyumbang terhadap perkembangan kanker mulut, terutama di negara-negara dengan akses kesehatan yang terbatas. Kanker mulut sering kali terdeteksi pada stadium lanjut karena gejalanya yang tidak jelas pada tahap awal, seperti luka atau sariawan yang tidak sembuh, rasa sakit yang menetap di area mulut, atau munculnya benjolan yang tidak biasa.

Upaya pencegahan dan pengobatan kanker mulut terus mengalami perkembangan seiring dengan kemajuan ilmu pengetahuan dan teknologi. Deteksi dini dan pendidikan pencegahan menjadi sangat penting dalam menekan angka kejadian dan kematian akibat kanker ini. Namun, tantangan besar masih ada, terutama di negara-negara berkembang yang menghadapi keterbatasan dalam infrastruktur kesehatan dan akses terhadap pengobatan yang memadai. Dengan meningkatnya pemahaman tentang faktor-faktor risiko serta pengembangan pendekatan-pendekatan baru dalam pengobatan, harapan hidup pasien kanker mulut diharapkan dapat meningkat di masa depan.

2.2.2. Oral Potentially Malignant Disorders (OPMD)

Oral Potentially Malignant Disorders (OPMD) merupakan sekelompok kondisi pada rongga mulut yang memiliki potensi untuk berkembang menjadi kanker, khususnya karsinoma sel skuamosa, yang merupakan bentuk paling umum dari kanker mulut. Menurut definisi konvensional, *OPMD* didefinisikan sebagai "sekelompok lesi dan kondisi yang ditandai dengan peningkatan risiko yang bervariasi untuk berkembang menjadi kanker pada bibir dan rongga mulut" (Warnakulasuriya et al., 2021). *OPMD* yang paling sering ditemukan meliputi *leukoplakia*, *eritroplakia*, *cheilitis actinica*, *fibrosis submukosa oral (OSF)*, serta *lichen planus oral* tipe erosif (Ahmada et al., 2024). *OPMD* menyebabkan ketidaknyamanan, masalah mengunyah dan makan, rasa sakit, serta ketakutan atau

kecemasan terhadap kemungkinan transformasi menjadi keganasan, yang memengaruhi kualitas hidup pasien secara keseluruhan (Kumar et al., 2021).



Gambar 2. *Oral potentially malignant disorders* (Piyarathne et al., 2024).

Diagnosis *Oral Potentially Malignant Disorders (OPMD)* ditegakkan melalui pemeriksaan klinis dan histopatologi (Ahmada et al., 2024). Secara histopatologis, *OPMD* dicirikan oleh adanya displasia epitel, yaitu kondisi di mana terjadi pertumbuhan sel yang tidak teratur dan perubahan pada struktur jaringan, yang menunjukkan tanda-tanda awal keganasan. Pemeriksaan klinis dan histopatologi memainkan peran yang sangat penting dalam menegakkan diagnosis, mengingat beberapa kelainan mungkin tampak jinak secara klinis tetapi menunjukkan perubahan displastik pada pemeriksaan mikroskopis. Risiko transformasi maligna dari *OPMD* bervariasi tergantung pada jenis kelainannya, dengan beberapa jenis memiliki risiko lebih tinggi dibandingkan yang lain.

Faktor-faktor risiko yang berkaitan dengan perkembangan *OPMD* termasuk kebiasaan merokok dan konsumsi alkohol yang berlebihan (Ahmada et al., 2024). Penggunaan tembakau dalam berbagai bentuk, baik itu rokok, tembakau kunyah, maupun snuff, telah diidentifikasi sebagai salah satu faktor paling signifikan dalam patogenesis *OPMD*. Demikian pula, konsumsi alkohol secara berlebihan berperan sinergis dengan tembakau dalam meningkatkan risiko kanker mulut pada individu dengan *OPMD*.

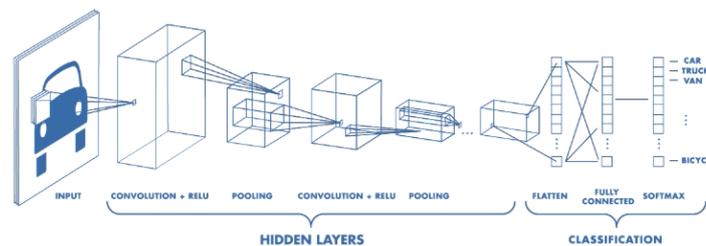
Tantangan utama dalam manajemen *OPMD* adalah ketidakpastian mengenai kemungkinan transformasi maligna dari setiap lesi. Oleh karena itu, pemantauan jangka panjang sangat diperlukan pada pasien dengan diagnosis *OPMD*, bahkan setelah intervensi awal. Lesi yang menunjukkan perubahan displastik berat atau

yang memiliki risiko transformasi lebih tinggi sering kali memerlukan biopsi berulang atau bahkan eksisi bedah untuk mengurangi risiko berkembang menjadi kanker. Selain itu, teknologi baru seperti pencitraan fluoresen dan pemetaan genetik dapat membantu dalam deteksi dini dan prediksi risiko transformasi maligna pada pasien dengan *OPMD*.

Dalam konteks pencegahan, modifikasi gaya hidup merupakan langkah yang sangat penting untuk mengurangi risiko perkembangan *OPMD* menjadi kanker. Individu yang memiliki kebiasaan merokok atau mengonsumsi alkohol dianjurkan untuk menghentikan kebiasaan tersebut, mengingat hubungan kuat antara kedua faktor ini dengan perkembangan *OPMD*. Edukasi masyarakat dan peningkatan kesadaran tentang faktor-faktor risiko *OPMD*, serta tanda-tanda awal keganasan, dapat berkontribusi dalam upaya pencegahan yang lebih efektif.

2.3. *Convolutional Neural Network (CNN)*

Convolutional Neural Network (CNN) adalah salah satu jenis arsitektur jaringan saraf tiruan (*Artificial Neural Network*) yang sangat populer dalam tugas-tugas pemrosesan citra, seperti pengenalan objek, klasifikasi gambar, dan deteksi wajah. CNN dirancang khusus untuk mengenali serta pembuatan gambar (Ding et al., 2023). CNN dirancang untuk secara otomatis dan adaptif mempelajari fitur dari data gambar melalui serangkaian layer berlapis. Arsitektur CNN terdiri dari beberapa komponen utama, termasuk layer konvolusi (*convolutional layer*), *layer pooling*, dan *layer fully-connected*, di mana masing-masing komponen berperan penting dalam proses ekstraksi fitur dan klasifikasi data citra.



Gambar 3. Ilustrasi *convolutional neural network* (MathWorks, 2021).

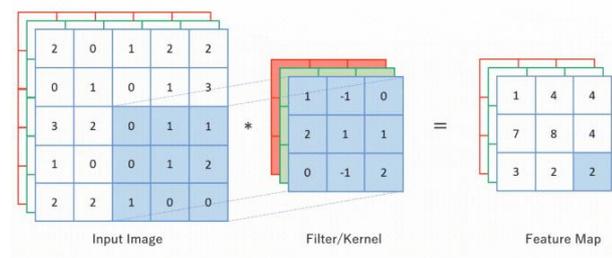
CNN secara khusus dirancang untuk belajar pola-pola spasial, seperti tepi, sudut, tekstur, hingga bentuk abstrak yang lebih kompleks, yang paling baik menggambarkan kelas target atau kuantitas yang diidentifikasi (Kattenborn et al., 2021). Melalui proses hierarkis, CNN mampu menangkap pola-pola dari level fitur yang lebih sederhana ke yang lebih kompleks, menghasilkan representasi spasial yang kaya dari input gambar. Oleh karena itu, CNN menjadi algoritma deep learning yang paling umum digunakan untuk analisis pola spasial dalam berbagai aplikasi pemrosesan citra.

2.3.1. *Feature Extraction Layer*

Feature extraction layer bertanggung jawab untuk mengekstraksi informasi penting dari gambar input yang akan digunakan untuk tugas-tugas lanjutan, seperti pengenalan objek atau klasifikasi. Pada CNN, proses ekstraksi fitur ini dilakukan oleh kombinasi *convolutional layer* dan *pooling layer*.

2.3.1.1. *Convolutional Layer*

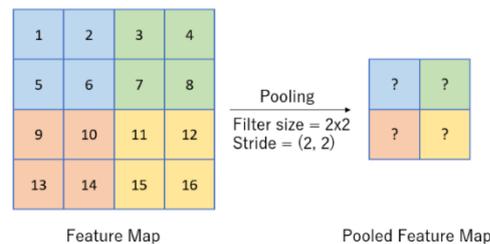
Convolutional layer atau layer konvolusi merupakan komponen inti dalam CNN yang bertanggung jawab untuk mengekstraksi fitur dari gambar input. Proses ini dilakukan dengan menerapkan filter atau kernel, yaitu matriks kecil, pada seluruh area gambar secara lokal. Kernel bergerak melintasi gambar, menghitung nilai dot product antara nilai piksel dan elemen kernel, sehingga menghasilkan peta fitur (feature map) yang mencerminkan pola atau fitur spesifik yang terdeteksi, seperti garis, tepi, atau bentuk sederhana. Pada layer awal CNN, filter biasanya mendeteksi fitur-fitur dasar, sedangkan layer yang lebih dalam dapat mengenali pola yang lebih kompleks, seperti objek atau bagian objek.



Gambar 4. Ilustrasi *convolutional layer*.

Fungsi aktivasi, seperti ReLU (*Rectified Linear Unit*), biasanya diterapkan setelah operasi konvolusi untuk memperkenalkan non-linearitas ke dalam model, memungkinkan jaringan untuk belajar lebih banyak variasi pola dari data. Hasil dari *convolutional layer* ini juga bisa disebut dengan *feature map* atau peta fitur (Pangestu et al., 2020).

2.3.1.2. *Pooling Layer*



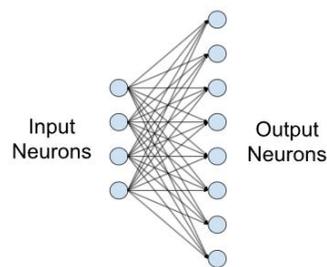
Gambar 5. Ilustrasi *pooling layer*.

Pooling layer bertanggung jawab untuk mereduksi dimensi dari peta fitur yang dihasilkan oleh layer konvolusi, sehingga mengurangi jumlah parameter yang harus dipelajari dan mempercepat proses komputasi. Pooling juga membantu membuat jaringan lebih tahan terhadap perubahan kecil pada input, seperti pergeseran atau rotasi objek pada gambar. Selain itu, pooling layer melakukan down sampling non-linear pada fitur yang telah dikonvolusi, mengurangi daya komputasi melalui reduksi dimensi (Sarvamangala & Kulkarni, 2022). Dengan mengurangi ukuran spasial dan mengagregasi data berdasarkan ruang atau jenis fitur, pooling membantu mengendalikan *overfitting* serta mengatasi variasi translasi dan rotasi pada gambar. Proses pooling ini membagi input menjadi serangkaian patch

berbentuk persegi panjang, di mana setiap patch digantikan oleh satu nilai tergantung pada jenis pooling yang dipilih, seperti *maximum pooling* atau *average pooling*.

2.3.2. *Fully-connected Layer*

Setelah fitur-fitur penting diekstraksi melalui serangkaian layer konvolusi dan pooling, hasilnya kemudian diratakan menjadi satu vektor yang akan dimasukkan ke *fully-connected layer*. Layer ini mirip dengan layer pada jaringan saraf tiruan (*artificial neural network*) tradisional, di mana setiap neuron di layer terhubung dengan semua neuron di layer sebelumnya, dan setiap koneksi memiliki bobot yang terkait dengannya (Sarvamangala & Kulkarni, 2022).



Gambar 6. Ilustrasi sederhana *fully-connected layer* (NVIDIA Corporation & affiliates, 2023).

Fully-connected layer bertanggung jawab untuk melakukan tugas klasifikasi berdasarkan fitur-fitur yang telah diekstraksi oleh layer sebelumnya. Pada tahap akhir, output dari *fully-connected layer* biasanya melewati fungsi aktivasi softmax (untuk masalah klasifikasi multikelas) atau sigmoid (untuk masalah klasifikasi biner), yang mengubah nilai output menjadi probabilitas untuk setiap kelas yang ada.

Fully-connected layer memiliki banyak parameter karena setiap neuron terhubung dengan semua neuron sebelumnya. Oleh karena itu, CNN biasanya menggunakan *fully-connected layer* hanya di bagian akhir untuk melakukan klasifikasi, sementara tugas ekstraksi fitur dilakukan oleh *convolutional layer* dan *pooling layer*.

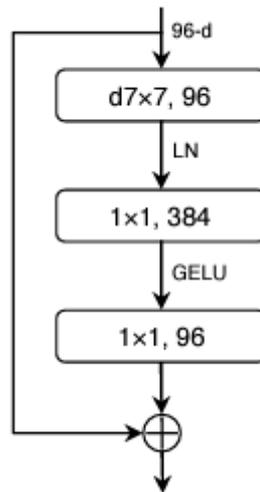
2.4. ConvNeXt

ConvNeXt merupakan arsitektur *convolutional neural network* yang diperkenalkan oleh Liu et al. (2022) dalam makalah berjudul "*A ConvNet for the 2020s*". Arsitektur ini dirancang untuk membawa CNN ke tingkat yang lebih tinggi, mengadaptasi beberapa inovasi dari model-model terbaru seperti Vision Transformer (ViT), yang telah menunjukkan performa unggul dalam berbagai tugas *computer vision*. ConvNeXt mempertahankan keunggulan dasar CNN sambil mengadopsi berbagai teknik modern, seperti yang digunakan dalam arsitektur transformer.

Pada dekade terakhir, CNN telah mendominasi bidang *computer vision* dengan berbagai arsitektur seperti AlexNet, VGG, dan ResNet. Namun, dengan munculnya model berbasis transformer seperti Vision Transformer (ViT), performa arsitektur CNN mulai tertinggal. Vision Transformer menggunakan mekanisme self-attention untuk menangkap informasi global, sementara CNN secara tradisional bekerja dengan mekanisme lokal melalui konvolusi. ConvNeXt muncul sebagai jawaban terhadap tantangan ini dengan memperbarui arsitektur CNN klasik, khususnya ResNet, untuk menyesuaikan performanya dengan standar teknologi terbaru.

ConvNeXt melakukan serangkaian modifikasi yang signifikan pada arsitektur CNN. Salah satu perubahan utamanya adalah penggunaan *depthwise convolution* dengan kernel berukuran besar (7×7), yang memungkinkan pengolahan fitur pada setiap channel secara independen, mengurangi kompleksitas parameter tanpa mengorbankan akurasi. ConvNeXt juga mengganti *batch normalization* dengan *layer normalization*, yang lebih cocok untuk arsitektur modern dan menghasilkan proses *training* yang lebih stabil.

Selain itu, ConvNeXt mengganti fungsi aktivasi ReLU yang tradisional dengan GELU (*Gaussian Error Linear Unit*), yang lebih halus dalam menentukan aktivasi neuron, sehingga meningkatkan kemampuan model dalam menangani tugas-tugas yang lebih kompleks. Penggunaan *pre-normalization* di mana normalisasi diterapkan sebelum konvolusi atau operasi lainnya mengambil inspirasi dari Vision Transformer dan terbukti mempercepat proses konvergensi selama *training*.



Gambar 7. Desain *block* ConvNeXt (Liu et al., 2022).

Arsitektur block ConvNeXt secara umum terdiri dari *depthwise convolution* (7x7) untuk ekstraksi fitur lokal, *normalization layer* untuk stabilitas, *pointwise convolution* (1x1) untuk interaksi antar-channel, dan fungsi aktivasi GELU. *Residual connection* tetap digunakan untuk mempertahankan informasi selama propagasi. Spesifikasi detail dari arsitektur ConvNeXt dapat dilihat pada Tabel 2.

Tabel 2. Spesifikasi detail arsitektur ConvNeXt Tiny

	<i>output size</i>	ConvNeXt-T
stem	56×56	$7 \times 7, 64$, stride 2 3×3 max pool, stride 2
res2	56×56	$\begin{bmatrix} d7 \times 7, 96 \\ 1 \times 1, 384 \\ 1 \times 1, 96 \end{bmatrix} \times 3$
res3	28×28	$\begin{bmatrix} d7 \times 7, 192 \\ 1 \times 1, 768 \\ 1 \times 1, 192 \end{bmatrix} \times 3$
res4	14×14	$\begin{bmatrix} d7 \times 7, 384 \\ 1 \times 1, 1536 \\ 1 \times 1, 384 \end{bmatrix} \times 9$
res5	7×7	$\begin{bmatrix} d7 \times 7, 768 \\ 1 \times 1, 3072 \\ 1 \times 1, 768 \end{bmatrix} \times 3$
	FLOPs	$4,5 \times 10^9$
	# <i>params</i>	$28,6 \times 10^6$

Dalam berbagai pengujian, ConvNeXt berhasil menunjukkan performa yang sangat kompetitif di dataset ImageNet, bahkan melampaui performa arsitektur ResNet dan Vision Transformer dalam beberapa tugas *computer vision*. ConvNeXt juga terbukti efisien dalam tugas-tugas seperti deteksi objek dan segmentasi citra, di mana ia menghasilkan hasil yang baik dengan komputasi yang lebih rendah dibandingkan transformer.

ConvNeXt menunjukkan bahwa dengan beberapa penyesuaian, arsitektur CNN masih sangat relevan dan dapat bersaing dengan teknologi terbaru dalam *computer vision*.

2.5. MobileNet

MobileNet merupakan arsitektur *convolutional neural network* (CNN) yang dirancang khusus untuk perangkat dengan keterbatasan daya dan komputasi, seperti ponsel pintar dan perangkat *Internet of Things* (IoT). Versi terbaru dari arsitektur ini, MobileNetV3, diperkenalkan oleh Howard et al. (2019) dalam makalah "*Searching for MobileNetV3*". MobileNetV3 menggabungkan efisiensi komputasi dengan akurasi yang tinggi, sehingga ideal untuk aplikasi yang dijalankan pada perangkat bergerak.

Dengan semakin meningkatnya permintaan untuk menjalankan model deep learning pada perangkat dengan keterbatasan sumber daya, seperti ponsel atau drone, efisiensi menjadi fokus utama. Model-model tradisional seperti VGG atau ResNet memiliki jumlah parameter yang besar dan memerlukan daya komputasi yang tinggi, membuat mereka kurang ideal untuk perangkat dengan keterbatasan daya. MobileNet mengatasi masalah ini dengan pendekatan yang lebih ringan dan hemat daya, tanpa mengorbankan performa. MobileNetV3 menggunakan kombinasi antara teknik *neural architecture search* (NAS) dan optimalisasi manual untuk menemukan struktur terbaik.

Salah satu inovasi utama yang diusung oleh MobileNetV3 adalah *depthwise separable convolution*, yang pertama kali diperkenalkan pada MobileNetV1. Teknik ini memisahkan operasi konvolusi menjadi dua tahap: *depthwise*

convolution, yang melakukan konvolusi pada setiap channel secara independen, dan *pointwise convolution*, yang menggabungkan hasil dari *depthwise convolution*. Teknik ini secara signifikan mengurangi jumlah parameter dan operasi komputasi tanpa menurunkan kualitas fitur yang diekstraksi dari gambar.

MobileNetV3 juga menggunakan *inverted residual block*, yang pertama kali diperkenalkan pada MobileNetV2. Blok ini memungkinkan pengurangan dimensi yang efisien melalui struktur *bottleneck*, di mana fitur yang lebih luas diproses melalui layer sempit sebelum dilanjutkan ke layer berikutnya. Hal ini meningkatkan efisiensi dalam hal jumlah parameter dan mempercepat komputasi.

Selain itu, MobileNetV3 memperkenalkan fungsi aktivasi baru yang disebut *Hard-Swish* (H-Swish), yang menggantikan fungsi aktivasi ReLU. H-Swish memberikan peningkatan akurasi tanpa meningkatkan kompleksitas komputasi secara signifikan, karena memiliki sifat non-linearitas yang lebih halus dibandingkan ReLU. Fungsi aktivasi ini sangat efektif dalam meningkatkan performa model pada perangkat bergerak.

MobileNetV3 juga mengintegrasikan *squeeze-and-excitation* (SE) modules, yang secara dinamis menyesuaikan bobot setiap channel berdasarkan pentingnya fitur yang diekstraksi. SE modules membantu memperkuat fitur-fitur penting sambil mengurangi noise, sehingga memungkinkan model untuk fokus pada bagian gambar yang lebih relevan. Inovasi ini semakin meningkatkan efisiensi komputasi, terutama pada aplikasi dengan sumber daya terbatas.

MobileNetV3 hadir dalam dua varian utama, yaitu MobileNetV3-Large dan MobileNetV3-Small. MobileNetV3-Large dirancang untuk aplikasi yang membutuhkan performa tinggi, sementara MobileNetV3-Small lebih dioptimalkan untuk perangkat dengan keterbatasan daya dan memori yang lebih ekstrem. Kedua varian ini memberikan fleksibilitas untuk berbagai jenis aplikasi, baik yang membutuhkan daya komputasi besar maupun yang membutuhkan efisiensi ekstrem. Spesifikasi detail dari arsitektur MobileNetV3-Large dan MobileNetV3-Small dapat dilihat pada Tabel 3 dan Tabel 4.

Tabel 3. Spesifikasi MobileNetV3-Large

Input	Operator	exp size	#out	SE	NL	s
$224^2 \times 3$	conv2d	-	16	-	HS	2
$112^2 \times 16$	bneck, 3×3	16	16	-	RE	1
$112^2 \times 16$	bneck, 3×3	64	24	-	RE	2
$56^2 \times 24$	bneck, 3×3	72	24	-	RE	1
$56^2 \times 24$	bneck, 5×5	72	40	✓	RE	2
$28^2 \times 40$	bneck, 5×5	120	40	✓	RE	1
$28^2 \times 40$	bneck, 5×5	120	40	✓	RE	1
$28^2 \times 40$	bneck, 3×3	240	80	-	HS	2
$14^2 \times 80$	bneck, 3×3	200	80	-	HS	1
$14^2 \times 80$	bneck, 3×3	184	80	-	HS	1
$14^2 \times 80$	bneck, 3×3	184	80	-	HS	1
$14^2 \times 80$	bneck, 3×3	480	112	✓	HS	1
$14^2 \times 112$	bneck, 3×3	672	112	✓	HS	1
$14^2 \times 112$	bneck, 5×5	672	160	✓	HS	2
$7^2 \times 160$	bneck, 5×5	960	160	✓	HS	1
$7^2 \times 160$	bneck, 5×5	960	160	✓	HS	1
$7^2 \times 160$	conv2d, 1×1	-	960	-	HS	1
$7^2 \times 960$	pool, 7×7	-	-	-	-	1
$1^2 \times 960$	conv2d, 1×1 , NBN	-	1280	-	HS	1
$1^2 \times 1280$	conv2d, 1×1 , NBN	-	k	-	-	1

Tabel 4. Spesifikasi MobileNetV3-Small

Input	Operator	exp size	#out	SE	NL	s
$224^2 \times 3$	conv2d	-	16	-	HS	2
$112^2 \times 16$	bneck, 3×3	16	16	✓	RE	2
$56^2 \times 16$	bneck, 3×3	72	24	-	RE	2
$28^2 \times 24$	bneck, 3×3	88	24	-	RE	1

Input	Operator	exp size	#out	SE	NL	s
$28^2 \times 24$	bneck, 5×5	96	40	✓	HS	2
$14^2 \times 40$	bneck, 5×5	240	40	✓	HS	1
$14^2 \times 40$	bneck, 5×5	240	40	✓	HS	1
$14^2 \times 40$	bneck, 5×5	120	48	✓	HS	1
$14^2 \times 40$	bneck, 5×5	144	48	✓	HS	1
$14^2 \times 48$	bneck, 5×5	288	96	✓	HS	2
$7^2 \times 48$	bneck, 5×5	576	96	✓	HS	1
$7^2 \times 96$	bneck, 5×5	576	96	✓	HS	1
$7^2 \times 96$	conv2d, 1×1	-	576	✓	HS	1
$7^2 \times 576$	pool, 7×7	-	-	-	-	1
$1^2 \times 576$	conv2d, 1×1 , NBN	-	1024	-	HS	1
$1^2 \times 1280$	conv2d, 1×1 , NBN	-	k	-	-	1

Dari segi jumlah parameter dan operasi komputasi, MobileNetV3 menunjukkan peningkatan efisiensi dibandingkan dengan model sebelumnya. MobileNetV3-Large memiliki sekitar 5.4 juta parameter dan membutuhkan sekitar 219 juta Floating Point Operations per Second (FLOPs) pada input gambar berukuran 224×224 piksel. MobileNetV3-Small lebih ringan, dengan sekitar 2.9 juta parameter dan hanya membutuhkan sekitar 66 juta FLOPs untuk input gambar berukuran yang sama.

Dari segi performa, MobileNetV3 diuji pada dataset ImageNet dan menunjukkan hasil yang sangat kompetitif, baik dari segi akurasi maupun efisiensi komputasi. Selain tugas klasifikasi gambar, MobileNetV3 juga dapat digunakan untuk tugas lain seperti deteksi objek dan segmentasi citra, dengan hasil yang tetap memadai pada perangkat dengan daya terbatas.

Secara keseluruhan, MobileNetV3 merupakan solusi optimal untuk penerapan *deep learning* pada perangkat bergerak. Dengan kombinasi teknik AutoML, *depthwise separable convolution*, *inverted residual block*, *H-Swish activation*, dan *SE modules*, MobileNetV3 menawarkan keseimbangan sempurna antara efisiensi dan

akurasi. Model ini membuka jalan untuk penerapan teknologi deep learning di perangkat yang sebelumnya terbatas, seperti ponsel pintar, drone, dan perangkat IoT, tanpa harus mengorbankan performa.

2.6. EfficientNet

Arsitektur EfficientNet diperkenalkan oleh Tan & Le (2019) dalam makalah “*EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*” sebagai jawaban terhadap tantangan skala model pada *Convolutional Neural Network* (CNN). Sebelum kemunculannya, model scaling pada CNN biasanya dilakukan dengan memperbesar *depth*, *width*, atau resolusi input secara terpisah, yang seringkali menghasilkan peningkatan sumber daya komputasi tanpa peningkatan akurasi yang signifikan. Pendekatan tradisional seperti ini terbukti tidak efisien, terutama ketika model diterapkan pada perangkat dengan keterbatasan komputasi, seperti ponsel pintar atau sistem embedded. Dalam white paper mereka, Tan dan Le memperkenalkan compound scaling sebagai solusi untuk meningkatkan efisiensi model secara menyeluruh.

EfficientNet memperkenalkan konsep compound scaling untuk menyelaraskan peningkatan *depth* (jumlah layer), *width* (jumlah filter per layer), dan resolusi input gambar secara proporsional. Metode ini berbeda dengan pendekatan sebelumnya, yang sering memperbesar salah satu dimensi tanpa mempertimbangkan pengaruh terhadap dimensi lainnya. Dengan memformulasikan skala model menggunakan rumus proporsionalitas yang spesifik, *compound scaling* memungkinkan peningkatan kinerja yang signifikan tanpa secara eksponensial meningkatkan jumlah parameter atau biaya komputasi. Rumus scaling yang digunakan dapat dilihat pada rumus (1).

$$d = \alpha^\phi, w = \beta^\phi, r = \gamma^\phi \quad (1)$$

Dimana d adalah *depth*, w adalah *width*, dan r adalah resolusi input gambar. Parameter α , β , dan γ merupakan konstanta scaling yang dihitung berdasarkan eksperimen, sementara ϕ adalah variabel yang mengontrol tingkat scaling.

Dalam pengembangannya, EfficientNet memanfaatkan *neural architecture search* (NAS) untuk menemukan arsitektur dasar yang optimal, disebut EfficientNet-B0, yang digunakan sebagai pondasi untuk model lain dalam keluarga EfficientNet (EfficientNet-B1 hingga EfficientNet-B7). NAS merupakan teknik yang secara otomatis menjelajahi ruang arsitektur model untuk menemukan desain yang paling efisien berdasarkan keterbatasan komputasi dan kebutuhan tugas spesifik. Dalam penelitian Tan dan Le, NAS digunakan untuk mengeksplorasi variasi arsitektur dasar yang dioptimalkan untuk ukuran model dan kinerja yang seimbang. Hasilnya, EfficientNet-B0 memiliki arsitektur dasar yang ringan dan efisien, sehingga memudahkan scaling ke versi yang lebih besar tanpa pemborosan sumber daya. Spesifikasi dasar dari EfficientNet bisa dilihat pada Tabel 5.

Tabel 5. EfficientNet-B0 *baseline network*

Stage i	Operator \hat{F}_i	Resolution $\hat{H}_i \times \hat{W}_i$	#Channels \hat{C}_i	#Layers \hat{L}_i
1	Conv3x3	224×224	32	1
2	MBCConv,k3x3	112×112	16	1
3	MBCConv,k3x3	112×112	24	2
4	MBCConv,k5x5	56×56	40	2
5	MBCConv,k3x3	28×28	80	3
6	MBCConv,k5x5	14×14	112	3
7	MBCConv,k5x5	14×14	192	4
8	MBCConv,k3x3	7×7	320	1
9	Conv1x1 & Pooling & FC	7×7	1280	1

Selain *compound scaling*, salah satu keunggulan utama dari arsitektur EfficientNet adalah efisiensi parameter. Model seperti ResNet, DenseNet, atau Inception seringkali mengalami pertumbuhan parameter yang tidak linear seiring dengan

peningkatan skala, yang menyebabkan tingginya kebutuhan memori dan daya komputasi. Sebagai perbandingan, EfficientNet mencapai trade-off yang lebih optimal antara akurasi dan jumlah parameter, yang ditunjukkan oleh kinerja unggul pada dataset ImageNet. Misalnya, EfficientNet-B7 mencapai akurasi top-1 sekitar 84.4% pada ImageNet, lebih tinggi daripada model-model seperti ResNet-152 atau Inception-ResNet, tetapi dengan jumlah parameter yang jauh lebih sedikit.

Untuk EfficientNet-B0, model ini memiliki sekitar 5.3 juta parameter dan membutuhkan sekitar 390 juta FLOPs (Floating Point Operations per Second). Dengan jumlah parameter yang relatif kecil, EfficientNet-B0 tetap mampu memberikan akurasi yang kompetitif dibandingkan dengan model lain yang lebih besar, menjadikannya pilihan ideal untuk aplikasi dengan keterbatasan komputasi.

Sebuah aspek yang sangat penting dari EfficientNet adalah kemampuannya untuk beroperasi dengan baik pada perangkat dengan keterbatasan komputasi, seperti ponsel pintar. Di era modern, banyak aplikasi kecerdasan buatan yang dijalankan pada perangkat bergerak, yang seringkali memiliki sumber daya komputasi terbatas seperti memori, daya prosesor, dan energi. EfficientNet memungkinkan pengembangan model yang ringan dan efisien, yang bisa diimplementasikan pada perangkat seluler tanpa mengorbankan akurasi. Sebagai contoh, varian EfficientNet-B0 digunakan dalam pengenalan gambar pada ponsel dan perangkat IoT karena ukuran modelnya yang kecil namun tetap mampu menghasilkan hasil yang akurat.

Lebih jauh lagi, model-model EfficientNet tidak hanya terbukti efisien dalam image classification, tetapi juga di berbagai tugas *computer vision* lainnya, seperti object detection dan segmentation. Dalam tugas object detection, kombinasi EfficientNet dengan *framework Faster R-CNN* telah menunjukkan peningkatan kinerja yang signifikan dibandingkan model *backbone* lain seperti ResNet atau Inception. Bahkan, dalam beberapa eksperimen pada dataset COCO, EfficientNet menunjukkan *mean Average Precision* (mAP) yang lebih tinggi dengan waktu inferensi yang lebih cepat, menjadikannya pilihan unggul untuk tugas deteksi objek yang membutuhkan presisi tinggi dan latensi rendah.

Selain itu, dalam konteks *transfer learning*, EfficientNet juga menawarkan efisiensi sangat baik. Model ini dapat di-*fine-tune* untuk tugas-tugas spesifik dengan biaya komputasi yang lebih rendah dibandingkan model lain yang lebih kompleks. Pada berbagai studi kasus di domain medical imaging dan autonomous driving, EfficientNet terbukti mudah diadaptasi dan cepat mencapai performa tinggi setelah di-*fine-tune* pada dataset yang lebih kecil. Hal ini membuat arsitektur ini sangat ideal untuk proyek-proyek yang melibatkan domain-spesifik data dan memerlukan pelatihan ulang dengan cepat.

Dari segi implementasi, salah satu alasan popularitas EfficientNet adalah ketersediaannya dalam berbagai *framework deep learning* populer seperti TensorFlow, PyTorch, dan Keras. Dengan dukungan yang kuat dari komunitas *open-source*, EfficientNet telah menjadi salah satu model yang paling sering digunakan dalam penelitian dan industri. Ini memudahkan integrasi ke dalam *pipeline machine learning* yang sudah ada dan memungkinkan penggunaan arsitektur ini pada berbagai skenario aplikasi, baik pada server besar maupun perangkat dengan keterbatasan sumber daya.

Dengan segala keunggulannya, EfficientNet telah membawa perubahan signifikan dalam cara kita memandang pengembangan arsitektur CNN modern. Tidak hanya memberikan peningkatan dalam akurasi, arsitektur ini juga menekankan pentingnya efisiensi dalam hal sumber daya komputasi, yang sangat krusial untuk aplikasi di dunia nyata. Model ini berhasil menunjukkan bahwa pendekatan yang lebih efisien dalam *scaling* model dapat mengurangi kompleksitas tanpa mengorbankan performa, suatu kemajuan yang relevan dalam perkembangan teknologi *machine learning* saat ini.

2.7. Augmentasi Data

Augmentasi data (*data augmentation*) adalah teknik yang menggunakan berbagai transformasi untuk menghasilkan data sintetis dari dataset yang sudah ada. Teknik ini memainkan peran penting dalam pelatihan model *deep neural networks* (DNN), karena memungkinkan model menerima variasi yang lebih luas dari data, tanpa harus mengumpulkan data baru (Shorten et al., 2021). Augmentasi data biasanya

melibatkan perubahan-perubahan kecil yang tidak mengubah label kelas, seperti rotasi, *flipping*, *zooming*, dan translasi. Tujuannya adalah untuk memperkaya keragaman data pelatihan, sehingga model menjadi lebih tangguh dalam menghadapi data baru yang belum pernah dilihat sebelumnya.

Teknik augmentasi data merupakan bagian dari metode regularisasi yang bertujuan untuk meningkatkan generalisasi model (Mumuni & Mumuni, 2022). Salah satu tantangan terbesar dalam membangun model *machine learning* adalah *overfitting*, di mana model terlalu terfokus pada detail spesifik dari data pelatihan sehingga kehilangan kemampuannya untuk memprediksi dengan baik pada data baru. Augmentasi data membantu mencegah *overfitting* dengan memperkenalkan variasi tambahan pada data pelatihan, yang membuat model lebih adaptif terhadap perubahan kecil dalam data input dan membantu menangkap pola-pola yang lebih umum dan relevan (Shorten et al., 2021).

2.8. *Hyperparameter*

Hyperparameter adalah variabel konfigurasi eksternal yang digunakan oleh data scientist untuk mengelola pelatihan suatu model *machine learning* dan perlu diatur sebelum pelatihan dimulai (Amazon Web Services, 2024). Berbeda dengan parameter model yang diperoleh melalui proses pembelajaran dari data, *hyperparameter* ditentukan secara manual oleh pengembang atau melalui metode otomatis seperti pencarian grid (*grid search*) atau pencarian acak (*random search*).

Pada dasarnya, *hyperparameter* berfungsi untuk mengontrol perilaku model, seperti struktur jaringan, kecepatan pembelajaran, jumlah iterasi, hingga cara optimalisasi dilakukan. Pemilihan dan penyetelan *hyperparameter* yang tepat sangat krusial dalam memaksimalkan performa model.

2.8.1. *Epoch*

Epoch adalah satu siklus lengkap di mana seluruh dataset dilalui oleh model selama pelatihan, dan selama setiap *epoch*, bobot model diperbarui berdasarkan kesalahan prediksi. Jumlah *epoch* yang digunakan memiliki pengaruh signifikan terhadap

hasil akhir; terlalu sedikit *epoch* dapat menyebabkan *underfitting*, sementara terlalu banyak *epoch* dapat menyebabkan *overfitting*. Jumlah *epoch* optimal untuk CNN bergantung pada kompleksitas masalah dan ukuran dataset—semakin besar dan kompleks dataset, semakin banyak *epoch* yang diperlukan untuk mencapai hasil yang optimal (Wibawa, 2017). Oleh karena itu, pemilihan jumlah *epoch* yang tepat sangat penting untuk mencapai keseimbangan antara pembelajaran dan generalisasi.

2.8.2. *Batch-size*

Batch size merujuk pada jumlah sampel data yang digunakan dalam satu iterasi pembaruan bobot selama pelatihan model. Iterasi adalah proses pembacaan data dari depan ke belakang (Rochmawati et al., 2021). Memilih ukuran batch yang tepat dapat memengaruhi kecepatan pelatihan dan akurasi model. Ukuran batch yang lebih kecil dapat memberikan pembaruan yang lebih sering dan lebih responsif terhadap perubahan dalam data, tetapi mungkin memperlambat proses pelatihan secara keseluruhan. Di sisi lain, ukuran batch yang lebih besar dapat mempercepat proses pelatihan dengan memungkinkan pemrosesan lebih banyak data sekaligus, tetapi bisa mengurangi kemampuan model untuk belajar dari variasi dalam data. Oleh karena itu, pemilihan ukuran batch yang sesuai harus mempertimbangkan *trade-off* antara kecepatan pelatihan dan kualitas model.

2.8.3. *Optimizer AdamW*

Optimizer adalah algoritma yang digunakan untuk memperbarui bobot model berdasarkan gradien yang dihitung. Salah satu *optimizer* yang banyak digunakan adalah AdamW (*Adaptive Moment Estimation with Weight Decay*). AdamW diperkenalkan oleh (Loshchilov & Hutter, 2017) dalam makalah mereka yang berjudul "*Decoupled Weight Decay Regularization*". AdamW adalah varian dari Adam dimana weight decay dilakukan hanya setelah mengendalikan ukuran langkah per parameter (Llugsi et al., 2021).

Weight decay adalah teknik regularisasi yang menambahkan penalti ke dalam pembaruan bobot untuk mengontrol pertumbuhan bobot agar tidak terlalu besar, membantu mencegah *overfitting* dan meningkatkan kemampuan generalisasi model. Dalam AdamW, komponen *weight decay* diterapkan secara terpisah setelah pembaruan bobot, berbeda dari Adam yang menggabungkan regularisasi L2 ke dalam update parameter. Dengan memisahkan *weight decay* dari proses *update* gradien, AdamW menjaga bobot model tetap optimal tanpa memengaruhi langkah pembelajaran itu sendiri, sehingga regulasi parameter menjadi lebih efektif.

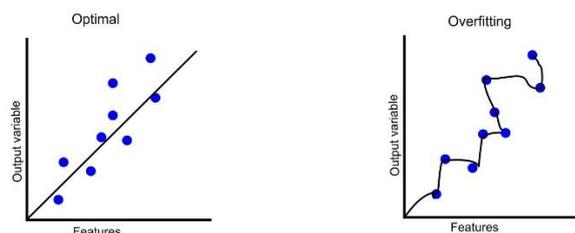
AdamW menggabungkan keuntungan dari algoritma Adam dengan tambahan regularisasi *weight decay* yang lebih optimal. Sama seperti Adam, AdamW memiliki *hyperparameter* tambahan seperti β_1 dan β_2 , yang mengontrol *decay rate* dari momentum, serta *learning rate* yang sangat mempengaruhi kecepatan dan stabilitas pembelajaran. Dengan penyesuaian yang tepat, *optimizer* AdamW sering kali memberikan konvergensi yang lebih baik dan stabil, terutama pada model yang kompleks dan data yang besar, berisik, atau sparse.

Dengan pendekatan *weight decay* yang terpisah ini, AdamW mendorong pembaruan bobot ke arah yang lebih seimbang dan mencegah mereka tumbuh terlalu besar selama proses pelatihan, yang meningkatkan generalisasi dan stabilitas model secara keseluruhan.

2.8.4. *Learning-rate*

Learning rate adalah *hyperparameter* penting yang mengontrol seberapa besar perubahan bobot model selama pembaruan dalam proses *training*. Pemilihan *learning rate* menjadi tantangan, karena nilai yang terlalu tinggi dapat menyebabkan pelatihan tidak stabil dan melewati minimum global, sementara nilai yang terlalu rendah memperlambat konvergensi dan membuat proses pelatihan lebih lama (Rochmawati et al., 2021). Oleh karena itu, pemilihan *learning rate* yang tepat sangat penting untuk performa optimal, dan banyak penelitian menunjukkan bahwa teknik penjadwalan *learning rate* dapat membantu mempercepat konvergensi dan meningkatkan hasil akhir.

2.9. *Overfitting*



Gambar 8. Ilustrasi *overfitting*.

Overfitting merupakan suatu kondisi dalam machine learning di mana model bekerja dengan sangat baik pada data pelatihan tetapi gagal tampil optimal pada data uji. Lebih spesifiknya, model mempelajari pola-pola noise yang ada dalam data pelatihan, sehingga terlihat adanya perbedaan yang besar antara kesalahan pada data pelatihan dan data uji (Salman & Liu, 2019). Dalam kondisi ini, model tidak hanya mengidentifikasi pola yang relevan untuk pembentukan data, tetapi juga menangkap gangguan atau noise spesifik pada sampel tersebut. Dengan kata lain, model menjadi terlalu terfokus pada detail acak yang mungkin tidak relevan, yang mengakibatkan model hanya "menghafal" pola dari data latih tanpa benar-benar mempelajari pola yang relevan untuk memprediksi data baru. Hal ini menyebabkan performa model yang buruk pada data uji atau data di luar sampel (*out-of-sample data*) (Peng & Nagata, 2020).

2.10. *Confusion matrix*

Confusion matrix adalah alat evaluasi yang digunakan untuk menilai kinerja algoritma klasifikasi terhadap masalah klasifikasi. Matriks ini memvisualisasikan hasil klasifikasi dalam bentuk tabel yang merinci prediksi model berdasarkan kategori sebenarnya dan prediksi model. Setiap baris dalam *confusion matrix* mewakili kelas yang sebenarnya, sementara setiap kolom mewakili kelas yang

diprediksi. Matriks ini menampilkan jumlah prediksi benar dan salah untuk setiap kelas, memungkinkan analisis lebih mendalam tentang kinerja model.

Confusion matrix terdiri dari empat metrik utama, yaitu *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN), yang masing-masing membantu mengevaluasi seberapa baik model dalam mengklasifikasikan data. Elemen diagonal pada matriks menunjukkan prediksi yang benar (TP dan TN), sedangkan elemen di luar diagonal menunjukkan prediksi yang salah (FP dan FN) (Maxwell et al., 2021). Selain itu, *confusion matrix* memungkinkan perhitungan metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengukur kualitas prediksi yang dihasilkan oleh model klasifikasi. Tabel *confusion matrix* dapat dilihat pada Tabel 6.

Tabel 6. *Confusion matrix* (Maxwell et al., 2021)

		<i>Predicted Class</i>	
		<i>Positive</i>	<i>Negative</i>
<i>Actual Class</i>	<i>Positive</i>	TP	FN
	<i>Negative</i>	FP	TN

Keterangan:

- 1) *True Positive* (TP) = Data positif yang terdeteksi positif
- 2) *False Positive* (FP) = Data positif yang terdeteksi negatif
- 3) *True Negative* (TN) = Data negatif yang terdeteksi negatif
- 4) *False Negative* (FN) = Data negatif yang terdeteksi positif

Confusion matrix telah lama digunakan dalam evaluasi model ilmiah dan aplikasi teknik, dan sering diterapkan di berbagai bidang seperti *computer vision*, pengenalan pola, dan deteksi objek (Li & Deng, 2019). Matriks ini memainkan peran penting dalam mengukur kinerja model klasifikasi dengan menunjukkan distribusi prediksi yang benar dan salah untuk setiap kategori. Ini memungkinkan para peneliti dan pengembang model untuk menganalisis kualitas model secara lebih mendalam dan akurat.

Confusion matrix paling umum digunakan dalam model klasifikasi *multi-class*, *single-label*, di mana setiap instance data hanya dapat tergolong dalam satu kelas pada satu waktu tertentu. Dengan struktur dua dimensi, matriks ini menyajikan hasil klasifikasi dalam bentuk tabel yang mengidentifikasi jumlah instance yang diklasifikasikan dengan benar dan salah oleh model, berdasarkan label yang benar dan prediksi model (Krstinić et al., 2020).

Selain itu, *confusion matrix* juga menghasilkan metrik evaluasi yang sangat penting dalam menilai kinerja sebuah *classifier*. Berikut adalah beberapa metrik evaluasi yang umum digunakan dan dapat dihitung berdasarkan *confusion matrix*.

2.10.1. Akurasi (*Accuracy*)

Akurasi adalah persentase dari total prediksi yang benar dibandingkan dengan semua prediksi yang dibuat. Ini memberikan gambaran umum tentang kinerja model secara keseluruhan. Rumus yang dipakai untuk mengukur tingkat akurasi dapat dilihat pada rumus (2).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (2)$$

2.10.2. Presisi (*Precision*)

Presisi mengukur seberapa banyak dari semua prediksi positif yang benar-benar positif. Ini sangat penting ketika konsekuensi dari false positives tinggi. Rumus yang dipakai untuk mengukur tingkat presisi dapat dilihat pada rumus (3).

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (3)$$

2.10.3. *Recall*

Recall mengukur seberapa baik model mengenali seluruh kelas positif, atau berapa banyak dari total kelas positif yang berhasil diidentifikasi dengan benar oleh model. Rumus yang dipakai untuk mengukur *recall* dapat dilihat pada rumus (4).

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (4)$$

2.10.4. *F1-score*

F1-score adalah rata-rata harmonis dari presisi dan *recall*. Metrik ini sering digunakan ketika terdapat ketidakseimbangan antara presisi dan *recall*, atau ketika kita ingin mendapatkan keseimbangan yang baik antara keduanya. Rumus yang dipakai untuk mengukur *F1-score* dapat dilihat pada rumus (5).

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100\% \quad (5)$$

2.11. *Framework OSEM N*



Gambar 9. *Framework OSEM N* (Data Science PM, 2024).

Framework OSEM N adalah salah satu pendekatan standar dalam proyek data science. *Framework* ini diperkenalkan oleh Hilary Mason dan Chris Wiggins pada artikel yang berjudul “*A Taxonomy of Data Science.*” *Framework* ini terdiri dari lima tahap utama: Obtain (Pengumpulan Data), Scrub (Pembersihan Data), Explore

(Eksplorasi Data), *Model* (Pemodelan), dan *Interpret* (Interpretasi Hasil) (Mason & Wiggins, 2010).

Framework OSEMN dirancang untuk memudahkan pengolahan dan analisis data secara sistematis sehingga dapat menghasilkan output yang akurat dan dapat diinterpretasikan dengan baik.

2.11.1. Obtain

Obtain adalah proses pengumpulan data dari berbagai sumber yang relevan untuk digunakan dalam analisis. Pada tahap ini, data dapat diperoleh dari berbagai sumber, seperti API, database, file CSV, atau bahkan melalui *web scraping*. Data yang dikumpulkan bisa berbentuk terstruktur, seperti tabel, atau tidak terstruktur, seperti gambar, teks, atau audio. Tantangan dalam tahap ini adalah mendapatkan data yang relevan, akurat, dan lengkap. Ketersediaan data yang tepat dan berkualitas pada tahap ini sangat penting untuk menjamin kesuksesan langkah-langkah selanjutnya dalam analisis data.

2.11.2. Scrub

Scrub adalah proses pembersihan data untuk memastikan data bebas dari kesalahan, duplikasi, atau ketidakkonsistenan yang bisa memengaruhi hasil analisis. Data mentah sering kali mengandung nilai yang hilang, duplikasi, atau format yang tidak seragam. Oleh karena itu, langkah ini melibatkan identifikasi dan perbaikan masalah-masalah tersebut. Proses ini dapat mencakup penghapusan data duplikat, penanganan data hilang, dan normalisasi atau standarisasi format data. Pembersihan data sangat penting untuk meningkatkan kualitas analisis dan menghindari kesalahan yang dapat muncul akibat data yang kotor.

2.11.3. Explore

Explore adalah tahap eksplorasi di mana data dianalisis secara mendalam untuk memahami pola, hubungan, dan karakteristik data secara umum. Setelah data dibersihkan, tahap eksplorasi memungkinkan data scientist untuk lebih memahami

data. Analisis eksploratif melibatkan penggunaan statistik deskriptif seperti *mean*, *median*, dan standar deviasi, serta visualisasi data seperti grafik distribusi, *scatter plot*, dan *heatmap*. Tujuannya adalah untuk menemukan pola, outlier, atau hubungan antara variabel. Wawasan yang diperoleh dari tahap ini sangat membantu dalam mengarahkan keputusan untuk pemodelan dan memungkinkan penyesuaian lebih lanjut terhadap data jika diperlukan.

2.11.4. Model

Model adalah proses membangun model prediktif atau deskriptif menggunakan teknik statistik atau algoritma *machine learning*. Tahap ini merupakan inti dari *framework* OSEMN, di mana data yang telah diproses digunakan untuk melatih model. Pemilihan algoritma yang tepat, seperti regresi, klasifikasi, atau clustering, sangat bergantung pada tujuan analisis. Setelah model dilatih, evaluasi performa dilakukan menggunakan metrik seperti *accuracy*, *precision*, *recall*, atau *F1-score*. Hasil dari model ini diharapkan dapat memberikan prediksi atau klasifikasi yang akurat yang bisa digunakan untuk pengambilan keputusan selanjutnya.

2.11.5. iNterpret

iNterpret adalah proses interpretasi hasil dari model yang telah dibangun untuk memberikan insight yang dapat diambil keputusan. Langkah terakhir ini adalah menganalisis hasil dari model yang telah dibuat dan menjelaskan implikasinya kepada pemangku kepentingan. Interpretasi hasil tidak hanya mencakup angka atau prediksi, tetapi juga memahami konteks dari hasil tersebut. Data scientist perlu menyampaikan insight ini dalam bentuk laporan atau presentasi yang mudah dipahami oleh pihak non-teknis. Selain itu, hasil ini juga dapat digunakan sebagai dasar rekomendasi atau tindakan lebih lanjut untuk mencapai tujuan bisnis atau riset.

III. METODOLOGI PENELITIAN

3.1. Waktu dan Tempat

3.1.1. Waktu Penelitian

Penelitian ini dilaksanakan pada semester ganjil tahun akademik 2024/2025, dimulai dari bulan September 2024 hingga bulan Desember 2024. Waktu penelitian akan dijelaskan dengan tabel *gantt chart* pada Tabel 7.

Tabel 7. Rencana penelitian

Tahapan	2024													
	Sep				Okt				Nov				Des	
	1	2	3	4	1	2	3	4	1	2	3	4	1	2
Obtain	■	■	■											
Scrub			■	■	■	■	■							
Explore							■							
Model							■	■						
iNterpret								■	■	■	■			
Penulisan Laporan												■	■	■

3.1.2. Tempat Penelitian

Penelitian ini dilaksanakan di lingkungan Jurusan Ilmu Komputer Universitas Lampung, Jalan Prof. Dr. Ir. Sumantri Brojonegoro No.1, Gedong Meneng, Kecamatan Rajabasa, Kota Bandar Lampung, Provinsi Lampung.

3.2. Perangkat Penelitian

3.2.1. Perangkat Keras

Perangkat keras yang digunakan dalam penelitian ini adalah sebuah laptop dengan spesifikasi dan detail dijelaskan pada Tabel 8 berikut.

Tabel 8. Spesifikasi perangkat keras

Spesifikasi	Detail
Merk	HP
Tipe	OMEN 15
Model	en1029AX
CPU	AMD Ryzen 7 5800H
GPU	NVIDIA RTX 3060
RAM	16GB DDR4
Penyimpanan	SSD M.2 NVME Gen.3 1 TB

3.2.2. Perangkat Lunak

Perangkat Lunak yang digunakan dalam proses penelitian ini adalah sebagai berikut:

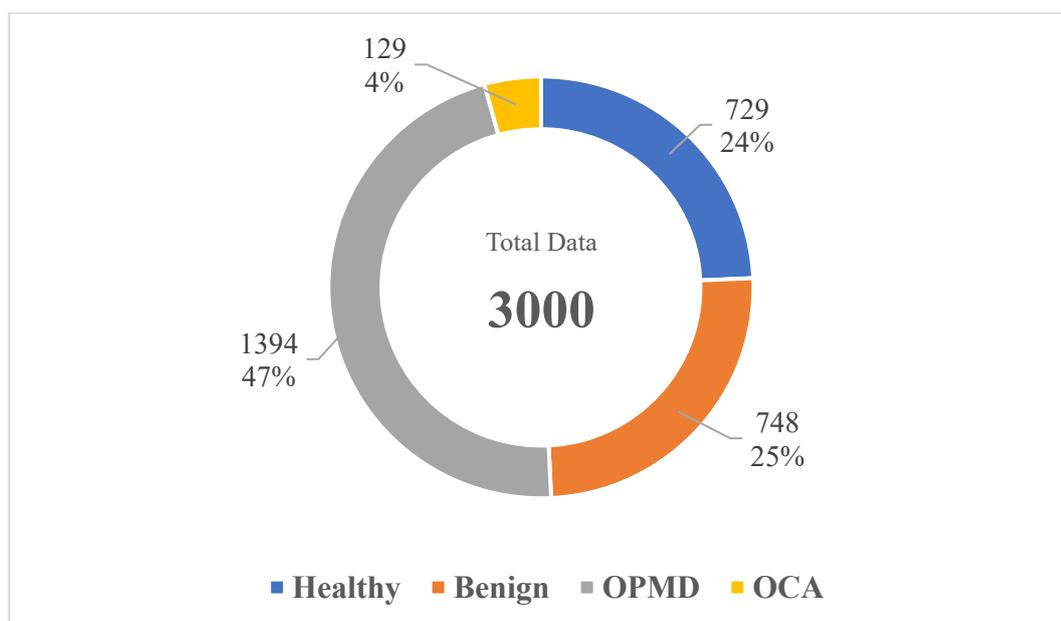
- 1) Sistem Operasi Windows 11 Home 64-bit
- 2) Visual Studio Code versi 1.95.3
- 3) Anaconda Navigator versi 2.6.0
- 4) Python versi 3.12.3

3.3. Dataset

Dataset yang digunakan dalam penelitian ini bersumber dari riset yang dilakukan oleh Piyaathne et al. (2024), yang dipublikasikan dalam artikel yang berjudul “*A comprehensive dataset of annotated oral cavity images for diagnosis of oral cancer*”

and oral potentially malignant disorders”. Dataset tersebut dapat diakses melalui *platform Zenodo Data Repository*.

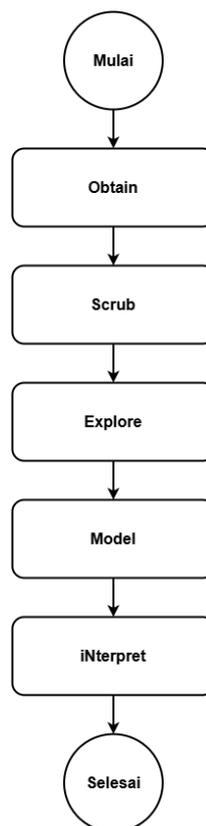
Data foto tersebut terbagi atas empat kondisi yaitu *Healthy* (sehat), *OCA* (*Oral Cancer*), *OPMD* (*Oral Potentially Malignant Disorders*), dan *Benign* (Kondisi sakit namun tidak termasuk kedalam kategori *OCA* maupun *OPMD*), dengan komposisi dapat dilihat pada Gambar 10.



Gambar 10. Komposisi dataset

3.4. Tahap Penelitian

Penelitian ini menggunakan *framework* OSEMN, yang diperkenalkan oleh Mason & Wiggins (2010), sebagai panduan alur kerja. *Framework* OSEMN dirancang untuk memudahkan pengolahan dan analisis data secara sistematis sehingga dapat menghasilkan output yang akurat dan dapat diinterpretasikan dengan baik. Alur dari tahapan yang akan dilakukan pada penelitian ini diilustrasikan pada Gambar 11.



Gambar 11. Tahap penelitian.

3.4.1. Obtain

Tahap pertama yang dilakukan dalam penelitian adalah Obtain. Tahap ini bertujuan untuk mengumpulkan dan mendapatkan data dari berbagai sumber sebagai bahan untuk penelitian. Dataset yang dikumpulkan sebanyak 3000 citra mulut dengan 4 kelas kondisi. Citra-citra tersebut disimpan dalam penyimpanan lokal untuk mempermudah pemanggilan dan penggunaan data ke *visual studio code*. Setiap citra mulut disimpan ke dalam folder yang berbeda sesuai dengan label masing-masing kelas dataset, yang tertera pada nama folder. Mesin akan membaca dataset di awal proses, baik citra mulut maupun nama kelas yang tertera pada tiap folder.

Setelah melakukan pembacaan dataset, pada tahap ini juga dilakukan visualisasi sampel gambar dari setiap kategori, menghitung rata-rata resolusi citra, dan melihat distribusi data per kelas. Hal ini bertujuan untuk memahami karakteristik dataset secara lebih mendalam sebelum melanjutkan ke tahap pemrosesan berikutnya.

3.4.2. Scrub

Pada tahap ini, dilakukan pembersihan dan preprocessing data agar siap digunakan dalam model CNN. Berikut adalah penjelasan lebih lanjut dari setiap langkah dalam tahap Scrub pada *framework* OSEMN, khususnya untuk proses pembersihan dan *preprocessing* data agar siap digunakan dalam model *machine learning*.

a. Melakukan Augmentasi Data

Augmentasi data merupakan teknik yang digunakan untuk memperluas variasi dalam dataset dengan tujuan untuk meningkatkan generalisasi model. Teknik augmentasi yang digunakan meliputi:

- 1) *Resizing*: Menyesuaikan ukuran gambar pada dataset agar konsisten dengan input yang dibutuhkan oleh model.
- 2) *Rotasi*: Memutar gambar pada berbagai sudut untuk menciptakan variasi dalam orientasi objek.
- 3) *Flip*: Membalik gambar secara horizontal atau vertikal untuk mensimulasikan perbedaan sudut pandang.
- 4) *Zoom*: Memperbesar atau memperkecil gambar untuk menambah variasi skala objek.
- 5) Penyesuaian Kontras dan Kecerahan: Menambah atau mengurangi kontras dan kecerahan gambar untuk mensimulasikan kondisi pencahayaan yang berbeda.

Dengan melakukan augmentasi, dataset menjadi lebih beragam tanpa perlu menambah jumlah data secara manual, sehingga membantu mencegah *overfitting* pada model.

b. Melakukan Pemisahan Dataset

Pemisahan dataset dilakukan untuk membagi data menjadi beberapa kelompok dengan tujuan yang berbeda, yaitu data *training*, data *validation*, dan data *testing*. Dalam penelitian ini, digunakan beberapa format pembagian, yaitu 80%:10%:10%, 70%:15%:15%, dan 60%:20%:20%.

- 1) *Data Training*: Data ini digunakan untuk melatih model. Ukurannya yang lebih besar dibandingkan set lainnya bertujuan agar model dapat belajar pola dari data sebanyak mungkin. Dalam pembagian yang digunakan, porsi data *training* adalah 80%, 70%, atau 60%, tergantung pada format pembagian yang dipilih.
- 2) *Data Validation*: Data ini digunakan untuk memantau performa model selama proses pelatihan dan menghindari *overfitting*. *Data validation* tidak digunakan untuk melatih model, tetapi untuk mengevaluasi seberapa baik model dapat memprediksi data yang belum pernah dilihat sebelumnya. Porsi data *validation* dalam pembagian yang digunakan adalah 10%, 15%, atau 20%.
- 3) *Data Testing*: *Data testing* digunakan setelah model selesai dilatih. Tujuannya adalah untuk mengevaluasi performa akhir model pada data yang benar-benar baru dan tidak terlibat dalam pelatihan atau validasi. Porsi data *testing* dalam pembagian yang digunakan adalah 10%, 15%, atau 20%.

c. Melakukan Normalisasi Data

Langkah terakhir adalah normalisasi, yang bertujuan untuk memastikan bahwa nilai piksel gambar berada dalam rentang yang sesuai untuk model *deep learning*. Pada umumnya, nilai piksel gambar berada dalam rentang 0 hingga 255. Namun, model *deep learning* bekerja lebih efisien jika nilai-nilai ini dinormalisasi ke dalam rentang $[0, 1]$ atau $[-1, 1]$, tergantung arsitektur modelnya. Normalisasi ini menghindari nilai piksel yang terlalu tinggi atau terlalu rendah yang dapat menyebabkan masalah pada proses pelatihan, seperti memperlambat konvergensi atau membuat model sulit belajar.

3.4.3. Explore

Tahap *Explore* digunakan untuk memahami lebih dalam karakteristik data sebelum digunakan dalam pemodelan *deep learning*. Proses ini melibatkan analisis visual dan statistik deskriptif.

a. Distribusi data per kelas penyakit.

Salah satu langkah pertama dalam eksplorasi data adalah menganalisis distribusi data per kelas. Ini sangat penting, terutama ketika bekerja dengan data klasifikasi seperti dalam pengklasifikasian penyakit mulut. Distribusi yang tidak seimbang antar kelas dapat menyebabkan model condong untuk lebih baik memprediksi kelas mayoritas, sementara mengabaikan kelas minoritas. Untuk itu, memahami jumlah sampel per kelas dapat memberikan gambaran mengenai potensi bias dalam dataset.

Implementasi pada tahap ini adalah melakukan pengecekan distribusi data dengan divisualisasikan menggunakan grafik batang (*bar plot*).

b. Visualisasi sampel gambar dari setiap kategori.

Selain menganalisis jumlah data, penting juga untuk melihat visualisasi contoh gambar dari setiap kategori atau kelas penyakit. Ini membantu peneliti untuk memahami karakteristik visual yang mungkin berbeda antar kelas. Gambar-gambar ini bisa diperiksa secara manual untuk melihat pola umum, kesamaan, atau perbedaan yang jelas antar kategori, yang nantinya bisa mempengaruhi strategi preprocessing atau augmentasi data.

c. Eksplorasi fitur utama dari arsitektur model yang digunakan

Selain memahami data, penting juga untuk mengeksplorasi fitur utama dari arsitektur model deep learning yang digunakan, yaitu ConvNeXt Tiny, MobileNet V3 Large, dan EfficientNet B0. Eksplorasi ini bertujuan untuk memahami fitur utama dari masing-masing model.

3.4.4. Model

Pada tahap ini, tiga arsitektur *convolutional neural network* (CNN) yang akan diuji adalah ConvNeXt Tiny, MobileNet V3 Large, dan EfficientNet B0. Proses modeling mencakup proses *training* dan evaluasi model.

a. *Training* model.

Tahap ini merupakan inti dari proses pembuatan model, di mana data gambar yang telah melalui preprocessing dan augmentasi siap digunakan untuk melatih model. Proses *training* bertujuan untuk mempelajari pola visual dalam gambar agar model dapat melakukan prediksi yang akurat pada data baru.

1) Pelatihan Model

- a) Data *training* digunakan untuk melatih model, sementara data *validation* memonitor performa model guna mencegah *overfitting*.
- b) Model dilatih menggunakan algoritma backpropagation untuk memperbarui bobot-bobot dalam jaringan selama beberapa *epoch* hingga mencapai *convergence*, yaitu saat performa model stabil pada data latih dan validasi.

2) *Transfer learning*

Menggunakan dua skenario, yaitu sebagai berikut.

- a) Tanpa *fine-tuning*: Hanya lapisan akhir model yang disesuaikan untuk klasifikasi. Lapisan pralatih (*pre-trained layers*) dibiarkan tetap beku (*frozen*) sehingga bobot-bobotnya tidak mengalami perubahan selama pelatihan. Pendekatan ini lebih efisien secara komputasi dan cocok jika dataset baru memiliki karakteristik yang mirip dengan dataset pralatih.
- b) Dengan *fine-tuning*: Seluruh lapisan, termasuk lapisan pralatih, dioptimalkan ulang untuk meningkatkan performa model pada dataset baru. Proses ini memungkinkan model untuk lebih beradaptasi dengan fitur spesifik dari dataset penyakit mulut, sehingga dapat menghasilkan prediksi yang lebih akurat meskipun memerlukan waktu pelatihan lebih lama dan sumber daya komputasi yang lebih besar.

3) Optimalisasi *hyperparameter*

Proses optimalisasi *hyperparameter* sangat penting untuk memastikan model berperforma optimal. *Hyperparameter* yang digunakan pada proses *training* dapat dilihat pada Tabel 9.

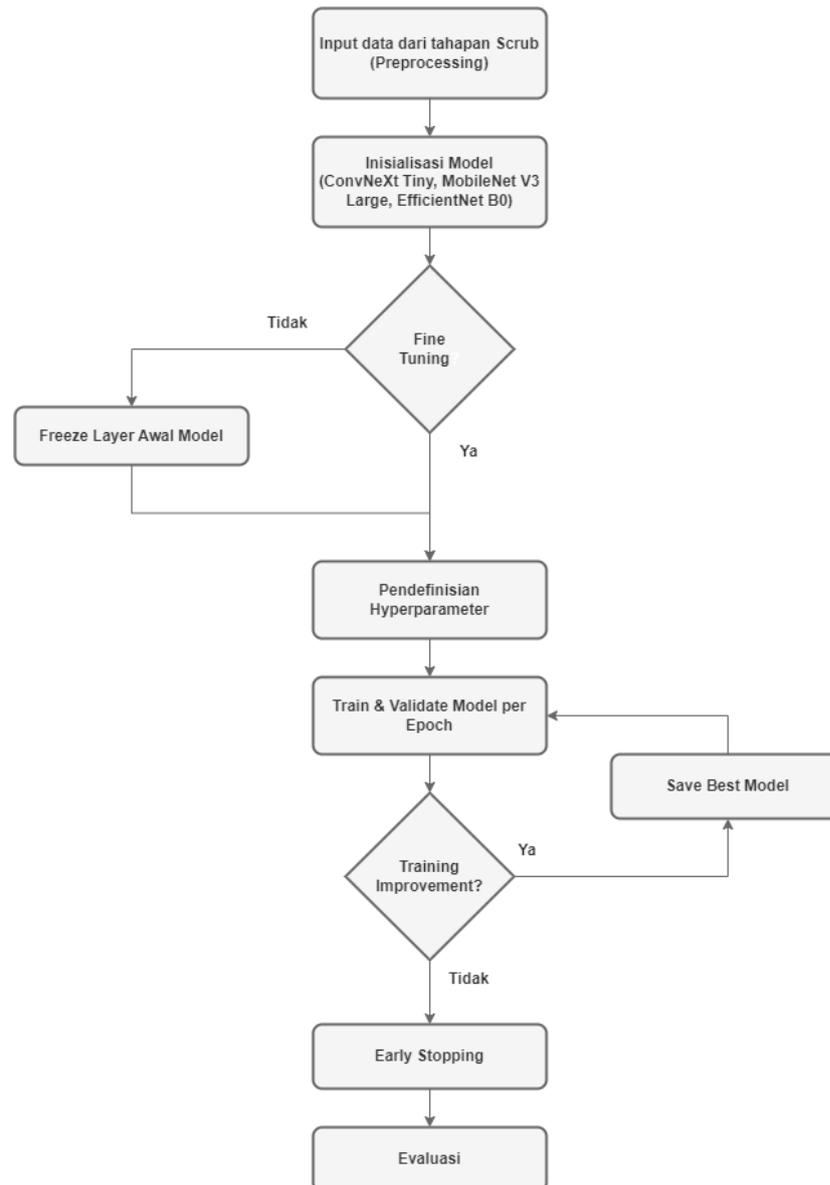
Tabel 9. *Hyperparameter training model*

<i>Hyperparameter</i>	Nilai
<i>Epoch</i>	20
<i>Batch-size</i>	32
<i>Optimizer</i>	AdamW
<i>Learning-rate</i>	0,0001
<i>Weight Decay</i>	0,001

b. Evaluasi performa model

Pada tahap evaluasi performa model, digunakan *classification report* dan *confusion matrix* untuk menilai kinerja model. *Classification report* memberikan metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score*, yang membantu memahami bagaimana model menangani prediksi di setiap kelas, terutama pada dataset yang mungkin tidak seimbang. *Confusion matrix* digunakan untuk memvisualisasikan hasil prediksi dengan melihat jumlah *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Dengan menggunakan kedua alat ini, dapat dilakukan analisis mendalam terhadap pola kesalahan dan kinerja model, sehingga memberikan wawasan apakah model perlu diperbaiki atau dioptimalkan lebih lanjut.

Dengan seluruh tahapan yang telah dijelaskan, mulai dari proses pelatihan model, penerapan *transfer learning*, optimalisasi *hyperparameter*, hingga evaluasi performa model, maka dapat digambarkan secara keseluruhan alur kerja model dalam bentuk diagram alir seperti pada Gambar 12 berikut.



Gambar 12. Diagram alir proses *training* dan evaluasi.

3.4.5. iNterpret

Tahap ini bertujuan untuk mengevaluasi dan membandingkan kinerja model ConvNeXt, MobileNet V3, dan EfficientNet dalam klasifikasi penyakit mulut. Analisis dilakukan menggunakan data uji untuk menilai akurasi dan efisiensi dari masing-masing model.

Hasil evaluasi menunjukkan performa tiap model dalam aspek akurasi (rasio prediksi benar terhadap total prediksi) dan kecepatan prediksi. Model yang dipilih tidak hanya didasarkan pada akurasi tertinggi, tetapi juga pada efisiensi, seperti waktu inferensi yang cepat dan hemat sumber daya komputasi.

Kesimpulan dari hasil ini memberikan rekomendasi mengenai model terbaik untuk klasifikasi penyakit mulut, dengan kriteria utama yaitu akurasi tinggi, kecepatan prediksi optimal, dan efisiensi dalam penggunaan sumber daya komputasi

V. SIMPULAN DAN SARAN

5.1. Simpulan

Simpulan yang didapatkan dari penelitian yang telah dilakukan adalah sebagai berikut:

1. Penelitian ini berhasil mengimplementasikan dan mengoptimalkan model CNN dengan arsitektur ConvNeXt Tiny, MobileNetV3 Large, dan EfficientNet B0 untuk mendeteksi penyakit mulut. Proses optimasi dilakukan melalui teknik augmentasi data dan *fine-tuning*, yang secara signifikan meningkatkan performa deteksi pada ketiga arsitektur. Hasil akhir menunjukkan bahwa ketiga model mencapai tingkat akurasi yang tinggi setelah optimasi, yaitu masing-masing sebesar 96,964% untuk ConvNeXt Tiny, 96,439% untuk MobileNetV3 Large, dan 98,214% untuk EfficientNet B0.
2. Berdasarkan perbandingan performa masing-masing model, EfficientNet B0 menunjukkan keseimbangan terbaik antara akurasi dan efisiensi sumber daya komputasi, dengan akurasi tertinggi dan waktu inferensi yang cepat serta penggunaan parameter dan FLOPs yang relatif rendah. MobileNetV3 Large memiliki efisiensi waktu *training* dan *testing* yang paling baik, menjadikannya ideal untuk perangkat dengan keterbatasan komputasi, meskipun akurasinya sedikit lebih rendah. ConvNeXt Tiny menghasilkan akurasi yang tinggi dan stabil, namun membutuhkan sumber daya komputasi yang lebih besar, sehingga lebih cocok digunakan di lingkungan dengan kapasitas komputasi tinggi.
3. EfficientNet B0 merupakan arsitektur CNN yang paling efisien dalam penggunaan sumber daya komputasi tanpa mengorbankan akurasi, menjadikannya pilihan yang paling tepat untuk klasifikasi penyakit mulut,

terutama dalam kondisi atau lingkungan dengan keterbatasan perangkat keras dan kebutuhan akan performa yang seimbang.

5.2. Saran

Saran yang diberikan untuk penelitian selanjutnya adalah sebagai berikut:

1. Menambahkan lebih banyak variasi kelas penyakit mulut serta memperluas dataset agar meningkatkan keberagaman data, yang dapat membantu model dalam menggeneralisasi lebih baik dan meningkatkan akurasi prediksi.
2. Melakukan uji coba menggunakan arsitektur lain, seperti Xception, ResNet, DenseNet, atau model *neural network* terkini seperti Vision Transformer (ViT) dan Swin Transformer untuk mengevaluasi performa model secara komprehensif.
3. Mengembangkan sistem yang berfungsi sebagai platform implementasi model yang telah dikembangkan agar dapat digunakan secara nyata dalam klasifikasi penyakit mulut, misalnya dalam bentuk aplikasi berbasis web atau mobile yang ramah pengguna.

DAFTAR PUSTAKA

- Ahmada, R. A., Ramadhanti, D. S., Surboyo, M. D. C., Parmadiati, A. E., Ernawati, D. S., Mahdani, F. Y., & Ayuningtyas, N. F. (2024). Analysis on the Knowledge about Oral Potentially Malignant Disorders of a Male Motorcycle Community in Indonesia. *Journal of Health and Allied Sciences NU*, 14(03), 340–345. <https://doi.org/10.1055/s-0043-1772681>
- Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). *Optuna: A Next-generation Hyperparameter Optimization Framework*. <http://arxiv.org/abs/1907.10902>
- Amazon Web Services. (2024). *What is Hyperparameter Tuning?* <https://aws.amazon.com/what-is/hyperparameter-tuning/>. Diakses pada 30 Oktober 2024.
- Awaji, B., Senan, E. M., Olayah, F., Alshari, E. A., Alsulami, M., Abosaq, H. A., Alqahtani, J., & Janrao, P. (2023). Hybrid Techniques of Facial Feature Image Analysis for Early Detection of Autism Spectrum Disorder Based on Combined CNN Features. *Diagnostics*, 13(18). <https://doi.org/10.3390/diagnostics13182948>
- Carrillo-Perez, F., Pecho, O. E., Morales, J. C., Paravina, R. D., Della Bona, A., Ghinea, R., Pulgar, R., Pérez, M. del M., & Herrera, L. J. (2022). Applications of artificial intelligence in dentistry: A comprehensive review. *Journal of Esthetic and Restorative Dentistry*, 34(1), 259–280. <https://doi.org/10.1111/jerd.12844>
- Chen, J., Kao, S.-H., He, H., Zhuo, W., Wen, S., Lee, C.-H., & Chan, S.-H. G. (2023). *Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks*. <https://github.com>.
- Data Science PM. (2024, March 31). *OSEMN Data Science Life Cycle*. <https://www.datascience-pm.com/osemn/>. <https://www.datascience-pm.com/osemn/>. Diakses pada 28 Oktober 2024.
- Devindi, G. A. I., Dissanayake, D. M. D. R., Liyanage, S. N., Francis, F. B. A. H., Pavithya, M. B. D., Piyarathne, N. S., Hettiarachchi, P. V. K. S., Rasnayaka, R. M. S. G. K., Jayasinghe, R. D., Ragel, R. G., & Nawinne, I. (2024). Multimodal Deep Convolutional Neural Network Pipeline for AI-Assisted Early Detection of Oral Cancer. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2024.3454338>

- Ding, H., Wu, J., Zhao, W., Matinlinna, J. P., Burrow, M. F., & Tsoi, J. K. H. (2023). Artificial intelligence in dentistry—A review. In *Frontiers in Dental Medicine* (Vol. 4). Frontiers Media S.A.
<https://doi.org/10.3389/fdmed.2023.1085251>
- Fawaz, P., Sayegh, P. El, & Vannet, B. Vande. (2023). What is the current state of artificial intelligence applications in dentistry and orthodontics? *Journal of Stomatology, Oral and Maxillofacial Surgery*, *124*(5), 101524.
<https://doi.org/10.1016/j.jormas.2023.101524>
- Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D. M., Piñeros, M., Znaor, A., & Bray, F. (2019). Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. In *International Journal of Cancer* (Vol. 144, Issue 8, pp. 1941–1953). Wiley-Liss Inc.
<https://doi.org/10.1002/ijc.31937>
- Ferlay, J., Ervik, M., Lam, F., Laversanne, M., Colombet, M., Mery, L., Piñeros, M., Znaor, A., Soerjomataram, I., & Bray, F. (2024, September 22). *Global Cancer Observatory: Cancer Today*.
<https://gco.iarc.who.int/Media/GlobOCAn/Factsheets/Cancers/1-Lip-Oral-Cavity-Fact-Sheet.Pdf>
- Frazier, P. I. (2018). *A Tutorial on Bayesian Optimization*.
<http://arxiv.org/abs/1807.02811>
- Greenspaneditors, S. J. (2020). *Textbook of Oral Cancer: Prevention, Diagnosis and Management*. <https://doi.org/https://doi.org/10.1007/978-3-030-32316-5>
- Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V., & Adam, H. (2019). *Searching for MobileNetV3*. <http://arxiv.org/abs/1905.02244>
- Inchingolo, F., Santacroce, L., Ballini, A., Topi, S., Dipalma, G., Haxhirexha, K., Bottalico, L., & Charitos, I. A. (2020). Oral cancer: A historical review. In *International Journal of Environmental Research and Public Health* (Vol. 17, Issue 9). MDPI AG. <https://doi.org/10.3390/ijerph17093168>
- Institute for Health Metrics and Evaluation (IHME). (2020). *Global Burden of Disease Collaborative Network Global burden of disease study 2019 (GBD 2019) results*.
- Jeyaraj, P. R., & Samuel Nadar, E. R. (2019). Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm. *Journal of Cancer Research and Clinical Oncology*, *145*(4), 829–837. <https://doi.org/10.1007/s00432-018-02834-7>
- Karacan, M. H., & Yucebas, S. C. (2022). A Deep Learning Model with Attention Mechanism for Dental Image Segmentation. *2022 International Congress on*

Human-Computer Interaction, Optimization and Robotic Applications (HORA), 1–4. <https://doi.org/10.1109/HORA55278.2022.9800072>

- Kasnanda Bintang, Y., & Imaduddin, H. (2024). Pengembangan Model Deep Learning untuk Deteksi Retinopati Diabetik Menggunakan Metode *Transfer learning*. *Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika*, 9(3), 1442–1455. <https://doi.org/10.29100/jipi.v9i3.5588>
- Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 173, pp. 24–49). Elsevier B.V. <https://doi.org/10.1016/j.isprsjprs.2020.12.010>
- Krstinić, D., Braović, M., Šerić, L., & Božić-Štulić, D. (2020). *Multi-label Classifier Performance Evaluation with Confusion Matrix*. 01–14. <https://doi.org/10.5121/csit.2020.100801>
- Kumar, K., Khandpur, M., Khandpur, S., Mehrotra, D., Chandra Tiwari, S., & Kumar, S. (2021). Quality of life among Oral Potentially Malignant Disorder (OPMD) patients: A prospective study. *Journal of Oral Biology and Craniofacial Research*, 11(1), 88–91. <https://doi.org/10.1016/j.jobcr.2020.11.009>
- Li, S., & Deng, W. (2019). Blended Emotion in-the-Wild: Multi-label Facial Expression Recognition Using Crowdsourced Annotations and Deep LOCALity Feature Learning. *International Journal of Computer Vision*, 127(6–7), 884–906. <https://doi.org/10.1007/s11263-018-1131-1>
- Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). *A ConvNet for the 2020s*. <http://arxiv.org/abs/2201.03545>
- Llusi, R., Yacoubi, S. El, Fontaine, A., & Lupera, P. (2021). Comparison between Adam, AdaMax and Adam W optimizers to implement a Weather Forecast based on Neural Networks for the Andean city of Quito. *2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM)*, 1–6. <https://doi.org/10.1109/ETCM53643.2021.9590681>
- Loshchilov, I., & Hutter, F. (2017). *Decoupled Weight Decay Regularization*. <http://arxiv.org/abs/1711.05101>
- Mason, H., & Wiggins, C. (2010, September 25). *A Taxonomy of Data Science*. <https://web.archive.org/web/20160220042455/dataists.com/2010/09/a-taxonomy-of-data-science/>. Diakses pada 2 Oktober 2024.
- MathWorks. (2021). *Introducing Deep Learning with MATLAB*. <https://www.mathworks.com/campaigns/offers/next/deep-learning-ebook.html>. Diakses pada 1 Oktober 2024.
- Maxwell, A. E., Warner, T. A., & Guillén, L. A. (2021). *Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—*

- part 1: Literature review. In *Remote Sensing* (Vol. 13, Issue 13). MDPI AG. <https://doi.org/10.3390/rs13132450>
- Mumuni, A., & Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. In *Array* (Vol. 16). Elsevier B.V. <https://doi.org/10.1016/j.array.2022.100258>
- NVIDIA Corporation & affiliates. (2023). *Optimizing Linear/Fully-Connected Layers User's Guide | NVIDIA Docs*. <https://docs.nvidia.com/deeplearning/performance/pdf/Optimizing-Linear-Fully-Connected-Layers-User-Guide.pdf>
- Oroh, E., & Lubis, C. (2023). Klasifikasi Penyakit Gigi Karies Dan Kalkulus Menggunakan Convolutional Neural Network. *Nusantara Journal of Multidisciplinary Science*, 1(4). <https://jurnal.intekom.id/index.php/njms>
- Panetta, K., Rajendran, R., Ramesh, A., Rao, S., & Aгаian, S. (2022). Tufts Dental Database: A Multimodal Panoramic X-Ray Dataset for Benchmarking Diagnostic Systems. *IEEE Journal of Biomedical and Health Informatics*, 26(4), 1650–1659. <https://doi.org/10.1109/JBHI.2021.3117575>
- Pangestu, R. A., Rahmat, B., & Anggraeny, F. T. (2020). Implementasi Algoritma CNN untuk Klasifikasi Citra Lahan dan Perhitungan Luas. *Jurnal Informatika Dan Sistem Informasi (JIFoSI)*, 1.
- Peng, Y., & Nagata, M. H. (2020). An empirical overview of nonlinearity and *overfitting* in machine learning using COVID-19 data. *Chaos, Solitons and Fractals*, 139. <https://doi.org/10.1016/j.chaos.2020.110055>
- Peres, M. A., Macpherson, L. M. D., Weyant, R. J., Daly, B., Venturelli, R., Mathur, M. R., Listl, S., Celeste, R. K., Guarnizo-Herreño, C. C., Kearns, C., Benzian, H., Allison, P., & Watt, R. G. (2019). Oral diseases: a global public health challenge. *The Lancet*, 394(10194), 249–260. [https://doi.org/10.1016/S0140-6736\(19\)31146-8](https://doi.org/10.1016/S0140-6736(19)31146-8)
- Piyarathne, N. S., Liyanage, S. N., Rasnayaka, R. M. S. G. K., Hettiarachchi, P. V. K. S., Devindi, G. A. I., Francis, F. B. A. H., Dissanayake, D. M. D. R., Ranasinghe, R. A. N. S., Pavithya, M. B. D., Nawinne, I. B., Ragel, R. G., & Jayasinghe, R. D. (2024). A comprehensive dataset of annotated oral cavity images for diagnosis of oral cancer and oral potentially malignant disorders. *Oral Oncology*, 156. <https://doi.org/10.1016/j.oraloncology.2024.106946>
- Rashid, J., Qaisar, B. S., Faheem, M., Akram, A., Amin, R. ul, & Hamid, M. (2023). Mouth and oral disease classification using InceptionResNetV2 method. *Multimedia Tools and Applications*, 83(11), 33903–33921. <https://doi.org/10.1007/s11042-023-16776-x>
- Rochmawati, N., Hidayati, H. B., Yamasari, Y., Peni, H., Tjahyaningtijas, A., Yustanti, W., & Prihanto, A. (2021). Analisa *Learning rate* dan *Batch size*

- Pada Klasifikasi Covid Menggunakan Deep learning dengan Optimizer Adam. *Journal Information Engineering and Educational Technology*, 05.
- Salman, S., & Liu, X. (2019). *Overfitting Mechanism and Avoidance in Deep Neural Networks*. <http://arxiv.org/abs/1901.06566>
- Sarvamangala, D. R., & Kulkarni, R. V. (2022). Convolutional neural networks in medical image understanding: a survey. In *Evolutionary Intelligence* (Vol. 15, Issue 1). Springer Science and Business Media Deutschland GmbH. <https://doi.org/10.1007/s12065-020-00540-3>
- Shorten, C., Khoshgoftaar, T. M., & Furht, B. (2021). Text Data Augmentation for Deep Learning. *Journal of Big Data*, 8(1), 101. <https://doi.org/10.1186/s40537-021-00492-0>
- Tan, M., & Le, Q. V. (2019). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. <http://arxiv.org/abs/1905.11946>
- Vodanović, M., Subašić, M., Milošević, D., & Savić Pavičin, I. (2023). Artificial Intelligence in Medicine and Dentistry. *Acta Stomatologica Croatica*, 57(1), 70–84. <https://doi.org/10.15644/asc57/1/8>
- Warnakulasuriya, S., Kujan, O., Aguirre-Urizar, J. M., Bagan, J. V., González-Moles, M. Á., Kerr, A. R., Lodi, G., Mello, F. W., Monteiro, L., Ogden, G. R., Sloan, P., & Johnson, N. W. (2021). Oral potentially malignant disorders: nomenclature and classification. In *Oral Diseases* (Vol. 27, Issue 8, pp. 1862–1880). John Wiley and Sons Inc. <https://doi.org/10.1111/odi.13704>
- Wibawa, M. S. (2017). Pengaruh Fungsi Aktivasi, Optimisasi dan Jumlah Epoch Terhadap Performa Jaringan Saraf Tiruan. *JURNAL SISTEM DAN INFORMATIKA*.
- World Health Organization. (2022). *Global oral health status report: Towards universal health coverage for oral health by 2030*.
- Xu, Z., Peng, J., Zeng, X., Xu, H., & Chen, Q. (2022). High-Accuracy Oral Squamous Cell Carcinoma Auxiliary Diagnosis System Based on EfficientNet. *Frontiers in Oncology*, 12. <https://doi.org/10.3389/fonc.2022.894978>