

BAB II

TINJAUAN PUSTAKA

2.1. Penelitian Terkait

Untuk menyusun landasan teori dan pemetaan penelitian yang komprehensif terkait klasifikasi citra sampah dan arsitektur model *deep learning*, penelitian ini menggunakan Systematic Literature Review (SLR) sebagai metode dokumentasi dan analisis literatur ilmiah secara sistematis. Metode SLR memungkinkan peneliti untuk mengidentifikasi, mengevaluasi, dan mensintesis temuan-temuan dari studi terdahulu dengan kriteria yang jelas dan terstandarisasi, sehingga hasilnya tidak hanya berupa kumpulan referensi, tetapi benar-benar mencerminkan gambaran ilmiah terkini tentang tren, performa model, dan gap penelitian yang belum terjawab. Menurut Khan dkk., proses tinjauan literatur disebut sistematis jika melibatkan pertanyaan yang dirumuskan dengan jelas, pemilihan studi relevan, penilaian kualitas penelitian tersebut, dan sintesis bukti ilmiah menggunakan metodologi yang eksplisit, sehingga dapat direplikasi oleh peneliti lain . Dengan demikian, penggunaan SLR dalam penelitian ini tidak hanya memberi konteks yang kuat terhadap kajian arsitektur *Convolutional Neural Network* pada klasifikasi citra sampah tetapi juga mendukung perumusan *research gap* dan *framework* metodologis yang kokoh untuk studi eksperimen yang akan dilakukan. (Ryan et al., 2024)

Tabel 2.1 SLR

No	Nama penulis Tahun	Judul penelitian	Metode penelitian	Masalah penelitian	Hasil
1	(Peradeniya, et al., 2025)	Studi Perbandingan Penggunaan Konvolusional Model Jaringan Saraf Tiruan dalam Klasifikasi Limbah	Model-model yang telah dilatih sebelumnya, Yaitu MobileNet, MobileNetV2, Xception, InceptionV3, EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, DenseNet121, DenseNet169, dan DenseNet201	klasifikasi sampah manual dapat menyebabkan ketegangan fisik dan masalah kesehatan bagi pekerja manusia	Model DenseNet169 akurasi model sebesar 87,68% serta skor terbaik untuk presisi, recall, dan f1-score.
2	(Kumala et al., 2025)	Perbandingan Arsitektur MobileNetV2 dan VGG16 dengan Transfer Pembelajaran untuk Klasifikasi Sampah Berbasis Citra Multikelas	Convolutional Neural Network (CNN), MobileNetV2 dan VGG16,	menentukan arsitektur CNN optimal untuk kendala penerapan di dunia nyata, khususnya mengenai trade-off efisiensi komputasi versus akurasi klasifikasi.	MobileNetV2 Akurasi 95,6% presisi 0,93, recall 0,93 F1-score 0,93, VGG16 Akurasi 89,13% presisi 0,91, recall 0,90, dan F1-score 0,90.
3	(Doly Ilham Saputra Huta Julu & Dewi Nurdiah, 2025))	Klasifikasi sampah organik dan non organik menggunakan transfer learning	VGG16, MobileNetV2, dan ResNet50V2.	Pengelolaan sampah di Indonesia menghadapi tantangan serius dengan 7,2 juta ton sampah belum terkelola dengan baik dari 202 kabupaten/kota, mencemari lingkungan	MobileNetV2 akurasi 90,13%, presisi 96,25%, dan F1-Score 87,88%, waktu inferensi 127,76 ms
4	(Thio & Susilo, 2025)	Identifikasi Pemilahan Sampah Berbasis Algoritma Transfer Learning CNN Menggunakan MobileNetV2 dan EfficientNetB0	MobileNetV2 dan EfficientNetB0.	Masalah pengelolaan sampah menjadi tantangan lingkungan yang tidak ada habisnya di kalangan masyarakat.	MobileNetV2 mencapai akurasi 87,31%, sementara EfficientNetB0 memperoleh akurasi 82,21%.

TABEL 2.1. SLR

Berdasarkan hasil kajian pada Tabel 2.1 dapat diketahui bahwa penelitian mengenai klasifikasi citra sampah telah berkembang cukup pesat dengan memanfaatkan berbagai arsitektur Convolutional Neural Network (CNN). Sebagian besar penelitian memanfaatkan model *Transfer Learning* seperti MobileNetV2, ResNet, EfficientNet, VGG16, maupun DenseNet karena mampu meningkatkan performa klasifikasi dibandingkan model CNN yang dibangun dari awal (*training from scratch*). Selain itu, penggunaan *Transfer Learning* juga mampu mengurangi waktu pelatihan serta meningkatkan kemampuan model dalam mengekstraksi fitur citra.

Meskipun demikian, sebagian besar penelitian terdahulu hanya berfokus pada penggunaan satu arsitektur model sehingga belum memberikan informasi mengenai perbandingan performa antararsitektur. Beberapa penelitian memang membandingkan beberapa model CNN, namun menggunakan dataset, jumlah kelas, maupun parameter pelatihan yang berbeda sehingga hasil yang diperoleh belum dapat dibandingkan secara langsung.

Berdasarkan kondisi tersebut, penelitian ini melakukan perbandingan dua arsitektur Transfer Learning, yaitu MobileNetV2 dan DenseNet169, dengan menggunakan dataset yang sama, jumlah kelas yang sama, ukuran citra yang sama, serta parameter pelatihan yang identik. Dengan demikian, hasil penelitian diharapkan mampu memberikan informasi yang lebih objektif mengenai karakteristik kedua model dalam melakukan klasifikasi citra sampah anorganik.

2.2.Landasan Teori

2.2.1 Deep Learning dan Convolutional Neural Network (CNN)

Perkembangan teknologi Artificial Intelligence (AI) telah mendorong lahirnya berbagai metode komputasi yang mampu menyelesaikan permasalahan kompleks secara otomatis, salah satunya adalah Deep Learning. Deep Learning merupakan cabang dari Machine Learning yang menggunakan jaringan saraf tiruan (*Artificial Neural Network*) dengan banyak lapisan (*hidden layer*) untuk mempelajari representasi data secara bertingkat. Berbeda dengan metode Machine Learning konvensional yang masih memerlukan proses ekstraksi fitur secara manual, Deep Learning mampu melakukan proses ekstraksi fitur dan klasifikasi secara otomatis melalui proses pembelajaran dari sejumlah besar data. Kemampuan tersebut menjadikan Deep Learning sebagai salah satu teknologi yang paling banyak digunakan dalam bidang pengolahan citra, pengenalan suara, analisis teks, serta berbagai aplikasi berbasis kecerdasan buatan lainnya. (Fotovvatikhah et al., 2025)

Pada bidang pengolahan citra digital, Deep Learning menunjukkan perkembangan yang sangat pesat karena mampu mengenali pola visual yang kompleks, seperti bentuk, warna, tekstur, dan objek dalam suatu citra. Melalui proses pembelajaran yang dilakukan secara berulang (*iterative learning*), model Deep Learning dapat membangun representasi fitur yang semakin abstrak pada setiap lapisan jaringan sehingga menghasilkan tingkat akurasi yang lebih tinggi dibandingkan

metode klasifikasi citra tradisional. Oleh karena itu, Deep Learning telah banyak diterapkan pada berbagai bidang, seperti deteksi penyakit, identifikasi wajah, kendaraan otonom, hingga klasifikasi citra sampah.

Salah satu arsitektur Deep Learning yang paling banyak digunakan dalam klasifikasi citra adalah Convolutional Neural Network (CNN). CNN merupakan pengembangan dari Artificial Neural Network yang dirancang khusus untuk mengolah data berbentuk citra dua dimensi. Arsitektur ini mampu melakukan ekstraksi fitur secara otomatis melalui proses konvolusi sehingga dapat mengenali karakteristik objek tanpa memerlukan proses ekstraksi fitur secara manual. Kemampuan tersebut menjadikan CNN sebagai metode yang efektif dalam berbagai tugas *computer vision*, termasuk klasifikasi citra sampah.

Secara umum, CNN terdiri atas beberapa lapisan utama, yaitu Convolution Layer, Activation Layer, Pooling Layer, Fully Connected Layer, dan Output Layer. Setiap lapisan memiliki fungsi yang berbeda dalam proses pembelajaran citra.

1. Convolution Layer

Convolution Layer merupakan lapisan utama dalam CNN yang bertugas melakukan proses ekstraksi fitur dari citra masukan menggunakan kernel atau filter. Filter akan bergerak pada seluruh bagian citra untuk menghasilkan feature map yang berisi informasi mengenai pola-pola penting, seperti tepi (*edge*), sudut (*corner*), tekstur, maupun bentuk objek. Pada lapisan awal, CNN umumnya

mempelajari fitur-fitur sederhana, sedangkan pada lapisan yang lebih dalam CNN mampu mengenali fitur yang lebih kompleks hingga membentuk representasi objek secara utuh.

2. Activation Layer

Setelah proses konvolusi dilakukan, feature map akan melewati fungsi aktivasi untuk memperkenalkan sifat nonlinier pada jaringan. Fungsi aktivasi yang paling banyak digunakan adalah Rectified Linear Unit (ReLU) karena mampu mempercepat proses pelatihan serta mengurangi permasalahan *vanishing gradient*. ReLU bekerja dengan mengubah seluruh nilai negatif menjadi nol, sedangkan nilai positif dipertahankan sehingga proses pembelajaran menjadi lebih stabil.

3. Pooling Layer

Pooling Layer digunakan untuk mengurangi dimensi feature map sehingga jumlah parameter yang diproses menjadi lebih sedikit. Selain mempercepat proses komputasi, pooling juga membantu mengurangi risiko *overfitting* dengan mempertahankan informasi penting dari feature map. Teknik pooling yang paling umum digunakan adalah Max Pooling, yaitu memilih nilai maksimum pada setiap area tertentu dari feature map.

4. Fully Connected Layer

Feature map yang dihasilkan dari proses pooling selanjutnya diubah menjadi bentuk vektor (*flatten*) sebelum diteruskan ke Fully

Connected Layer. Pada tahap ini seluruh neuron saling terhubung sehingga jaringan dapat mempelajari hubungan antarfitur yang telah diekstraksi sebelumnya dan menghasilkan keputusan klasifikasi.

5. Output Layer

Output Layer merupakan lapisan terakhir yang menghasilkan prediksi kelas. Pada penelitian klasifikasi biner seperti penelitian ini, Output Layer menggunakan fungsi aktivasi Sigmoid sehingga menghasilkan nilai probabilitas antara 0 dan 1 untuk menentukan kelas cardboard atau paper.

6. Kelebihan Convolutional Neural Network

CNN memiliki beberapa keunggulan dibandingkan metode klasifikasi citra konvensional, antara lain:

1. Mampu melakukan ekstraksi fitur secara otomatis tanpa memerlukan proses *feature engineering*.
2. Memiliki tingkat akurasi yang tinggi pada berbagai permasalahan klasifikasi citra.
3. Mampu mengenali pola visual yang kompleks meskipun terdapat variasi posisi, ukuran, maupun pencahayaan objek.
4. Sangat sesuai diterapkan pada pendekatan Transfer Learning, sehingga waktu pelatihan menjadi lebih singkat dengan tetap mempertahankan performa yang tinggi.

7. Kekurangan Convolutional Neural Network

Di samping berbagai kelebihan, CNN juga memiliki beberapa keterbatasan, di antaranya:

1. Membutuhkan jumlah data pelatihan yang relatif besar untuk memperoleh performa optimal.
2. Memerlukan sumber daya komputasi yang tinggi, terutama pada arsitektur dengan jumlah parameter yang besar.
3. Berpotensi mengalami overfitting apabila data pelatihan terbatas atau proses regularisasi tidak dilakukan dengan baik.

Dalam penelitian ini, CNN digunakan melalui pendekatan Transfer Learning dengan dua arsitektur berbeda, yaitu MobileNetV2 dan DenseNet169. Kedua model dipilih karena memiliki karakteristik yang berbeda. MobileNetV2 merupakan arsitektur ringan yang dirancang untuk efisiensi komputasi, sedangkan DenseNet169 memiliki konektivitas antarlapisan yang lebih padat sehingga mampu menghasilkan representasi fitur yang lebih kaya. Perbandingan kedua model tersebut diharapkan dapat memberikan gambaran mengenai pengaruh kompleksitas arsitektur terhadap performa klasifikasi citra sampah anorganik.

2.2.2 MobileNetV2

MobileNetV2 merupakan salah satu arsitektur Convolutional Neural Network (CNN) yang dikembangkan oleh Google untuk menghasilkan model klasifikasi citra dengan kebutuhan komputasi yang

rendah namun tetap memiliki performa yang tinggi. Arsitektur ini dirancang untuk mengatasi keterbatasan perangkat dengan sumber daya komputasi terbatas, seperti telepon pintar, perangkat *embedded*, dan perangkat berbasis Internet of Things (IoT). Dibandingkan dengan arsitektur CNN konvensional, MobileNetV2 mampu mengurangi jumlah parameter dan operasi komputasi secara signifikan sehingga proses pelatihan maupun inferensi dapat dilakukan lebih cepat tanpa penurunan akurasi yang berarti. Konsep tersebut menjadikan MobileNetV2 sebagai salah satu model yang paling banyak digunakan dalam berbagai aplikasi *computer vision*, termasuk klasifikasi citra sampah.

Arsitektur MobileNetV2 dibangun berdasarkan dua komponen utama, yaitu Depthwise Separable Convolution dan Inverted Residual with Linear Bottleneck. Depthwise Separable Convolution memisahkan proses konvolusi menjadi dua tahap, yaitu *depthwise convolution* dan *pointwise convolution*. Pada tahap pertama, setiap kanal citra diproses secara independen untuk mengekstraksi informasi spasial, sedangkan tahap kedua menggunakan konvolusi berukuran 1×1 untuk menggabungkan informasi antar-kanal. Pendekatan ini mampu menurunkan jumlah operasi komputasi secara signifikan dibandingkan konvolusi standar, sehingga MobileNetV2 menjadi lebih ringan dan efisien.

Secara matematis, proses *Depthwise Convolution* pada MobileNetV2 dapat dinyatakan sebagai berikut:

$$Y_{i,j,k} = \sum_{m=1}^K \sum_{n=1}^K X_{i+m,j+n,k} \times W_{m,n,k}$$

- X = citra masukan (*input feature map*)
- W = kernel *depthwise convolution*
- Y = hasil *feature map*
- K = ukuran kernel konvolusi
- i, j = posisi piksel
- k = kanal (*channel*)

Persamaan tersebut menunjukkan bahwa proses *Depthwise Convolution* melakukan operasi konvolusi secara terpisah pada setiap kanal citra masukan. Berbeda dengan konvolusi standar yang menggabungkan seluruh kanal secara bersamaan, metode ini hanya memproses satu kanal pada setiap operasi sehingga jumlah parameter dan operasi komputasi menjadi jauh lebih kecil.

Setelah proses tersebut selesai dilakukan, hasil *Depthwise Convolution* kemudian digabungkan menggunakan *Pointwise Convolution* dengan kernel berukuran 1×1 , yang secara matematis dapat dituliskan sebagai berikut:

$$Y_{i,j,l} = \sum_{k=1}^C X_{i,j,k} \times W_{k,l}$$

- C = jumlah kanal masukan
- l = jumlah kanal keluaran

Persamaan tersebut menjelaskan bahwa *Pointwise Convolution* berfungsi menggabungkan informasi dari seluruh kanal hasil *Depthwise Convolution* sehingga diperoleh representasi fitur yang lebih kaya tanpa menambah kompleksitas komputasi secara signifikan. Kombinasi kedua proses tersebut dikenal sebagai *Depthwise Separable Convolution*, yang menjadi karakteristik utama MobileNetV2 dalam menghasilkan model yang ringan namun tetap memiliki performa klasifikasi yang tinggi.

Selain menggunakan *Depthwise Separable Convolution*, MobileNetV2 juga memperkenalkan konsep *Inverted Residual* dan *Linear Bottleneck*. Berbeda dengan residual block pada ResNet yang terlebih dahulu mengurangi dimensi fitur, MobileNetV2 justru memperluas dimensi fitur sebelum melakukan proses ekstraksi, kemudian mengompresinya kembali menggunakan *Linear Bottleneck*. Pendekatan ini memungkinkan jaringan mempertahankan informasi penting selama proses propagasi tanpa meningkatkan kompleksitas model secara berlebihan. Dengan demikian, MobileNetV2 mampu menghasilkan representasi fitur yang lebih baik meskipun memiliki jumlah parameter yang relatif kecil. (Yong et al., 2023)

Selain menggunakan *Depthwise Separable Convolution*, MobileNetV2 juga menerapkan mekanisme *Inverted Residual with Linear Bottleneck*. Secara matematis, hubungan residual tersebut dapat dituliskan sebagai berikut.

$$y = x + F(x)$$

- x = masukan (*input feature map*)
- $F(x)$ = hasil transformasi melalui beberapa lapisan konvolusi
- y = keluaran (*output feature map*)

Persamaan tersebut menunjukkan bahwa informasi dari lapisan sebelumnya tetap dipertahankan melalui koneksi residual sehingga proses propagasi informasi menjadi lebih stabil dan risiko hilangnya informasi penting (*information loss*) dapat dikurangi. Berbeda dengan ResNet yang melakukan koneksi residual pada dimensi fitur yang besar, MobileNetV2 terlebih dahulu memperluas dimensi fitur (*expansion layer*), kemudian melakukan ekstraksi fitur menggunakan *Depthwise Convolution*, dan akhirnya mengompresi kembali dimensi fitur menggunakan *Linear Bottleneck*. Fungsi *Linear Bottleneck* sendiri menggunakan fungsi identitas yang dapat dinyatakan sebagai:

$$f(x) = x$$

Penggunaan fungsi linear pada lapisan bottleneck bertujuan untuk mempertahankan informasi penting selama proses reduksi dimensi sehingga representasi fitur yang dihasilkan tetap optimal meskipun jumlah parameter model relatif kecil.

Keunggulan utama MobileNetV2 terletak pada keseimbangan antara akurasi dan efisiensi komputasi. Model ini memiliki ukuran yang jauh lebih kecil dibandingkan banyak arsitektur CNN lainnya sehingga memerlukan memori yang lebih sedikit serta waktu inferensi yang lebih cepat. Oleh karena itu, MobileNetV2 sangat sesuai digunakan pada sistem klasifikasi citra secara real-time maupun pada perangkat dengan spesifikasi perangkat keras yang terbatas. Selain itu, MobileNetV2 juga sangat kompatibel dengan pendekatan Transfer Learning, yaitu memanfaatkan bobot awal (*pre-trained weights*) dari dataset ImageNet sehingga proses pelatihan menjadi lebih cepat dan membutuhkan jumlah data pelatihan yang lebih sedikit dibandingkan pelatihan dari awal (*training from scratch*).

Sejumlah penelitian terkini menunjukkan bahwa MobileNetV2 memiliki performa yang kompetitif pada berbagai permasalahan klasifikasi sampah. Penelitian oleh Liying Yong mengembangkan sistem klasifikasi sampah rumah tangga menggunakan MobileNetV2 dan menunjukkan bahwa model tersebut mampu meningkatkan akurasi secara signifikan dibandingkan CNN konvensional, sekaligus mempertahankan ukuran model yang ringan sehingga sesuai diterapkan pada perangkat bergerak.

Penelitian lain yang dilakukan oleh Universitas Dian Nuswantoro menerapkan MobileNetV2 untuk klasifikasi sampah organik dan anorganik menggunakan teknik Transfer Learning. Hasil penelitian menunjukkan bahwa kombinasi MobileNetV2 dengan augmentasi data mampu menghasilkan performa klasifikasi yang baik serta meningkatkan kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya. Penelitian tersebut juga menegaskan bahwa penggunaan pembagian dataset yang tepat (*training*, *validation*, dan *testing*) serta augmentasi hanya pada data pelatihan dapat mengurangi risiko *overfitting*.

Hasil serupa juga dilaporkan pada penelitian yang menambahkan kategori plastic sachets pada sistem klasifikasi sampah berbasis MobileNetV2. Penelitian tersebut menerapkan *stratified data splitting*, augmentasi citra, serta normalisasi piksel sehingga model mampu mempertahankan performa klasifikasi yang tinggi meskipun jumlah kelas ditingkatkan. Hal ini menunjukkan bahwa MobileNetV2 memiliki kemampuan adaptasi yang baik terhadap variasi dataset dan dapat digunakan untuk berbagai skenario klasifikasi sampah. (Pritama et al., 2025)

Pada penelitian ini, MobileNetV2 digunakan sebagai salah satu model pembandingan terhadap DenseNet169 melalui pendekatan Transfer Learning. Kedua model dilatih menggunakan konfigurasi yang identik, meliputi ukuran citra 224×224 piksel, pembagian dataset 80% data

pelatihan, 10% data validasi, dan 10% data pengujian, penggunaan teknik data augmentation hanya pada data pelatihan, optimizer Adam, *learning rate* sebesar 0,00001, serta jumlah *epoch* yang sama. Penyamaan konfigurasi tersebut dilakukan agar perbandingan performa kedua model berlangsung secara objektif dan tidak dipengaruhi oleh perbedaan parameter pelatihan.

Berdasarkan hasil implementasi pada penelitian ini, MobileNetV2 mampu memberikan performa klasifikasi yang baik dengan ukuran model yang relatif kecil dan waktu pelatihan yang lebih singkat dibandingkan DenseNet169. Karakteristik tersebut menjadikan MobileNetV2 sebagai alternatif yang tepat apabila sistem klasifikasi citra sampah akan diterapkan pada perangkat dengan keterbatasan sumber daya komputasi, seperti perangkat mobile, *smart bin*, maupun sistem berbasis IoT.

2.2.3 DenseNet169

DenseNet169 merupakan salah satu varian dari arsitektur Dense Convolutional Network (DenseNet) yang diperkenalkan untuk mengatasi permasalahan *vanishing gradient*, meningkatkan aliran informasi antar-lapisan, serta mendorong penggunaan kembali (*feature reuse*) fitur yang telah dipelajari sebelumnya. Berbeda dengan arsitektur CNN konvensional yang hanya menghubungkan suatu lapisan dengan lapisan berikutnya, DenseNet menghubungkan setiap lapisan dengan seluruh lapisan sebelumnya. Dengan demikian, setiap lapisan memperoleh

informasi dari seluruh fitur yang telah dipelajari sebelumnya sehingga mampu menghasilkan representasi fitur yang lebih kaya dan lebih diskriminatif. Pendekatan ini membuat DenseNet menjadi salah satu arsitektur CNN dengan performa tinggi pada berbagai tugas klasifikasi citra.

DenseNet169 terdiri atas 169 lapisan yang tersusun dalam beberapa Dense Block dan Transition Layer. Pada setiap Dense Block, keluaran dari seluruh lapisan sebelumnya digabungkan (*concatenate*) sebagai masukan bagi lapisan berikutnya. Strategi tersebut memungkinkan informasi penting tetap dipertahankan sepanjang proses propagasi tanpa mengalami degradasi kualitas. Sementara itu, Transition Layer digunakan untuk mengurangi ukuran feature map melalui proses konvolusi dan *pooling*, sehingga kebutuhan memori tetap dapat dikendalikan meskipun jumlah koneksi antar-lapisan sangat banyak.

Salah satu karakteristik utama DenseNet169 adalah mekanisme Dense Connectivity, yaitu setiap lapisan (*layer*) menerima masukan (*input*) dari seluruh lapisan sebelumnya. Secara matematis hubungan tersebut dapat dinyatakan sebagai berikut:

$$x_l = H_l([x_0, x_1, x_2, \dots, x_{l-1}])$$

- x_l = keluaran (*output*) pada lapisan ke- l
- H_l = fungsi transformasi pada lapisan ke- l
- $[x_0, x_1, \dots, x_{l-1}]$ = proses *concatenation* seluruh keluaran lapisan

Persamaan tersebut menunjukkan bahwa setiap lapisan pada DenseNet tidak hanya menerima keluaran dari lapisan sebelumnya, tetapi juga seluruh keluaran dari lapisan-lapisan sebelumnya. Pendekatan ini memungkinkan informasi penting tetap dipertahankan selama proses propagasi sehingga mampu mengurangi masalah *vanishing gradient* dan meningkatkan proses pembelajaran jaringan.

Salah satu keunggulan utama DenseNet169 adalah kemampuannya dalam melakukan feature reuse, yaitu menggunakan kembali fitur yang telah dipelajari pada lapisan sebelumnya tanpa harus mempelajarinya kembali pada lapisan berikutnya. Pendekatan ini membuat jumlah parameter DenseNet relatif lebih efisien dibandingkan beberapa arsitektur CNN lain yang memiliki kedalaman serupa. Selain itu, konektivitas padat (*dense connectivity*) membantu memperlancar proses propagasi gradien sehingga proses pelatihan menjadi lebih stabil dan risiko *vanishing gradient* dapat diminimalkan.

Fungsi transformasi pada setiap lapisan DenseNet terdiri atas beberapa operasi yang dilakukan secara berurutan, yaitu *Batch Normalization (BN)*, *Rectified Linear Unit (ReLU)*, dan *Convolution*. Secara matematis fungsi tersebut dapat dituliskan sebagai berikut.

$$H_l(x) = W_l(BN(ReLU(x)))$$

- $ReLU(x)$ = fungsi aktivasi
- BN = Batch Normalization
- W_l = operasi konvolusi
- $H_l(x)$ = keluaran transformasi pada lapisan ke- l

Persamaan tersebut menunjukkan bahwa sebelum proses konvolusi dilakukan, data terlebih dahulu dinormalisasi menggunakan *Batch Normalization*, kemudian diproses menggunakan fungsi aktivasi *ReLU*. Kombinasi ketiga proses tersebut menghasilkan representasi fitur yang lebih stabil sehingga mempercepat proses pelatihan model.

Dalam bidang klasifikasi citra, DenseNet169 telah banyak digunakan karena mampu menghasilkan akurasi yang tinggi pada objek dengan karakteristik visual yang kompleks. Kemampuan tersebut menjadikan DenseNet169 sesuai untuk mengidentifikasi objek yang memiliki variasi warna, bentuk, tekstur, maupun pencahayaan, termasuk pada klasifikasi citra sampah. Dibandingkan model CNN yang lebih ringan seperti MobileNetV2, DenseNet169 umumnya menghasilkan representasi fitur yang lebih lengkap, meskipun membutuhkan sumber daya komputasi yang lebih besar.

Selain menerapkan konektivitas padat, DenseNet juga menggunakan konsep *Growth Rate*, yaitu jumlah *feature map* baru yang dihasilkan pada setiap lapisan. Hubungan tersebut dapat dinyatakan sebagai berikut.

$$C_l = C_0 + k \times l$$

Keterangan:

- C_l = jumlah kanal pada lapisan ke- l
- C_0 = jumlah kanal awal
- k = *growth rate*
- l = nomor lapisan

Persamaan tersebut menunjukkan bahwa jumlah *feature map* akan bertambah secara bertahap sesuai nilai *growth rate*. Semakin besar nilai *growth rate*, semakin banyak informasi visual yang dapat dipelajari oleh jaringan, namun kebutuhan memori dan waktu komputasi juga akan meningkat.

Beberapa penelitian terkini menunjukkan bahwa DenseNet169 memiliki performa yang sangat baik pada sistem klasifikasi sampah. Zhang dkk. menggunakan DenseNet169 pada dataset NWN-Trash dan memperoleh hasil klasifikasi yang kompetitif dibandingkan beberapa arsitektur CNN lainnya. Penelitian tersebut menunjukkan bahwa konektivitas padat pada DenseNet mampu meningkatkan kemampuan ekstraksi fitur sehingga model lebih mudah membedakan objek dengan karakteristik visual yang hampir serupa. (sciencedirect.com)

Penelitian lain yang dilakukan oleh Fotovvatikhah dkk. dalam tinjauan sistematis mengenai klasifikasi sampah berbasis kecerdasan buatan menyatakan bahwa keluarga DenseNet merupakan salah satu arsitektur yang paling konsisten menghasilkan performa tinggi pada

berbagai dataset klasifikasi sampah. Hal tersebut disebabkan oleh kemampuan DenseNet dalam mempertahankan informasi visual selama proses propagasi serta mengurangi kehilangan informasi akibat kedalaman jaringan.

Selain itu, penelitian oleh Darussalam dan Arief menunjukkan bahwa penggunaan DenseNet pada sistem klasifikasi sampah dengan pendekatan *ensemble learning* mampu meningkatkan akurasi hingga lebih dari 95% setelah dilakukan optimasi proses pelatihan dan penyeimbangan dataset. Hasil tersebut menunjukkan bahwa DenseNet tidak hanya efektif digunakan sebagai model tunggal, tetapi juga memiliki potensi besar untuk dikembangkan pada sistem klasifikasi citra yang lebih kompleks.

Pada penelitian ini, DenseNet169 digunakan sebagai model pembanding terhadap MobileNetV2 menggunakan pendekatan Transfer Learning. Kedua model dilatih menggunakan konfigurasi pelatihan yang identik, meliputi ukuran citra 224×224 piksel, pembagian dataset 80% data pelatihan, 10% data validasi, dan 10% data pengujian, teknik data augmentation hanya pada data pelatihan, optimizer Adam, *learning rate* sebesar 0,00001, serta jumlah *epoch* yang sama. Penyamaan konfigurasi tersebut bertujuan agar perbandingan performa kedua model berlangsung secara objektif. (Fotovvatikhah et al., 2025)

Untuk mengurangi kompleksitas komputasi akibat bertambahnya jumlah *feature map*, DenseNet169 menggunakan *Transition Layer* yang berfungsi melakukan reduksi dimensi melalui operasi *Convolution* berukuran 1×1 dan *Average Pooling*. Secara matematis proses tersebut dapat dinyatakan sebagai berikut.

$$y = AvgPool(Conv_{1 \times 1}(x))$$

Keterangan:

- $Conv_{1 \times 1}$ = konvolusi berukuran 1×1
- $AvgPool$ = operasi *Average Pooling*
- x = masukan
- y = keluaran

Penggunaan *Transition Layer* bertujuan mengurangi ukuran *feature map* sehingga jumlah parameter tetap terkendali dan kebutuhan memori dapat diminimalkan tanpa mengurangi kualitas representasi fitur yang telah dipelajari.

Berdasarkan hasil implementasi penelitian ini, DenseNet169 menunjukkan performa klasifikasi yang lebih tinggi dibandingkan MobileNetV2. Model ini memperoleh nilai Accuracy sebesar 96,14%, Precision sebesar 95,70%, Recall sebesar 97,09%, dan F1-Score sebesar 96,39% pada data pengujian. Hasil tersebut menunjukkan bahwa kemampuan DenseNet169 dalam mengekstraksi fitur yang lebih mendalam memberikan kontribusi terhadap peningkatan performa klasifikasi citra sampah anorganik. Namun demikian, peningkatan

performa tersebut diikuti oleh kebutuhan waktu pelatihan yang lebih lama serta ukuran model yang lebih besar dibandingkan MobileNetV2. Oleh karena itu, DenseNet169 lebih sesuai diterapkan pada sistem yang memiliki sumber daya komputasi yang memadai dan mengutamakan tingkat akurasi dibandingkan efisiensi komputasi.

2.2.4 Transfer Learning

Transfer Learning merupakan salah satu teknik dalam Deep Learning yang memanfaatkan pengetahuan yang telah dipelajari oleh suatu model pada tugas sebelumnya untuk digunakan kembali pada tugas baru yang memiliki karakteristik serupa. Pada klasifikasi citra, teknik ini biasanya dilakukan dengan menggunakan model yang telah dilatih (*pre-trained model*) pada dataset berskala besar, seperti ImageNet, kemudian dilakukan proses *fine-tuning* menggunakan dataset penelitian. Pendekatan ini memungkinkan model memperoleh kemampuan ekstraksi fitur yang baik tanpa harus memulai proses pelatihan dari awal (*training from scratch*).

Penggunaan Transfer Learning menjadi solusi yang efektif ketika jumlah data pelatihan relatif terbatas. Model yang telah dilatih pada jutaan citra mampu mengenali fitur-fitur dasar seperti tepi (*edges*), tekstur, warna, hingga bentuk objek. Selanjutnya, selama proses *fine-tuning*, model akan menyesuaikan parameter tertentu agar mampu mengenali karakteristik khusus pada dataset penelitian. Dengan demikian, Transfer Learning dapat mempercepat proses pelatihan

Secara umum, proses Transfer Learning terdiri atas dua tahapan utama, yaitu feature extraction dan fine-tuning. Pada tahap *feature extraction*, sebagian besar lapisan dasar (*base model*) dibekukan (*freeze*) sehingga bobot yang telah dipelajari sebelumnya tidak berubah selama proses pelatihan. Selanjutnya, pada tahap *fine-tuning*, beberapa lapisan terakhir dibuka kembali (*unfreeze*) sehingga model dapat mempelajari karakteristik khusus dari dataset baru tanpa kehilangan kemampuan ekstraksi fitur yang telah dimiliki sebelumnya.

Menurut (Tan & Le, 2019) penggunaan Transfer Learning terbukti mampu meningkatkan performa klasifikasi citra sekaligus mengurangi waktu pelatihan dibandingkan pelatihan dari awal. Teknik ini juga membantu mengurangi risiko *overfitting*, terutama ketika jumlah data pelatihan tidak terlalu besar. Oleh karena itu, Transfer Learning telah menjadi pendekatan yang banyak digunakan pada berbagai penelitian klasifikasi citra, termasuk klasifikasi objek medis, tanaman, kendaraan, hingga klasifikasi sampah. (Tan & Le, 2019)

Dalam bidang klasifikasi sampah, beberapa penelitian menunjukkan bahwa penggunaan Transfer Learning menghasilkan performa yang lebih baik dibandingkan model CNN yang dilatih dari awal. (Fotovvatikhah et al., 2025) melaporkan bahwa sebagian besar penelitian klasifikasi sampah modern menggunakan model *pre-trained* seperti MobileNetV2, DenseNet, EfficientNet, dan ResNet karena mampu memberikan tingkat akurasi yang tinggi dengan waktu pelatihan

yang relatif singkat. Selain itu, penggunaan Transfer Learning juga membuat model lebih stabil terhadap variasi data dan perubahan kondisi pencahayaan pada citra sampah.

Pada penelitian ini, pendekatan Transfer Learning diterapkan pada kedua model yang dibandingkan, yaitu MobileNetV2 dan DenseNet169. Kedua model menggunakan bobot awal (*pre-trained weights*) dari ImageNet, kemudian dilakukan proses *fine-tuning* pada beberapa lapisan terakhir. Seluruh konfigurasi pelatihan disamakan, meliputi ukuran citra 224×224 piksel, optimizer Adam, *learning rate* sebesar 0,00001, ukuran *batch* 32, serta pembagian dataset menjadi 80% data pelatihan, 10% data validasi, dan 10% data pengujian. Dengan konfigurasi tersebut, perbandingan performa kedua model dapat dilakukan secara objektif.

Implementasi Transfer Learning pada penelitian ini terbukti mampu menghasilkan performa klasifikasi yang baik. DenseNet169 memperoleh Accuracy sebesar 96,14%, sedangkan MobileNetV2 juga menunjukkan performa yang kompetitif dengan ukuran model yang jauh lebih kecil. Hasil tersebut menunjukkan bahwa Transfer Learning mampu meningkatkan kemampuan kedua model dalam mengenali karakteristik visual sampah anorganik tanpa memerlukan proses pelatihan dari awal.

2.2.5 Data Augmentation

Data augmentation merupakan teknik prapemrosesan data yang bertujuan untuk meningkatkan jumlah dan keragaman data pelatihan melalui transformasi citra tanpa mengubah identitas kelasnya. Teknik ini banyak digunakan dalam pengembangan model Deep Learning untuk mengatasi keterbatasan jumlah dataset serta meningkatkan kemampuan generalisasi model terhadap data baru. Dengan melakukan augmentasi, model memperoleh variasi bentuk objek yang lebih beragam sehingga mampu mempelajari karakteristik visual secara lebih baik dan mengurangi risiko overfitting.

Pada klasifikasi citra, data augmentation dilakukan dengan menerapkan berbagai transformasi geometris maupun fotometris terhadap citra asli. Transformasi tersebut meliputi rotasi (*rotation*), pergeseran (*width shift* dan *height shift*), pembesaran (*zoom*), pencerminan horizontal (*horizontal flip*), serta perubahan posisi objek di dalam citra. Seluruh transformasi tersebut dilakukan tanpa mengubah label kelas sehingga model memperoleh variasi data pelatihan yang lebih representatif terhadap kondisi nyata. (Essa & Murshid, 2023)

Penggunaan data augmentation telah menjadi praktik standar dalam penelitian klasifikasi citra modern. Sebuah tinjauan mengenai model Transfer Learning untuk klasifikasi citra menunjukkan bahwa augmentasi data mampu meningkatkan kemampuan model dalam mengenali objek yang memiliki variasi posisi, ukuran, pencahayaan,

maupun orientasi. Selain meningkatkan akurasi, data augmentation juga terbukti membantu mengurangi *overfitting* ketika jumlah dataset relatif terbatas.

Dalam penelitian klasifikasi sampah, penggunaan data augmentation juga memberikan dampak yang signifikan terhadap peningkatan performa model. Penelitian mengenai klasifikasi sampah berbasis Transfer Learning menunjukkan bahwa kombinasi augmentasi citra dengan model CNN menghasilkan model yang lebih stabil ketika diuji menggunakan data baru. Model menjadi lebih mampu mengenali karakteristik visual sampah meskipun terdapat perubahan sudut pengambilan gambar, ukuran objek, maupun kondisi pencahayaan. (Ajumon et al., 2023)

Pada penelitian ini, data augmentation hanya diterapkan pada data pelatihan (training set) sesuai dengan rekomendasi berbagai penelitian terdahulu dan arahan dosen pembimbing. Data validasi dan data pengujian tidak diberikan proses augmentasi, melainkan hanya dilakukan normalisasi piksel ($rescale = 1/255$). Pendekatan tersebut bertujuan agar proses evaluasi model dilakukan menggunakan data yang benar-benar merepresentasikan kondisi sebenarnya sehingga hasil pengujian tidak dipengaruhi oleh transformasi buatan.

Bentuk augmentasi yang digunakan dalam penelitian ini meliputi:

- Rotation Range = 20° , untuk meningkatkan variasi orientasi objek.
- Zoom Range = 0,2, untuk variasi jarak pengambilan gambar.
- Width Shift Range = 0,2, untuk menggeser posisi objek secara horizontal.
- Height Shift Range = 0,2, untuk menggeser posisi objek secara vertikal.
- Horizontal Flip, untuk meningkatkan variasi bentuk objek.
- Fill Mode = nearest, mengisi piksel kosong setelah transformasi.

Konfigurasi tersebut diterapkan secara identik pada pelatihan MobileNetV2 dan DenseNet169 sehingga kedua model memperoleh kondisi pelatihan yang sama. Penyamaan konfigurasi ini dilakukan agar perbandingan performa kedua arsitektur benar-benar dipengaruhi oleh karakteristik model, bukan oleh perbedaan proses prapemrosesan data.

Berdasarkan implementasi penelitian, penggunaan data augmentation menghasilkan model yang memiliki kemampuan generalisasi yang lebih baik. Hal ini terlihat dari selisih nilai akurasi pelatihan dan validasi yang relatif kecil serta peningkatan performa klasifikasi pada data pengujian. Dengan demikian, data augmentation menjadi salah satu komponen penting dalam membangun model klasifikasi citra sampah anorganik yang lebih robust terhadap variasi kondisi citra.

2.2.6 Confusion Matrix

Confusion Matrix merupakan salah satu metode evaluasi yang paling umum digunakan untuk mengukur performa model klasifikasi. Confusion Matrix menyajikan hasil prediksi model dalam bentuk tabel yang membandingkan kelas aktual dengan kelas hasil prediksi sehingga dapat diketahui jumlah prediksi yang benar maupun prediksi yang salah. Melalui Confusion Matrix, peneliti dapat menganalisis jenis kesalahan yang dilakukan model serta menghitung berbagai metrik evaluasi seperti Accuracy, Precision, Recall, dan F1-Score. (Chu et al., 2018)

Pada permasalahan klasifikasi biner, Confusion Matrix terdiri atas empat komponen utama, yaitu:

- True Positive (TP), yaitu jumlah data positif yang berhasil diprediksi sebagai kelas positif.
- True Negative (TN), yaitu jumlah data negatif yang berhasil diprediksi sebagai kelas negatif.
- False Positive (FP), yaitu jumlah data negatif yang salah diprediksi sebagai kelas positif.
- False Negative (FN), yaitu jumlah data positif yang salah diprediksi sebagai kelas negative

Hubungan keempat komponen tersebut dapat digambarkan pada Tabel 2.2.

Tabel 2.2 Komponen(Chu et al., 2018)

Kondisi Aktual	Prediksi Positif	Prediksi Negatif
Positif	True Positive (TP)	False Negative (FN)
Negatif	False Positive (FP)	True Negative (TN)

Pada penelitian ini, kelas paper diperlakukan sebagai kelas positif, sedangkan cardboard sebagai kelas negatif ketika proses evaluasi dilakukan menggunakan *binary classification*. Nilai TP, TN, FP, dan FN kemudian digunakan untuk menghitung Accuracy, Precision, Recall, dan F1-Score sebagai indikator performa model.

Berdasarkan hasil implementasi penelitian, Confusion Matrix digunakan untuk mengevaluasi performa MobileNetV2 maupun DenseNet169 pada data pengujian (test set). Penggunaan data pengujian dilakukan agar hasil evaluasi benar-benar menggambarkan kemampuan model dalam mengklasifikasikan citra baru yang belum pernah digunakan selama proses pelatihan maupun validasi. Selain itu, visualisasi Confusion Matrix membantu mengidentifikasi kelas yang paling sering mengalami kesalahan klasifikasi sehingga dapat menjadi dasar analisis

2.2.7 Accuracy

Accuracy merupakan ukuran yang menunjukkan persentase jumlah prediksi yang benar dibandingkan dengan seluruh data yang diuji. Semakin tinggi nilai Accuracy, semakin baik kemampuan model dalam melakukan klasifikasi. Meskipun demikian, Accuracy tidak selalu mencerminkan performa model secara menyeluruh, terutama pada dataset yang tidak seimbang (*imbalanced dataset*). Oleh karena itu, Accuracy perlu diinterpretasikan bersama metrik lain seperti Precision, Recall, dan F1-Score. (Chu et al., 2018)

Persamaan Accuracy adalah:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Pada penelitian ini, Accuracy digunakan sebagai indikator umum untuk mengetahui tingkat keberhasilan MobileNetV2 dan DenseNet169 dalam mengklasifikasikan citra sampah anorganik.

2.2.8 Precision

Precision menunjukkan tingkat ketepatan model ketika memprediksi suatu kelas sebagai positif. Nilai Precision yang tinggi menunjukkan bahwa sebagian besar prediksi positif yang dihasilkan model memang benar-benar merupakan kelas positif. Dengan kata lain, Precision berhubungan langsung dengan kemampuan model dalam meminimalkan False Positive.

Persamaan Precision adalah:

$$Precision = \frac{TP}{TP + FP}$$

Pada penelitian ini, Precision digunakan untuk mengetahui seberapa tepat model dalam mengidentifikasi citra sampah sesuai kelas sebenarnya.

2.2.9 Recall

Recall atau Sensitivity menunjukkan kemampuan model dalam menemukan seluruh data yang benar-benar termasuk ke dalam kelas positif. Semakin tinggi nilai Recall, semakin sedikit data positif yang gagal dikenali oleh model. Nilai Recall berkaitan langsung dengan jumlah False Negative yang dihasilkan model.

Persamaan Recall adalah:

$$Recall = \frac{TP}{TP + FN}$$

Dalam penelitian ini, Recall digunakan untuk mengukur kemampuan MobileNetV2 dan DenseNet169 dalam mendeteksi seluruh citra sampah pada masing-masing kelas.

2.2.10 F1-Score

F1-Score merupakan rata-rata harmonik (*harmonic mean*) antara Precision dan Recall. Metrik ini digunakan ketika kedua indikator tersebut sama-sama penting untuk dipertimbangkan. Nilai F1-Score yang tinggi menunjukkan bahwa model memiliki keseimbangan yang baik antara ketepatan prediksi dan kemampuan mendeteksi seluruh data

positif. Oleh karena itu, F1-Score menjadi salah satu metrik yang paling banyak digunakan pada penelitian klasifikasi citra modern.

Persamaan F1-Score adalah:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Pada penelitian ini, F1-Score digunakan sebagai indikator utama dalam membandingkan performa MobileNetV2 dan DenseNet169 karena mampu memberikan gambaran yang lebih seimbang dibandingkan Accuracy saja.

2.2.11 Research Gap

Berdasarkan hasil Systematic Literature Review (SLR) terhadap berbagai penelitian mengenai klasifikasi citra sampah menggunakan metode *Deep Learning*, diketahui bahwa sebagian besar penelitian berfokus pada pengembangan model dengan tingkat akurasi setinggi mungkin melalui penggunaan berbagai arsitektur Convolutional Neural Network (CNN) seperti VGG16, ResNet50, EfficientNet, MobileNetV2, dan DenseNet. Berbagai penelitian tersebut menunjukkan bahwa CNN mampu memberikan performa yang baik dalam mengidentifikasi jenis sampah berdasarkan karakteristik visual citra. Namun demikian, masih terdapat beberapa kesenjangan penelitian (*research gap*) yang belum banyak dibahas secara komprehensif.

Kesenjangan pertama terletak pada minimnya penelitian yang secara langsung membandingkan performa MobileNetV2 dan DenseNet169 menggunakan dataset, konfigurasi pelatihan, serta metode

evaluasi yang identik. Sebagian besar penelitian hanya mengevaluasi satu arsitektur atau membandingkannya dengan model lain seperti ResNet atau EfficientNet, sehingga sulit memperoleh kesimpulan objektif mengenai keunggulan relatif antara MobileNetV2 dan DenseNet169 pada klasifikasi citra sampah.

Kesenjangan kedua berkaitan dengan perbedaan konfigurasi eksperimen yang digunakan pada penelitian terdahulu. Variasi jumlah *epoch*, ukuran *batch*, *learning rate*, teknik augmentasi, maupun pembagian dataset menyebabkan hasil penelitian sulit dibandingkan secara langsung. Akibatnya, peningkatan atau penurunan performa model belum tentu disebabkan oleh karakteristik arsitektur CNN, tetapi dapat dipengaruhi oleh perbedaan konfigurasi pelatihan.

Kesenjangan ketiga adalah belum banyak penelitian yang mengevaluasi keseimbangan antara akurasi dan efisiensi komputasi. Sebagian besar penelitian hanya berfokus pada peningkatan nilai akurasi tanpa mempertimbangkan ukuran model, jumlah parameter, waktu pelatihan, maupun potensi implementasi pada perangkat dengan sumber daya terbatas. Padahal, aspek efisiensi sangat penting apabila sistem klasifikasi sampah akan diterapkan pada perangkat bergerak, *embedded system*, atau teknologi berbasis Internet of Things (IoT). (Mahdi, 2024). Kesenjangan berikutnya adalah kurangnya konsistensi penggunaan metrik evaluasi. Beberapa penelitian hanya melaporkan nilai akurasi, sementara metrik lain seperti *Precision*, *Recall*, *F1-Score*, dan *Confusion*

Matrix tidak selalu disajikan secara lengkap. Kondisi ini menyebabkan evaluasi performa model menjadi kurang komprehensif karena tidak mampu menggambarkan kemampuan model dalam mengidentifikasi setiap kelas secara menyeluruh. (Azzahra et al., 2024)

Berdasarkan berbagai kesenjangan tersebut, penelitian ini dilakukan untuk memberikan kontribusi melalui perbandingan langsung antara MobileNetV2 dan DenseNet169 menggunakan dataset citra sampah anorganik yang sama dengan konfigurasi pelatihan yang identik, meliputi pembagian dataset 80% data pelatihan, 10% data validasi, dan 10% data pengujian, penggunaan *data augmentation* hanya pada data pelatihan, optimizer Adam, *learning rate* yang sama, serta jumlah *epoch* yang sama. Evaluasi dilakukan menggunakan metrik Accuracy, Precision, Recall, F1-Score, dan Confusion Matrix sehingga hasil perbandingan dapat menggambarkan secara objektif kelebihan dan keterbatasan masing-masing arsitektur. Dengan demikian, penelitian ini diharapkan dapat memberikan rekomendasi model yang paling sesuai berdasarkan kebutuhan antara akurasi klasifikasi dan efisiensi komputasi.