

## BAB II

### LANDASAN TEORI

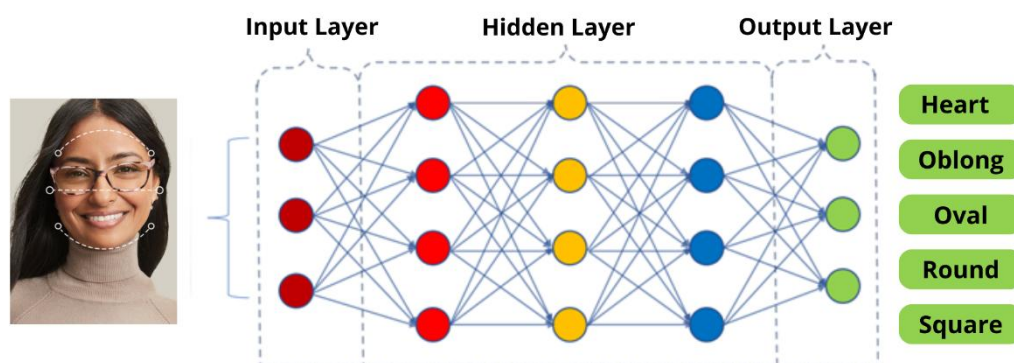
#### 2.1. *Convolutional Neural Network*

Ilmu biologi telah menjadi banyak inspirasi untuk berbagai teknik kecerdasan buatan, salah satunya adalah *Artificial Neural Network* (ANN). Penemuan dari neuron yang dilakukan oleh peneliti Hubel dan Wiesel pada tahun 1959 dengan memberikan kontribusi dengan melakukan eksperimen neurobiologis. Penemuan tersebut menjadi inspirasi awal dalam memodelkan sistem penglihatan yang didasarkan pada fitur sistem saraf pusat (Murinto dkk., 2023).

Pada tahun 1979, peneliti Kunihiko Fukushima mengembangkan model dengan konsep *neocognitron* dengan tujuan untuk mengenali pola sederhana tanpa mempengaruhi distorsi bentuk. Selain itu, tahun 1970 sampai 1986 telah dilakukan sebuah penelitian yang dilakukan oleh Linnainmaa dkk untuk mengembangkan dan mempopulerkan konsep algoritma *backpropagation*, sehingga pada tahun 1989 penelitian yang dilakukan oleh Yann LeCun berhasil menggabungkan konsep *Neocognitron* dan algoritma *backpropagation*. Penelitian tersebut telah menjadi awal mula yang penting bagi perkembangan *Convolutional Neural Network* dan menjadi dasar untuk berbagai arsitektur *deep learning* modern.

*Convolutional Neural Network* (CNN) adalah jenis *neural network* yang dapat mengekstrak fitur dan mampu memproses data spasial seperti gambar dengan cara kerja yang meniru otak manusia dalam mengenali pola visual melalui lapisan-lapisan konvolusi. Tujuan utama dari CNN ini adalah mengekstraksi dan

memahami pola visual untuk klasifikasi, deteksi objek, dan pengenalan pola (Li Z dkk., 2021).



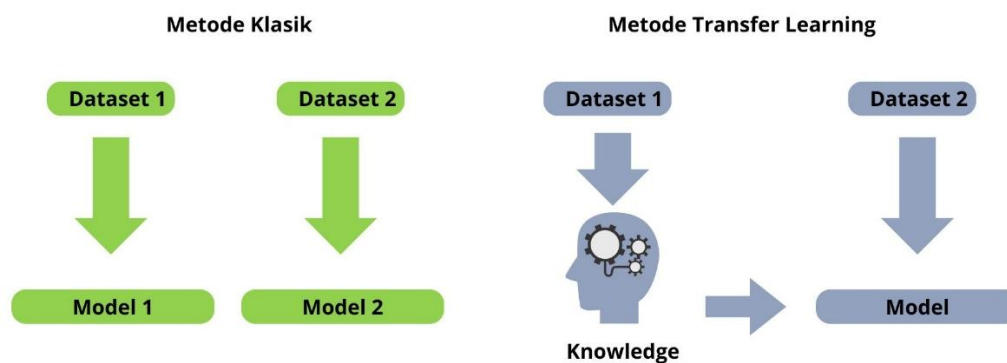
**Gambar 2.1.** Lapisan Layer Arsitektur CNN (Dimodifikasi dari (Krichen, 2023))

Gambar 2.1 menunjukkan lapisan-lapisan pada arsitektur CNN yang digunakan untuk mengklasifikasikan gambar. Input dari *layer* tersebut merupakan sebuah data gambar dalam bentuk PNG atau JPG yang kemudian dimasukkan ke *layer input*. *Layer* tersebut merupakan tempat sistem menerima informasi dari dunia luar dan akan meneruskan ke lapisan tersembunyi atau *hidden layer*. Setelah data masuk ke *hidden layer*, informasi akan diproses dan akan mencari pola penting. Semakin mendalam, semakin kompleks pola yang dapat dipelajari. Setelah data diproses, sistem akan memberikan hasil dengan cara mengklasifikasikan gambar (Krichen, 2023).

## 2.2. Transfer Learning

*Transfer learning* adalah pendekatan model CNN yang telah dilatih pada *dataset* besar kemudian disesuaikan secara menyeluruh untuk berfungsi dengan baik pada jumlah data yang terbatas. Teknik ini mengacu pada situasi dengan pengetahuan yang telah dipelajari dalam satu konteks. Hal ini dapat digunakan

untuk meningkatkan proses optimisasi dalam konteks lain. Ketika *dataset* baru yang lebih kecil digunakan untuk melatih model pra-latih, *transfer learning* biasanya digunakan (Shin dkk., 2016).

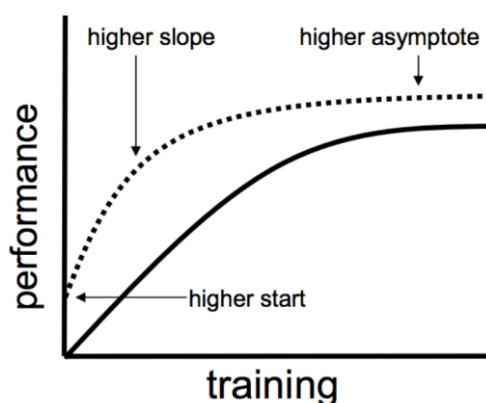


**Gambar 2.2.** Perbandingan metode klasik dan metode *transfer learning*

Perbandingan dari dua metode yang berbeda yaitu metode *learning* secara tradisional dan metode *learning* secara *transfer learning* divisualisasikan pada Gambar 2.2 Secara tradisional, sistem belajar yang dihasilkan dari *sample* atau tugas berbeda dengan hasil yang sesuai berdasarkan bentuk yang sudah dipelajari. Dalam konsep *machine learning*, setiap tugas yang berbeda harus dilatih dengan model yang berbeda. Sehingga, ketika terdapat beberapa karakteristik yang tidak sama, sistem pembelajaran dibutuhkan tanpa memanfaatkan pengetahuan dari model lain. Sedangkan, pada *transfer learning*, sistem belajar dari beberapa bentuk yang akhirnya dapat menghasilkan *knowledge*. Setelah itu, apabila diberi tugas atau *task* baru, sistem akan mengetahui atau mengenal bentuk karena sudah mendapatkan *knowledge* dari bentuk yang sudah dipelajari sebelumnya. Dalam konsep *machine learning*, proses *transfer learning* membuat pelatihan lebih cepat, efisien, dan tetap mencapai hasil yang baik karena memanfaatkan pengalaman yang

telah diperoleh dari tugas awal yang memiliki banyak data, sehingga model tidak perlu belajar dari nol (Hosna dkk., 2022).

Prinsip utama pada *transfer learning* adalah menjaga dan mempertahankan informasi berupa pengetahuan umum yang telah dipelajari dari pelatihan sebelumnya pada lapisan awal dan lapisan tengah. Sementara, lapisan akhir disesuaikan dengan tugas yang relevan untuk menyelesaikan tugas baru. Terdapat dua jenis lapisan ketika menerapkan *transfer learning*, yaitu *frozen layers* dan *trainable layers*. *Frozen layers* merupakan lapisan yang dibekukan selama proses *retraining* dan mempertahankan pengetahuan umum, sehingga model AI dapat memanfaatkan pengetahuan tersebut sebagai fondasi untuk belajar model baru. Sedangkan, *trainable layers* merupakan lapisan yang dilatih ulang selama penyempurnaan dengan cara *fine-tuned*. Sehingga, model dapat menyesuaikan pengetahuannya dengan tugas baru yang bersifat lebih spesifik.



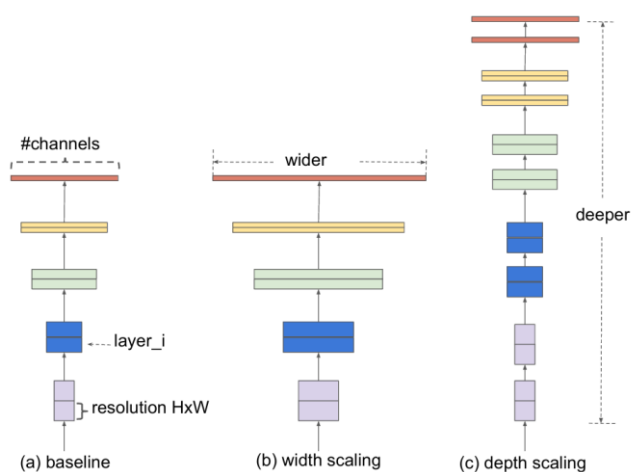
**Gambar 2.3.** Perbandingan performa *transfer learning* (Thiodorus dkk., 2021)

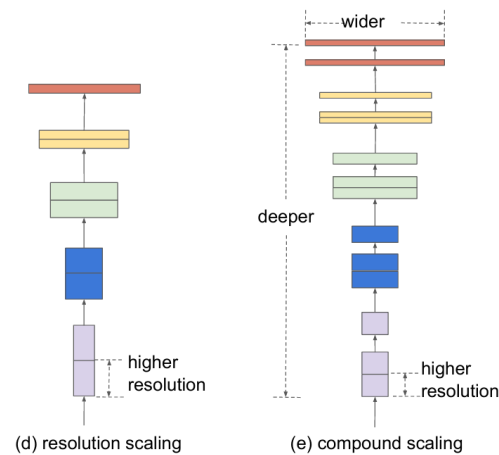
Penggunaan *transfer learning* menjadi sebuah strategi dalam membangun model dengan memanfaatkan model yang sudah belajar dari tugas lain sehingga performa cenderung lebih baik daripada model tanpa menggunakan *transfer*

*learning*, seperti yang ditunjukkan pada Gambar 2.3. Penggunaan *transfer learning* akan membuat model dengan *higher start*, *higher slope*, dan *higher asymptote*. Sehingga model lebih unggul dalam memulai *training*, proses *training* lebih cepat dan efisien, dan performa akhir yang lebih tinggi dibandingkan dengan model tanpa menggunakan *transfer learning* dengan ditunjukkan garis *asymptote* yang lebih tinggi (Thiodorus dkk., 2021).

### 2.3. *EfficientNet*

*EfficientNet* dikemukakan oleh *Google Brain* (Tan & Le, 2020) pada tahun 2019 sebagai salah satu bentuk varian dari model *EfficientNet*. Penelitian tersebut memperkenalkan *EfficientNet* yang dirancang dengan arsitektur yang optimal dengan menggunakan *compound scaling*. Teknik ini dapat melakukan penyesuaian kedalaman, lebar, dan resolusi secara simultan dengan menggunakan satu koefisien skala sehingga dihasilkan *EfficientNet-B0* sebagai model Dasar dan varian B1 sampai B7.





**Gambar 2.4.** Metode *Scaling* pada *EfficientNet* (Tan & Le, 2020)

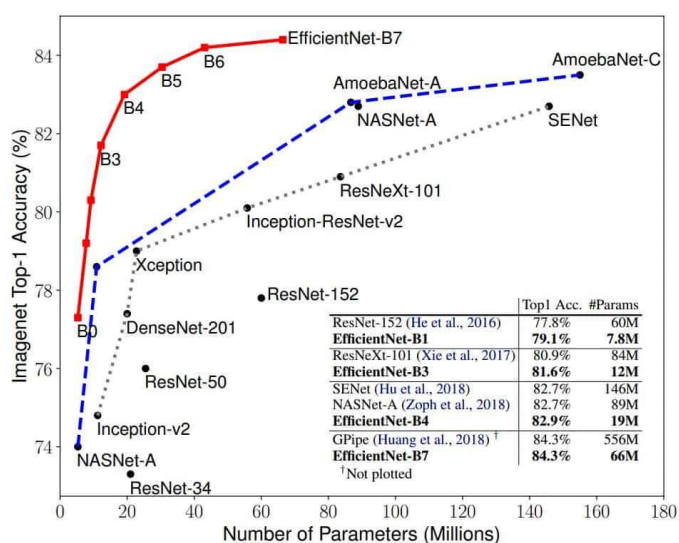
Penelitian yang dilakukan sering kali penskalaan model dilakukan dengan cara meningkatkan satu dimensi antara dari dimensi lebar, kedalaman, maupun secara resolusi. Akan tetapi, penelitian yang dilakukan oleh (Tan & Le, 2020) menunjukkan *compound scaling* bisa diterapkan dengan menggabungkan ketiga dimensi dalam penskalaan *channels* dengan dimensi lebar, kedalaman, dan resolusi, seperti yang ditunjukkan oleh Gambar 2.4.

Pendekatan *depth*, *width*, dan *resolution* menunjukkan sebuah target peningkatan kapasitas komputasi setiap model dinaikan satu tingkat skala nilai, sehingga secara seragam melakukan penskalaan dari ketiga dimensi dengan rasio tetap. Ketika nilai  $\phi$  dinaikan, maka ketiga dimensi akan ikut ditingkatkan secara proporsional dengan hasil ukuran komputasi oleh model yang meningkat sekitar 2 kali lipat. Penskalaan tersebut dilakukan dengan rasio tetap dengan menggunakan persamaan (2.1).

$$\begin{aligned}
 \text{depth: } d &= \alpha^\phi \\
 \text{width: } w &= \beta^\phi \\
 \text{resolution: } r &= \gamma^\phi \\
 \alpha^\phi \beta^\phi \gamma^\phi &\approx 2
 \end{aligned}
 \tag{2.1}$$

Dimana,  
 $\alpha$  = Konstanta *Scaling Depth*,  
 $\beta$  = Konstanta *Scaling Width*,  
 $\gamma$  = Konstanta *Scaling Resolution*,  
 $\varphi$  = *Compound Coefficient*

Akurasi model *deep learning* dapat ditingkatkan dengan cara memperbesar ukuran model dalam segi menambah jumlah *layer* dan parameter. Namun, pada Gambar 2.5 menunjukkan peningkatan performa dari *EfficientNet* ke *Gpipe* pada tahun 2018 menggunakan dataset *ImageNet-1k* yang digunakan sebagai dataset standar *pretraining* CNN (Krizhevsky dkk., 2020) dengan hasil akurasi dari 74,8% menjadi 84,3% dan jumlah parameter dari 6.8M menjadi 66M sehingga dampak komputasi menjadi lebih efisien namun memberikan akurasi yang maksimal.



**Gambar 2.5.** Perbandingan Performa *EfficientNet* (Tan & Le, 2020)

*EfficientNet* sangat relevan dengan *Transfer Learning* karena telah dilatih dengan menggunakan *dataset* besar seperti *ImageNet* yang memungkinkannya untuk mengenali pola visual yang umum. Pengetahuan atau *knowledge* yang didapat oleh model ini kemudian digunakan untuk tugas baru, terutama ketika *dataset* yang digunakan jauh lebih kecil. *EfficientNet* berfungsi sebagai model

*pretrained* yang menyediakan fitur dasar untuk diterapkan pada berbagai jenis klasifikasi gambar. Saat dipadukan dengan *Transfer Learning*, bobot *pretrained EfficientNet* bertindak sebagai dasar yang telah mengenali pola visual umum. Hal ini, yang memudahkan model untuk beradaptasi dengan efektif di domain yang baru, meskipun jumlah contoh sasaran yang tersedia terbatas. Kinerja setiap model *EfficientNet* yang diskalakan dari *baseline EfficientNet-B0* yang sama ditunjukkan dalam Tabel 2.1 dengan #Params adalah jumlah bobot (Tan & Le, 2020).

**Tabel 2.1.** Performa Varian *EfficientNet* (Dataset ImageNet)

<i>Model</i>	<i>Top-1 Acc.</i>	<i>Top-5 Acc.</i>	<i>#Params</i>
EfficientNet-B0	77.1%	93.3%	5.3M
EfficientNet-B1	79.1%	94.4%	7.8M
EfficientNet-B2	80.1%	94.9%	9.2M
EfficientNet-B3	81.6%	95.7%	12M
EfficientNet-B4	82.9%	96.4%	19M
EfficientNet-B5	83.6%	96.7%	30M
EfficientNet-B6	84.0%	96.8%	43M
EfficientNet-B7	84.3%	97.0%	66M

Pemilihan varian *EfficientNet* dapat dilakukan dengan menyesuaikan dengan kondisi yang akan digunakan. Semakin tinggi varian *EfficientNet* berarti akan semakin tinggi nilai parameter yang ditunjukkan dengan jumlah bobot yang dipelajari oleh model sehingga kapasitas model akan semakin besar dan dapat menangkap pola yang lebih kompleks, namun ukuran model yang lebih berat.. Maka, varian *EfficientNet* dengan varian yang rendah lebih cocok dalam melakukan

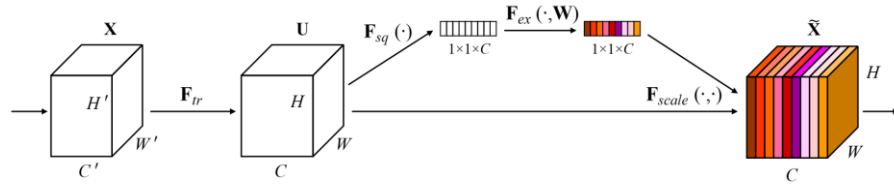
eferensi model *low device* seperti *website*, *mobile*, dan IoT sedangkan varian yang tinggi dapat diimplementasikan pada konteks yang bukan *device user*, tetapi lebih pada studi kasus di lapangan dengan kondisi yang sangat kompleks.

**Tabel 2.2.** Arsitektur *EfficientNet*

<i>Stage</i>	<i>Operator</i>	<i>Resolution</i>	<i>#Channels</i>	<i>#Layers</i>
1	Conv3x3	224x224	32	1
2	MBCConv1, k3x3	112x112	16	1
3	MBCConv6, k3x3	112x112	24	2
4	MBCConv6, k5x5	56x56	40	2
5	MBCConv6, k3x3	28x28	80	3
6	MBCConv6, k5x5	14x14	112	3
7	MBCConv6, k5x5	14x14	192	4
8	MBCConv6, k3x3	7x7	320	1
9	Conv1x1 & Pooling & FC	7x7	1280	1

Arsitektur *EfficientNet* yang ditunjukkan pada Tabel 2.2 dibangun dari serangkaian *Mobile Inverted Bottleneck Convolution (MBCConv)* tersusun secara bertingkat dengan penurunan resolusi dan tambahan jumlah *channel*. Varian *EfficientNet* tidak akan mengubah bentuk arsitekturnya, tetapi diperbesar skalanya menggunakan *compound scaling* dengan Persamaan (2.1). Gambar 2.6 menunjukan visualisasi dari arsitektur *EfficientNet-B0* sebagai *baseline* dari varian *EfficientNet*.





**Gambar 2.7.** Arsitektur *Squeeze and Excitation Attention*

Tahap *squeeze* menjadi tahap pertama dengan blok SE akan menangkap informasi global dari setiap kanal dengan melakukan pemadatan menjadi satu nilai numerik. Penerapan *squeeze* akan menghasilkan sebuah vektor *scalar* untuk setiap *channel* dengan vektor berdimensi  $1 \times 1 \times C$ . Persamaan (2.2) menunjukkan rata-rata seluruh nilai dalam satu kanal.

$$F_{sq}(uc) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W uc(i, j) \quad (2.2)$$

Dimana,

$uc$  = *Feature map* pada *channel* ke- $c$ ,

$H$  = Tinggi *feature map*,

$W$  = Lebar *feature map*,

$I$  = indeks baris,  $i = 1, 2, 3, \dots, H$

$J$  = Indeks kolom,  $j = 1, 2, 3, \dots, W$

Tahap *excitation* merupakan tahapan kedua dari tahap SE dengan nilai  $n$  didapatkan dari tahap *squeeze*, kemudian diteruskan ke jaringan dua lapis *feed forward neural network*. Jaringan tersebut akan menangkap pola dependensi dan hubungan kompleks. Sehingga, *output* dari tahap *excitation* adalah nilai  $n$  yang merupakan vektor dengan ukuran yang sama pada *output* tahap *squeeze*, tetapi memiliki bobot yang telah dipelajari untuk setiap kanal. Perhitungan nilai *excitation* mengacu persamaan (2.3).

$$F_{ex}(z, W) = Sc = \sigma(g(z, W)) = \sigma(W_2 \sigma(W_1 z)) \quad (2.3)$$

Dimana,  
 $z$  = vektor hasil *squeeze*,  
 $W1; W2$  = *weight fully connected*,  
 $\delta$  = Fungsi aktivasi ReLU,  
 $\sigma$  = Fungsi *sigmoid*  
 $Sc$  = Vektor bobot *channel* hasil *excitation*

Tahap terakhir yaitu, *scale* yang melibatkan penggunaan bobot "n" yang diperoleh dari tahap *excitation* yang bertujuan untuk menskalakan setiap kanal dari tensor *input*. Dengan menerapkan bobot adaptif ini pada *channel* asli, blok SE secara efektif menekan informasi yang kurang penting. Persamaan (2.4) menunjukkan pemberian bobot untuk setiap *channel*.

$$F_{scale}(uc, sc) = sc uc \quad (2.4)$$

Dimana,  
 $uc$  = *Feature map* pada *channel* ke-c,  
 $Sc$  = Vektor bobot *channel* hasil *excitation*  
 $F_{scale}$  = *Feature map* hasil kalibrasi *channel*

*Convolutional Block Attention Module* (CBAM) adalah salah satu jenis *attention mechanism* yang ringan dan adaptif, yang dirancang untuk meningkatkan representasi *feature map* dalam CNN. Mekanisme ini bekerja dengan menerapkan *attention* secara bertahap pada aspek *channel* dan *spatial*, sehingga CBAM disebut sebagai *dual block attention mechanism* dengan hadirnya dua blok yang terdiri dari *Channel Attention* dan *Spatial Attention*. Melalui pendekatan ini, CBAM membantu model memilih fitur penting untuk diperkuat dengan karakteristik *what to emphasize* dan lokasi mana yang relevan untuk diperhatikan (Woo dkk., 2018). Penerapan *Channel Attention* dilakukan dari nilai rata-rata dan nilai terbesar yang dihubungkan dengan MLP seperti yang ditunjukkan pada Persamaan (2.5).

$$Mc(F) = \sigma(MLP(AvgPool(F)) + (MaxPool(F)) \quad (2.5)$$

Dimana:

$F$  = *feature map input*

$M_c$  = *Output channel attention*

*AvgPool* = rata-rata dari semua *channel*

*MaxPool* = nilai maksimum dari semua *channel*

Fitur-fitur input sudah diperkuat oleh *Channel attention* yang dilakukan oleh *EfficientNet* dan *Channel Attention Block* dari CBAM kemudian selanjutnya dibutuhkan *attention spatial attention* yang melakukan *pooling* yaitu *MaxPool* untuk mengambil nilai terbesar untuk setiap *pixel* dan *AvgPool* untuk mengambil nilai rata-rata untuk semua piksel menggunakan Persamaan (2.6). *Spatial attention* akan memberikan atensi lebih pada fitur yang penting.

$$M_s(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \quad (2.6)$$

Dimana:

$F$  = *feature map input*

$M_s$  = *Output spatial attention*

*AvgPool* = rata-rata dari semua *channel*

*MaxPool* = nilai maksimum dari semua *channel*

Setelah dilakukan operasi *pooling*, *pooling map* tersebut dilakukan *concatenate* atau penggabungan yang dilanjutkan dengan melewati lapisan konvolusi untuk menangkap pola dan hubungan antar piksel. Hasil konvolusi diteruskan ke lapisan aktivasi *sigmoid* untuk memetakan nilai ke rentang 0 dan 1. Sehingga, fitur akhir dari hasil konvolusi akan memperkuat area penting dan tidak penting (Liao & Wang, 2025).

## 2.5. Rule-Based Approach

*Rule Based Approach* digunakan untuk membuat keputusan secara otomatis menggunakan kumpulan aturan *IF-Then* secara eksplisit. Metode ini membutuhkan kumpulan-kumpulan fakta yang berisi data yang akan dikaitkan dengan aturan

logika, *If* sebagai kondisi dan *Then* sebagai aksi (Masri dkk., 2019). Dalam konteks prediksi klasifikasi gambar, *Rule Based Approach* dilakukan setelah model telah dilatih, sehingga menghasilkan rekomendasi berdasarkan *output* prediksi dari model.

## 2.6. Evaluasi Model

Evaluasi pada pengujian model dilakukan menggunakan data uji melalui perhitungan akurasi, presisi, *recall*, dan *F1-Score*. Pada Gambar 2.8 merupakan *confussion matrix* yang menunjukkan *True Positive*, *False Positive*, *False Negative*, *True Negative*.

		Actual Values	
		Positive	Negative
Predicted Values	Positive	TP	FP
	Negative	FN	TN

Gambar 2.8. *Confussion Matrix*

Persamaan TP (*True Positive*) merupakan objek positif yang diprediksi dengan benar, TN (*True Negative*) menunjukkan objek negatif yang diprediksi atau terklasifikasi dengan benar, persamaan FP (*False Positive*) merupakan objek negatif yang diprediksi secara benar, dan FN (*False Negative*) menunjukkan objek negatif yang diprediksi atau terklasifikasi secara salah. Hasil perhitungan *confussion matrix* memperoleh dua jenis nilai yang berbeda, yaitu *Macro-Average* sebagai rata-rata

seluruh metrik tiap kelas, dan *Weighted-Average* yang menghitung rata-rata nilai TP berdasarkan jumlah data tiap kelas (Jude Chukwura Obi, 2023).

a. *Accuracy*

*Accuracy* atau akurasi adalah persentase sampel yang diprediksi dengan benar. Metrik ini mengukur seberapa banyak prediksi yang benar dibandingkan dengan seluruh data. Definisi matematis pada metrik akurasi dinyatakan dalam persamaan (2.6).

$$\mathbf{Accuracy} = \frac{\mathbf{TP} + \mathbf{TN}}{\mathbf{Total\ Prediction}} \quad (2.7)$$

Dimana:

*TP* = *True Positive*;

*TN* = *True Negative*

Tingkat akurasi menunjukkan kebalikan dari tingkat kesalahan dan jumlah klasifikasi yang benar dari seluruh kumpulan data uji atau fraksi prediksi yang benar dari suatu klasifikasi dari seluruh kumpulan data uji. Kedua tingkat ini memberikan informasi yang sama tentang kekuatan atau kelemahan suatu klasifikasi (Miller dkk, 2024).

b. *Precision*

*Precision* adalah metrik yang menunjukkan persentase sampel yang diprediksi sebagai positif yang sebenarnya positif. Metrik ini akan mengevaluasi tingkat ketepatan model dalam melakukan klasifikasi kelas positif. (Jude Chukwura Obi, 2023). *Precision* akan menunjukkan rasio antara semua positif benar yang diamati

bersama dengan beberapa negatif benar yang salah dianggap positif (FP). (Miller dkk., 2024). Precision didefinisikan secara matematis melalui Persamaan (2.7).

$$\mathbf{Precision} = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FP}} \quad (2.8)$$

Dimana;

$TP = True\ Positive$

$FP = False\ Positive$

### c. *Recall*

*Recall* adalah persentase sampel positif yang diprediksi dengan benar. Nilai *recall* yang tinggi menunjukkan bahwa jumlah TP lebih besar dibandingkan dengan FN, sedangkan nilai *recall* yang rendah menunjukkan bahwa kesalahan prediksi terhadap data positif lebih dominan. (Miller dkk., 2024). Formulasi *Recall* dinyatakan dalam Persamaan (2.8).

$$\mathbf{Recall} = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FN}} \quad (2.9)$$

Dimana;

$TP = True\ Positive;$

$FN = False\ Negative$

### d. *F1-Score*

*F1-Score* adalah rata-rata *harmonic* dari *precision* dan *recall*. *F1-Score* yang tinggi menandakan model dapat menyeimbangkan antara menangkap TP dan mampu menghindari terjadinya FP. (Miller dkk., 2024) Nilai *F1-Score* dihitung berdasarkan *precision* dan *recall* berdasarkan persamaan (2.9).

$$F1 = 2 \times \frac{\textit{Precision} \times \textit{Recall}}{\textit{Precision} + \textit{Recall}} \quad (2.10)$$

## 2.7. Bentuk Wajah dan Bingkai Kacamata

Secara antropometri, bentuk wajah menurut (Sforza dkk., 2009) adalah konfigurasi geometris khusus yang ditentukan oleh struktur tulang kraniofasial dan distribusi jaringan lunak. Bentuk wajah diklasifikasikan secara sistematis berdasarkan perbandingan antara lebar dahi, tulang pipi (*bizygomatic*), garis rahang (*bigonial*), dan tinggi total wajah. Berikut penjelasan beberapa bentuk wajah.

### 1. Wajah Hati

Bentuk wajah hati ditandai dengan dahi yang lebar tulang pipi yang tinggi, dan dagu yang sempit dan runcing, menyerupai segitiga terbalik. Bagian wajah yang paling lebar berada di di dahi dan bagian tersempit ada pada dagu dengan garis rahang yang tegas (Budianto dkk., 2023). Dalam memilih bingkai kacamata direkomendasikan untuk menggunakan kacamata yang dapat menyeimbangkan fitur wajah. Dalam konteks ini, (Optik PEEP, 2026) menyarankan untuk menghindari bingkai kacamata yang terlalu besar dan memilih bingkai yang menambah lebar bagian bawah wajah. Rekomendasi kacamata berdasarkan bentuk wajah hati terdiri dari *Frame* bulat, *Frame oval*, *Frame Aviator*, *Frame Cat Eye* (Optik PEEP, 2026).

### 2. Wajah Persegi

Bentuk wajah persegi ditandai dengan garis rahang yang tegas dan dahi yang memiliki lebar yang sama sehingga memberikan kesan pada wajah dengan sudut yang khas seperti struktur tulang yang jelas (Budianto dkk., 2023). Sehingga bingkai kacamata yang cocok untuk dipilih oleh bentuk wajah persegi adalah

bingkai kacamata yang terdiri dari sedikit lekukan yang dapat melembutkan fitur wajah. Rekomendasi kacamata berdasarkan bentuk wajah persegi, yaitu *Frame* bulat, *Frame* Oval, dan *Frame* Aviator untuk menghindari bingkai dengan sudut yang tajam (Optik PEEP, 2026).

### 3. Wajah Bulat

Bentuk wajah bulat memiliki panjang yang sama dengan lebar wajahnya dengan titik terlebar berada di pipi. Wajah bulat memiliki penampilan bentuknya yang melingkar diantar kontur wajahnya (Budianto dkk., 2023). Garis rahang yang membulat dengan dagunya lebih sempit dari dahi. Rekomendasi kacamata berdasarkan bentuk wajah hati terdiri dari *Frame* Kotak dan *Frame* Aviator (Optik PEEP, 2026).

### 4. Wajah Lonjong

Bentuk wajah yang berbentuk lonjong atau oblong memiliki fitur wajah yang memanjang dengan ukuran wajah yang lebih panjang dari lebarnya, sehingga bingkai kacamata yang cocok adalah bingkai yang menambah lebar dan volume wajah untuk membantu menyeimbangkan proporsi wajah (Budianto dkk., 2023). Sehingga untuk menghindari bingkai kacamata yang menonjolkan panjang wajah contohnya, *frame* Oversized dan *frame* Kotak (Optik PEEP, 2026).

### 5. Wajah Oval

Bentuk wajah oval memiliki bentuk yang ideal dengan proporsi yang seimbang. Wajah dengan bentuk oval memiliki ukuran yang lebih panjang dibandingkan luasnya dengan proporsional yang membentuk telur terbalik (Budianto dkk., 2023).

Kacamata yang dapat digunakan adalah bingkai yang lebih eksploratif dari bentuk wajah lainnya, namun perlu menghindari bingkai yang berukuran besar. Sehingga bingkai yang cocok adalah bingkai yang dapat mempertahankan keseimbangan alami wajah *Frame Persegi, Frame Bulat, Frame Cat Eye, Frame Aviator* (Optik PEEP, 2026).

## 2.8. Penelitian Terkait

Penelitian terkait disajikan dalam bentuk tabel *State of The Art* (SOTA) yang bertujuan untuk memberikan perbandingan serta gambaran tentang penelitian yang sudah dilakukan sebelumnya sehingga dapat diketahui posisi penelitian ini dan dapat mengidentifikasi celah penelitian berdasarkan permasalahan, kontribusi, hasil, dan kekurangan penelitian. Berikut Tabel 2.3 mengenai penelitian terkait sebagai dasar dalam penelitian ini.

**Tabel 2.3** *State of The Art* (SOTA)

No.	Konten	Penjelasan
1.	Judul	<i>Facial Shape-Based Eyeglass Recommendation Using Convolutional Neural Networks</i> (Rifat dkk., 2023)
	Permasalahan	Permasalahan pada penelitian ini adalah banyak orang yang merasa kesulitan dalam memilih kacamata yang cocok berdasarkan bentuk wajah.
	Kontribusi/ Keterbaharuan	Penelitian yang dilakukan memberikan kontribusi dalam mengusulkan arsitektur baru dengan membangun <i>pipeline</i> yang merekomendasikan kacamata.
	Hasil Utama	Arsitektur CNN yang digunakan antara lain <i>InceptionV4, ViT Small, DenseNet121, ResNet50, dan VGG16</i> . <i>Test Accuracy</i> yang didapatkan paling besar 75% dengan jumlah data 5500 gambar bentuk wajah.
	Kekurangan	Jumlah dataset yang digunakan relatif tidak banyak atau sedikit dengan kualitas gambar yang kurang konsisten sehingga mengakibatkan akurasi model yang masih terbatas.

2.	Judul	Klasifikasi Bentuk Bingkai ( <i>Frame</i> ) Kacamata Menggunakan CNN dengan Arsitektur <i>Inception V3</i> dan <i>Argumented Reality</i> Berbasis Android. (Ramadhan dkk., 2024)
	Permasalahan	Pemilihan <i>frame</i> kacamata secara manual yang mengakibatkan menghabiskan banyak waktu dan tidak akurat dalam mencocokkan kacamata.
	Kontribusi/ Keterbaharuan	Kontribusi penelitian ini adalah mengembangkan model CNN dengan memberikan rekomendasi bingkai kacamata berdasarkan bentuk wajah.
	Hasil Utama	Menggunakan Arsitektur <i>InceptionV3</i> untuk memprediksi <i>frame</i> kacamata dan penggunaan aplikasi AR dalam memberikan kesempatan dalam mencoba <i>frame</i> kacamata yang direkomendasikan. Hasil <i>test accuracy</i> pada penelitian ini sekitar 78.6% sedangkan 96% untuk <i>training accuracy</i> .
	Kekurangan	Menggunakan arsitektur <i>InceptionV3</i> dengan akurasi <i>training</i> dan <i>validation</i> yang berkesenjangan sehingga terindikasi <i>overfitting</i> .
3.	Judul	Pengenalan Bentuk Wajah dengan Metode <i>Convolutional Neural Network</i> untuk Pemilihan Model Kacamata Secara Online (Budianto dkk., 2023)
	Permasalahan	Sebagai akibat dari pandemi, ada dua masalah yang diangkat. Yang pertama adalah tantangan bagi pelanggan untuk mencoba kacamata secara langsung. Yang kedua adalah tantangan saat berbelanja online untuk mengetahui model kacamata yang tepat untuk bentuk wajah mereka.
	Kontribusi/ Keterbaharuan	Penelitian ini berkontribusi dalam eksplorasi model yaitu dengan dua pendekatan yaitu menggunakan model klasik dan CNN
	Hasil Utama	Metode <i>Viola-Jones</i> yang memberikan pengelompokan dasar menggunakan <i>Adaptive Boosting</i> dengan penggabungan <i>classifier</i> menggunakan <i>Cascade Classifier</i> dan CNN yang memiliki dua bagian yaitu Feature Extraction Layer dan Fully Connected Layer dengan akurasi 84,38%.
	Kekurangan	Penulis menyebutkan bahwa proses pelatihan model CNN membutuhkan waktu yang cukup lama.

4.	Judul	Sistem Identifikasi Bentuk Wajah untuk pemulihan <i>frame kacamata</i> menggunakan metode <i>Transfer Learning</i> (Tahta Hidayatillah dkk., 2022)
	Permasalahan	Permasalahan yang diangkat di penelitian ini adalah pada Optik Kusuma Wonosobo tidak adanya sistem komputerisasi dalam proses pemilihan bingkai kacamata. Sehingga, pelanggan harus berkonsultasi langsung dengan ahli, yang dianggap tidak efisien karena membutuhkan waktu identifikasi yang lama.
	Kontribusi/ Keterbaharuan	Kontribusi penelitian ini melakukan eksplorasi model yaitu penggunaan dari arsitektur <i>MobileNet</i>
	Hasil Utama	Dibuatnya website sebagai sistem rekomendasi frame kacamata dengan menerapkan model yang dilatih menggunakan CNN berarsitektur <i>MobileNet</i> dengan akurasi tes sebesar 74%
	Kekurangan	Penelitian ini menunjukkan model yang belum maksimal sehingga akurasi model perlu ditingkatkan kembali melalui segi dataset yang digunakan serta penggunaan sistem dengan menggunakan perangkat <i>mobile</i> .
5.	Judul	<i>A Deep Learning Approach to Face Shape Classification for Hairstyle, Beauty and Eyewear Recommendations.</i> (Sagarika & Mokashi, 2024)
	Permasalahan	Permasalahan yang diangkat pada penelitian ini adalah ketidakakuratan metode tradisional, yang bergantung pada pengukuran manual dan penilaian subjektif untuk menentukan bentuk wajah untuk memberikan rekomendasi individu, adalah masalah utama dalam penelitian ini.
	Kontribusi/ Keterbaharuan	Penelitian ini memperkenalkan metode deep learning untuk klasifikasi bentuk wajah yang akurat dengan <i>GoogLeNet</i> sebagai model terbaik mencapai akurasi 91.03%, sedangkan <i>MobileNetV3</i> mencapai 89.74%, ResNet50 mencapai 88.54%, dan Sequential CNN mencapai 80.02%.
	Hasil Utama	Penelitian ini menggunakan CNN tingkat lanjut yaitu GoogLeNet, MobileNetV3, ResNet50, dan Sequential CNN.
	Kekurangan	Dataset yang digunakan hanya terdiri dari 5000 gambar, sehingga perbedaan bentuk wajah dan kondisi pencahayaan tidak terwakili dengan baik.

6.	Judul	<i>Face Shape Classification on Bilinear Network with Attention Mechanism (Duan dkk., 2022)</i>
	Permasalahan	Penelitian ini membahas permasalahan mengenai kemampuan model dalam mengklasifikasikan bentuk wajah yang dapat mempengaruhi performa kinerja aplikasi, termasuk <i>face recognition</i> dan keperluan rekomendasi yang dipersonalisasi untuk berbagai keperluan.
	Kontribusi/ Keterbaharuan	Penelitian ini berkontribusi terhadap perkembangan algoritma deteksi wajah yang digunakan untuk merekomendasikan bentuk wajah.
	Hasil Utama	Penelitian ini menggunakan jaringan M-retinaFace untuk penyelarasan gambar wajah ( <i>alignment face</i> ) dan Jaringan AB-CNN yang digunakan dalam ekstraksi fitur menggunakan <i>attention mechanism</i> dengan metode yang didapat mencapai akurasi sebesar 89,8%.
	Kekurangan	Metode yang diusulkan pada penelitian ini memiliki batasan dalam pemilihan gambar bentuk wajah, sehingga diperlukan perluasan dataset dan bentuk wajah yang lebih mendetail.
7.	Judul	<i>Face Shape Classification using Inception v3 (Emmanuel &amp; Tio, 2019)</i>
	Permasalahan	Permasalahan ini menentukan cara dalam mengklasifikasikan bentuk wajah secara akurat
	Kontribusi/ Keterbaharuan	Penelitian ini melakukan eksperimen dengan melakukan pelatihan ulang ( <i>retraining</i> ) dari lapisan terakhir pada InceptionV3, kemudian dilakukan perbandingan antara metode lain seperti Linear Discriminant Analysis (LD), Support Vector Mechine (SVM), Artificial Neural Network, dan K-Nearest Neighbors (KNN)
	Hasil Utama	Salah satu eksplorasi yang ditentukan pada penelitian ini adalah InceptionV3 dengan akurasi model InceptionV3 yang dilatih ulang mencapai training accuracy mencapai 98%
	Kekurangan	Penelitian ini memiliki keterbatasan pada data label yang dilatih yang memiliki 500 gambar sehingga <i>cross validation</i> tidak dapat dilakukan.
8.	Judul	<i>Automatic Face Shape Classification Via Facial Landmark Measurements (Marinescu, 2021)</i>

	Permasalahan	Permasalahan penelitian ini adalah menentukan metode klasifikasi bentuk wajah manusia secara akurat berdasarkan citra 2D dengan identifikasi bentuk wajah yang bersifat subjektif.
	Kontribusi/ Keterbaharuan	Kontribusi penelitian ini adalah dengan menciptakan sistem yang kuat dan otomatis dalam melakukan klasifikasi bentuk wajah dengan memanfaatkan pengukuran landmark wajah.
	Hasil Utama	Penggunaan Facial Landmark dalam memahami kontur wajah dan pengembangan model prediksi menggunakan Naïve Bayes dan SVM.
	Kekurangan	Sifat subjektif klasifikasi bentuk wajah manusia menimbulkan identifikasi yang kurang konsisten dan akurat.
9.	Judul	<i>Human Facial Pattern Shape Classification Using a Retraining Stragey and Convolutional Neural Network Architecture</i> (Hidayat dkk., 2025)
	Permasalahan	Keberagaman bentuk wajah mengakibatkan kesulitan dalam membedakan bentuk wajah
	Kontribusi/ Keterbaharuan	Penelitian ini berkontribusi pada pengembangan pengetahuan dengan melakukan eksperimen dalam penelitian ini menghasilkan sebuah model yang dibangun berdasarkan lapisan akhir dari hasil <i>retraining</i> .
	Hasil Utama	Penggunaan InceptionV3 dalam melakukan pelatihan model dan penerapan prediksi klasifikasi bentuk wajah dan menerapkan <i>face alignment</i> dengan mendeteksi node yang membentuk kontur wajah dari hasil <i>retraining</i> InceptionV3 dengan akurasi maksimal sebesar 91.7%.
	Kekurangan	Jumlah dataset yang terbatas, hanya memiliki 747 gambar. Sehingga, keterbatasan akurasi dan pengenalan bentuk wajah.
10.	Judul	<i>Human Face Shape Classification with Machine Learning</i> (Mehta & Mohamed mahmoudt, 2022)
	Permasalahan	Permasalahan pada penelitian yang diangkat mengenai banyaknya keberagaman dan kemiripan bentuk wajah khususnya pada gambar wajah yang frontal dan terdeformasi
	Kontribusi/ Keterbaharuan	Penelitian ini berkontribusi dalam memprediksi bentuk wajah menggunakan jumlah landmark yang minimal dan teknik yang efisien.

	Hasil Utama	Penelitian ini menggunakan menggunakan tiga belas tanda wajah untuk menghitung empat belas fitur, yang terdiri dari lima jarak ter-normalisasi, tujuh rasio, dan dua sudut, menggunakan perhitungan yang sangat sederhana dan hanya menggunakan satu model klasifikasi untuk fitur tersebut.
	Kekurangan	Gambar yang terbatas dari sudut yang lurus dan tanpa adanya kemiringan, sehingga menentukan keakuratan bentuk wajah.
11.	Judul	<i>Bit-Plane and Correlation Spatial Attention Modules for Plant Disease Classification</i> (Wang & Cao, 2023)
	Permasalahan	Permasalahan penelitian ini adalah klasifikasi penyakit tanaman belum cukup efektif dan akurat sehingga masih menyebabkan tingginya ketergantungan pada ahli tanaman dalam produksi pertanian.
	Kontribusi/ Keterbaharuan	Penelitian ini memberikan kontribusi untuk meningkatkan identifikasi area penyakit pada tanaman, penelitian ini menawarkan model perhatian baru yang menggabungkan perhatian bit-plane dan perhatian korelasi ruang. Ditunjukkan oleh peningkatan akurasi pada dua dataset benchmark yang signifikan, model ini menunjukkan keunggulan dari pendekatan sebelumnya.
	Hasil Utama	Solusi yang ditawarkan adalah penggunaan bit plane attention untuk melokalisasi area penyakit dan correlation spatial attention untuk memberikan bobot area penting serta penggunaan Resnet101 sebagai arsitektur model
	Kekurangan	Penelitian ini hanya menguji model <i>attention</i> yang diusulkan pada dua dataset publik: AI Challenger 2018 dan PlantVillage. Selain itu, pengujian dilakukan dengan mengintegrasikan model perhatian ke dalam arsitektur ResNet101, sehingga efektivitasnya pada arsitektur CNN lain belum dipelajari secara menyeluruh.
12.	Judul	<i>CBAM: Convolutional Block Attention Module</i> (Woo dkk., 2018)
	Permasalahan	Permasalahan yang mendasari penelitian ini adalah belum efektifnya integrasi mekanisme perhatian ke dalam jaringan saraf konvolusi, sehingga penyempurnaan fitur dan kinerja dalam tugas seperti klasifikasi dan deteksi belum optimal.

	Kontribusi/ Keterbaruan	Penelitian ini berkontribusi pada penggabungan mekanisme perhatian yang efektif ke dalam arsitektur Convolutional Neural Networks (CNN) untuk meningkatkan kemampuan penyempurnaan fitur, yang akan memungkinkan peningkatan kinerja dalam tugas klasifikasi dan deteksi.
	Hasil Utama	Penerapanan inferensi <i>attention maps</i> pada dimensi channel dan spatial dengan mengalikannya dengan feature map untuk melakukan <i>adaptive feature refinement</i> .
	Kekurangan	Evaluasi pada penelitian ini hanya melihat dataset dan arsitektur yang sudah biasa digunakan. Selain itu, meskipun overhead dianggap kecil, penelitian ini tidak memeriksa secara menyeluruh dampak CBAM pada dataset yang lebih beragam.
13.	Judul	<i>Soft-Attention Improves Skin Cancer Classification Performance</i> (Datta dkk., 2021)
	Permasalahan	Permasalahan yang mendasari penelitian ini adalah bahwa jaringan saraf dalam masih belum mampu memfokuskan perhatian secara efektif pada karakteristik penting dari gambar kanker kulit. Akibatnya, kinerja klasifikasi pada gambar kanker kulit masih belum optimal.
	Kontribusi/ Keterbaruan	Penelitian ini berkontribusi dalam mengetahui efektivitas <i>Soft Attention</i> dalam mengklasifikasikan kanker kulit dengan menunjukkan <i>Soft Attention</i> meningkatkan kinerja sebesar 4,7% dari baseline menggunakan dataset HAM10000. Penelitian ini juga memberikan kontribusi dalam meningkatkan skor sensitivitas pada deteksi kanker yang mencapai 91,6% dengan menggunakan dataset ISIC-2017
	Hasil Utama	Penggunaan <i>soft attention</i> yang terbukti dapat meningkatkan kinerja model untuk menyelesaikan masalah neural network yang belum mampu memfokuskan perhatian secara efektif.
	Kekurangan	Terdapat ketidakseimbangan kelas dalam dataset yang digunakan sehingga memengaruhi kinerja model. Penelitian ini menekankan keterbatasan data beranotasi menambahkan tantangan dalam membangun model.
14.	Judul	<i>Reinforeced Self-Attention Network: a Hybrid of Hard and Soft Attention for Sequence Modeling</i> (Shen dkk., 2018)

	Permasalahan	Pemodelan dependensi dalam urutan panjang pada NLP masih belum optimal karena integrasi mekanisme <i>soft attention</i> dan <i>hard attention</i> belum dimanfaatkan secara efektif.
	Kontribusi/ Keterbaharuan	Untuk meningkatkan pemodelan sekuens, penelitian ini memperkenalkan model perhatian hibrida Reinforced Self-Attention (ReSA), yang menggabungkan <i>hard attention</i> dan <i>soft attention</i> .
	Hasil Utama	Dengan jumlah parameter yang lebih sedikit dan waktu komputasi yang lebih efisien, model ReSAN berbasis reinforced self-attention dapat mencapai performa canggih pada dataset SNLI dan SICK.
	Kekurangan	Metode ReSA masih bergantung pada pelatihan <i>hard attention</i> berbasis reinforcement learning, yang agak sulit dan sensitif terhadap stabilitas proses pembelajaran.
15.	Judul	SCSA: <i>Exploring the Synergistic Effects Between Spatial and Channel Attention</i> (Si dkk., 2024)
	Permasalahan	Penelitian ini memfokuskan permasalahan pada sulitnya memahami bagaimana integrasi perhatian saluran dan perhatian spasial dapat bekerja sama dalam meningkatkan kualitas representasi fitur dan meminimalkan kesenjangan semantik antar level multi-semantik.
	Kontribusi/ Keterbaharuan	Penelitian ini memberikan usulan dalam penggunaan modul Spatial and Channel Synergistic Attention module (SCSA) yang menggabungkan <i>spatial attention</i> dan <i>Spasial Multi-Semantik</i> yang Dapat Dibagikan untuk mengungkapkan hubungan sinergis antara <i>Shareable Multi-Semantic Spatial Attention and Progressive Channel-wise Self-Attention</i> .
	Hasil Utama	Eksperimen dilakukan dengan berbagai backbone yang menunjukkan perbandingan antar metode lain dengan metode SCSA dengan keunggulan akurasi yang signifikan pada beberapa arsitektur, yaitu ResNet-101, MobileNetV2, RepVGG-A0, Swim-T
	Kekurangan	Penggunaan model dalam lingkungan nyata menghadapi masalah latensi inferensi yang tinggi dan penggunaan konvolusi mendalam dengan lebar jaringan yang lebih besar dapat menyebabkan proses komputasi menjadi lebih lama.

16.	Judul	<i>Attention Mechansim-based CNN for Facial Expression Recognition</i> (Li dkk., 2020)
	Permasalahan	Model yang konvensional belum mampu mengekstraksi perubahan wajah yang berskala kecil
	Kontribusi/ Keterbaharuan	Kontribusi pada penelitian ini adalah mengusulkan CNN berbasis <i>attention</i> yang digunakan untuk pengenalan ekspresi wajah. Selain itu, penelitian ini melakukan sebuah fitur gabungan <i>Local Binary Pattern</i> (LBP) untuk meningkatkan <i>attention model</i> .
	Hasil Utama	Penelitian ini menggunakan beberapa dataset dalam eksperimennya seperti FER2013, CK+ dataset, Oulu-CASIA, dan JAFFE dengan masing-masing mengalami kenaikan Recognition Rates setelah menggunakan <i>attention module</i> dan <i>Local Binary Pattern</i> (LBP).
	Kekurangan	Penelitian ini memiliki keterbatasan pada metode yaitu hanya dapat diterapkan pada gambar wajah 2D sehingga belum mampu memanfaatkan data temporal, 3 dimensi, dan kedalaman wajah.
17.	Judul	<i>Performance-Efficiency Comparisons of Channel Attention Modules for ResNets</i> (Klomp dkk., 2023)
	Permasalahan	Belum diketahui secara pasti bagaimana penambahan <i>attention module</i> dalam mempengaruhi kinerja dan efisiensi komputasi arsitektur jaringan saraf, terutama untuk tugas klasifikasi gambar dan deteksi objek.
	Kontribusi/ Keterbaharuan	Kontribusi dari penelitian ini adalah untuk membandingkan berbagai modul perhatian yang efektif, memperkenalkan modul <i>Attention Bias</i> baru yang memiliki beban komputasi yang lebih rendah, dan melakukan penerapan fine-tuning pada jaringan pralatih yang diintegrasikan dengan <i>attention module</i> .
	Hasil Utama	Hasil penelitian menunjukkan bahwa arsitektur, framework, resolusi input, dan ukuran batch sangat berpengaruh pada efisiensi komputasinya, karena penambahan modul perhatian meningkatkan waktu inferensi sebesar sekitar 5 hingga 50%, meskipun kenaikan FLOP hanya sekitar 0,2 hingga 3%.
	Kekurangan	Penelitian ini memiliki kekurangan bahwa kenaikan waktu inferensi dan peningkatan akurasi modul perhatian tidak sebesar peningkatan akurasi, sehingga keuntungan

		performa tidak selalu sebanding dengan biaya komputasi yang ditambahkan.
18.	Judul	<i>A Comparative Study of Different CNN Models and Transfer Learning Effect for Underwater Object Classification in Side-Scan Sonar Images</i> (Du dkk., 2023)
	Permasalahan	Permasalahan yang mendasari penelitian ini adalah bahwa pengaruh berbagai model convolutional neural network (CNN) dan penerapan transfer learning terhadap akurasi klasifikasi objek bawah air pada citra side-scan sonar belum dipahami secara optimal, sehingga kinerja identifikasi objek masih belum mencapai tingkat akurasi yang diharapkan.
	Kontribusi/ Keterbaharuan	Penelitian ini memberikan kontribusi berupa analisis model CNN yang berbeda untuk klasifikasi objek bawah air dan melihat bagaimana penerapan transfer learning berdampak pada akurasi prediksi model.
	Hasil Utama	Transfer learning secara signifikan meningkatkan akurasi klasifikasi objek bawah laut pada citra side-scan sonar, dengan GoogleNet memiliki performa terbaik sebesar 94.81% pada data uji dan keseimbangan yang baik antara akurasi dan kompleksitas komputasi.
	Kekurangan	Penelitian ini hanya berfokus pada klasifikasi kapal dan pesawat, yang masih jarang digunakan dalam rekayasa kelautan yang tidak dirancang secara khusus untuk kebutuhan geofisika kelautan
19.	Judul	<i>A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope.</i> (Salehi dkk., 2023)
	Permasalahan	Kelangkaan data pelatihan dan risiko overfitting masih menjadi hambatan utama dalam proses pengembangan model pembelajaran mendalam. Di sisi lain, proses analisis dan validasi hasil menjadi kurang dapat diandalkan karena keterbatasan interpretabilitas model.
	Kontribusi/ Keterbaharuan	Penelitian ini berkontribusi dalam meningkatkan pemahaman tentang CNN dan transfer learning dengan menekankan elemen penting CNN. Ini juga menunjukkan bagaimana model CNN berbasis analisis dan klasifikasi gambar dapat diterapkan pada model pralatih dengan transfer learning.

	Hasil Utama	Penelitian ini menunjukkan bahwa analisis citra medis secara signifikan lebih akurat dan efisien dengan kombinasi CNN dan transfer learning, khususnya dalam situasi di mana data dan sumber daya komputasi terbatas. Oleh karena itu, kombinasi ini dapat menjadi alat pendukung diagnosis yang andal bagi radiolog.
	Kekurangan	Keterbatasan penelitian ini ada pada keterbatasan data pelatihan, yang dapat menyebabkan overfitting dan masalah vanishing gradient.
20.	Judul	<i>Multi-Level Circulation Pattern Classification Based on the Transfer Learning CNN Network</i> (Y. Liu dkk., 2022)
	Permasalahan	Klasifikasi pola sirkulasi udara masih mengalami kekurangan ketepatan karena biasanya hanya mengandalkan pendekatan tunggal, yang membuatnya kurang efisien dalam meramalkan cuaca ekstrem, terutama hujan deras yang berkepanjangan. Jadi, diperlukan sebuah metode klasifikasi pola sirkulasi yang bertingkat dan dapat meningkatkan ketepatan ramalan untuk peristiwa cuaca ekstrem serta iklim.
	Kontribusi/ Keterbaharuan	Penelitian ini berkontribusi pada pengembangan model CNN yang menggunakan transfer learning dengan variabel atmosfer tingkat multi (500 hPa, 200 hPa, 850 hPa) untuk memaksimalkan pengelompokan pola sirkulasi. Model tersebut menghadirkan dasar fisika yang lebih solid untuk memperbaiki proyeksi cuaca.
	Hasil Utama	Penelitian ini menghasilkan model CNN untuk klasifikasi sirkulasi dengan tiga level, yang menunjukkan tingkat akurasi yang lebih baik (92,5%) jika dibandingkan dengan model satu level (85%).
	Kekurangan	Penelitian ini masih fokus pada satu wilayah (Jianghuai Basin) dan jumlah kasus historis yang relatif terbatas (72 hari kejadian hujan lebat). Oleh karena itu, kemampuan untuk menggeneralisasi model multi-level transfer learning CNN terhadap wilayah lain, pola sirkulasi yang berbeda, dan kondisi iklim ekstrem yang lebih beragam belum sepenuhnya diuji.

Banyak penelitian mengenai klasifikasi bentuk wajah dan rekomendasi kacamata dengan menggunakan *computer vision*. Beberapa penelitian seperti (Budianto dkk., 2023; Ramadhan dkk., 2024; Rifat dkk., 2023; Sagarika & Mokashi, 2024; Tahta Hidayatillah dkk., 2022) telah menggunakan CNN dan *Transfer Learning* dengan berbagai arsitektur seperti *InceptionV3*, *GoogleNet*, *ResNet*, *MobileNet*, dan *DenseNet* dalam mengekstraksi fitur bentuk wajah. Akurasi yang dihasilkan dari penelitian tersebut sangat bervariasi. Walaupun, penelitian tersebut berhasil meningkatkan performa dibandingkan dengan metode tradisional (Marinescu, 2021), terdapat permasalahan yang identik, contohnya pada keterbatasan keberagaman dataset, ketidakseimbangan data, dan beberapa penelitian terindikasi *underfitting*.

Berbagai pendekatan telah diusulkan dalam mengintegrasikan attention mechanism dalam meningkatkan kualitas representasi fitur CNN, melalui *channel attention*, *spatial attention*, maupun kombinasi keduanya (Duan dkk., 2022; Si dkk., 2024; Wang & Cao, 2023; Woo dkk., 2018). Terbukti bahwa CBAM dan *attention module* lainnya, dapat meningkatkan kinerja dalam berbagai tugas klasifikasi dan deteksi. Namun, aplikasinya untuk klasifikasi bentuk wajah masih terbatas dan belum diuji secara menyeluruh pada arsitektur CNN. Akibatnya, peluang untuk meningkatkan kinerja model belum dimaksimalkan. (Klomp dkk., 2023). Oleh karena itu, penelitian ini berkontribusi dengan mengembangkan arsitektur CNN menggunakan *transfer learning* dengan *attention mechanism* untuk meningkatkan akurasi dan kemampuan generalisasi model dalam klasifikasi bentuk wajah.

Tabel 2.4 menunjukkan beberapa penelitian yang menyajikan informasi mengenai perbedaan penelitian yang diusulkan dengan penelitian yang sudah ada sebelumnya. Matriks penelitian tersebut menunjukkan perbandingan penelitian yang serupa dengan penelitian ini. Tujuan penelitian yang ada pada tabel 2.4. menunjukkan beberapa penelitian telah melakukan klasifikasi bentuk wajah dan melakukan rekomendasi kacamata. Penelitian yang dilakukan oleh (Budianto dkk., 2023; Ramadhan dkk., 2024; Rifat dkk., 2023; Sagarika & Mokashi, 2024; Tahta Hidayatillah dkk., 2022) telah melakukan klasifikasi dan rekomendasi kacamata dengan berbagai arsitektur seperti Inception, Resnet, DenseNet, VGG, ViT, GoogleNet, dan MobileNet. Sedangkan penelitian (Duan dkk., 2022; Emmanuel & Tio, 2019; Marinescu, 2021) hanya melakukan klasifikasi wajah, namun dengan eksplorasi pendekatan atau metode seperti Penggunaan *M-Retina Face*, dan penggunaan *machine learning* seperti *Naive Bayes* dan *Support Vector Machine* (SVM) dengan menggabungkan ekstraksi fitur. Namun terdapat sebuah celah, yaitu eksplorasi arsitektur dan mengurangi sifat *overlap* yang ada pada karakteristik wajah sehingga model dapat melakukan generalisasi dengan ekstraksi fitur yang optimal pada pemberian atensi terhadap dataset. Sehingga, penelitian ini menerapkan *dual block attention mechanism* yang terdiri dari *channel attention* dan *spatial attention*.

**Tabel 2.4** Matriks Penelitian

Peneliti	Tujuan		Algoritma	Arsitektur CNN								Attention	
	Klasifikasi	Rekomendasi		Inception	ResNet	DenseNet	VGG	ViT	MobileNet	EfficientNet	GooleNet	Spatial	Channel
(Rifat dkk., 2023)	✓	✓	<i>CNN</i>	✓	✓	✓	✓	✓	-	-	-	-	-
(Ramadhan dkk., 2024)	✓	✓	<i>CNN</i>	✓	-	-	-	-	-	-	-	-	-
(Budianto dkk., 2023)	✓	✓	<i>CNN + Viola Jones</i>	-	-	-	-	-	-	-	-	-	-
(Tahta Hidayatillah dkk., 2022)	✓	✓	<i>CNN</i>	-	-	-	-	-	✓	-	-	-	-
(Sagarika & Mokashi, 2024)	✓	✓	<i>CNN</i>	-	✓	-	-	-	✓	-	✓	-	-
(Duan dkk., 2022)	✓	-	<i>CNN + M-Retina Face</i>	-	-	-	-	-	-	-	-	-	-
(Emmanuel & Tio, 2019)	✓	-	<i>CNN</i>	✓	-	-	-	-	-	-	-	-	-
(Marinescu, 2021)	✓	-	<i>Naïve Bayes &amp; SVM</i>	-	-	-	-	-	-	-	-	-	-
<b><i>(Ours)</i></b>	✓	✓	<b><i>CNN</i></b>	-	-	-	-	-	-	✓	-	✓	✓