

BAB II

TINJAUAN PUSTAKA & LANDASAN TEORI

2.1 Tinjauan Pustaka

2.1.1 Penelitian Terdahulu

Terdapat berbagai penelitian terkait dengan klasifikasi tumor otak. Penelitian yang dilakukan oleh Amou et al. [14] menerangkan bahwa CNN berhasil dalam mengklasifikasi citra MRI ke dalam tiga jenis tumor otak (*Glioma*, *Meningioma*, dan *Pituitary*). Karena CNN memiliki jumlah lapisan yang tinggi dan memungkinkan untuk memahami fitur secara efektif. Untuk mencari hyperparameter yang optimal maka diusulkan teknik optimasi hyperparameter efisien berbasis *Bayesian Optimization* untuk CNN. Penelitian ini membandingkan lima model *deep learning pre-trained* dengan model CNN yang dioptimalkan. Model yang diusulkan berhasil mendapatkan akurasi validasi sebesar 98,70% sedangkan model VGG16, VGG19, ResNet50, InceptionV3, dan DenseNet201 mencapai akurasi validasi sebesar 97,08%, 96,43%, 89,29%, 92,86%, dan 94. 81%. Wahlang et al. [15] menggunakan arsitektur *deep learning* untuk mengklasifikasikan citra MRI otak menjadi normal atau abnormal. Penelitian ini juga menambahkan atribut seperti jenis kelamin dan usia dalam klasifikasi yang lebih akurat. Arsitektur yang diuji seperti LeNet, AlexNet, ResNet, dan pendekatan tradisional seperti SVM. Metode lain yang diusulkan untuk klasifikasi yaitu menggabungkan LeNet dan ResNet menjadi CNN-DNN. Bias usia dan jenis kelamin ditemukan lebih berguna dan berperan penting dalam klasifikasi. Hasil penelitian menunjukkan model yang terinspirasi dari LeNet dengan *convolutional layer*, *pooling*, dan *fully connected layer* mengungguli SVM dan AlexNet. Akurasi yang didapatkan dari metode yang diusulkan yaitu SVM, AlexNet, ResNet, LeNet, LIM, dan CNN-DNN adalah 82%, 64%, 44%, 87%, 88%, dan 80%.

Ahmad et al. [16] mengusulkan gabungan antara dua model generatif yaitu *variational autoencoders* (VAEs) dan *generative adversarial network* (GAN). Untuk mengklasifikasikan tumor otak dengan menggunakan citra MRI. Metode yang diusulkan untuk membantu GAN dalam menghasilkan gambar MRI otak yang tampak realistis. Gambar yang dihasilkan artifisial ini dapat mengatasi keterbatasan kumpulan data medis. Dengan menggunakan tambahan gambar MRI tumor otak hasil artifisial maka akurasi rata-rata klasifikasi meningkat dari 72.63% menjadi 96.25%. El Kader et al. [17] menjelaskan tentang bagaimana penggunaan *Deep Wavelet Autoencoder Model* (DWAE model) untuk melakukan klasifikasi tumor otak (abnormal) dan tidak ada tumor otak (normal) dengan total 2500 citra otak MRI. Dengan

menggunakan model DWAE dapat mendeteksi dan mengklasifikasikan tumor otak dengan akurasi tinggi, waktu yang relatif singkat, dan *loss validation* yang rendah. Model mendapatkan akurasi sebesar 99.3% dengan nilai *validation loss* 0.1 ini menunjukkan bahwa model dapat memfasilitasi deteksi tumor otak secara otomatis.

Alanazi et al. [18] mengusulkan metode *isolated deep learning network* dimana jaringan ini mempelajari tugas dari awal dan tidak memerlukan pengetahuan yang dipelajari sebelumnya. CNN akan dilatih dari awal menggunakan dataset gambar MRI otak dan kemudian digunakan kembali untuk melatih model *deep learning* melalui *transfer learning*. Arsitektur *Isolated CNN* dibedakan menjadi 19, 22, dan 25 lapisan. *Isolated CNN* 22-layer memiliki akurasi terbaik untuk mengklasifikasikan gambar MRI otak ke dalam kelas tumor dan non-tumor. Hasil yang didapatkan untuk mengklasifikasikan tumor otak dengan dataset I menjadi dua kelas adalah 99.33% sedangkan mengklasifikasikan tumor otak ke dalam empat kelas atau dataset II mendapatkan akurasi sebesar 95.75%. Model yang diusulkan yaitu hasil dari *isolated CNN* digabungkan dengan model *transfer learning* secara akurat berhasil mengklasifikasikan sebanyak 2969 citra otak MRI dari 3064 dengan akurasi 96.6%.

Gab Allah et al. [19] menerapkan sebuah augmentasi *Progressive Growing Generative Adversarial Network* (PGGAN) yang digunakan untuk menghasilkan citra MRI tumor otak yang realistis. Serta membandingkan augmentasi klasik dengan PGGAN. Untuk mencari kombinasi terbaik dari arsitektur VGG19 maka menguji tiga arsitektur yang berbeda: VGG19 + CNN, VGG19 + GRU, dan VGG19 + Bi-GRU. Semua arsitektur diuji dengan empat optimisasi yang berbeda yaitu Adam, Adamax, RMSprop, dan Nadam. Dengan menggabungkan arsitektur VGG19 + CNN dan menggunakan augmentasi data berbasis PGGAN menunjukkan kemampuan untuk klasifikasi glioma, meningioma, dan tumor hipofisis mendapatkan akurasi sebesar 98.54% dimana model ini mengungguli model lainnya. Badža dan Barjaktarovi [20] membangun sebuah model CNN dengan total 22 lapisan dan menggunakan augmentasi data. Model dilatih dengan menggunakan optimasi Adam. Dataset yang digunakan sebanyak 3064 gambar MRI dengan kontras yang ditingkatkan dengan pembobotan T1. Metode yang diusulkan yaitu menggunakan *k-fold cross-validation* dengan dataset dibagi menjadi 60% untuk pelatihan, 20% untuk validasi, dan 20% untuk uji coba mendapatkan akurasi sebesar 95.40%. Dengan penambahan dataset yang berbeda model diujikan dan menggunakan *k-fold cross-validation* mendapatkan akurasi sebesar 96.56%.

Alrashedy et al. [21] mengusulkan BrainGan, yaitu kerangka kerja untuk menghasilkan citra MRI otak berdasarkan Vanilla GAN dan DCGAN. Selain itu penelitian ini juga menerapkan *transfer learning* dengan menggunakan arsitektur CNN, MobileNetV2, dan

ResNet152V2. Hasil percobaan yang dilakukan menunjukkan bahwa model ResNet152V2 mengungguli dua model lainnya dalam hal kinerja keseluruhan. Berdasarkan gambar MRI otak yang dihasilkan oleh DCGAN, ResNet152V2 mendapatkan akurasi sebesar 99.09% untuk mengklasifikasikan dua kelas normal dan kanker. Ghassemis et al. [22] melatih DCGAN untuk menghasilkan citra MRI otak untuk mengatasi pembatasan ukuran dataset. Penelitian melakukan dua percobaan yaitu GAN + ConvNet dengan *introduced split* dan GAN + ConvNet dengan *random split*. Mendapatkan akurasi sebesar 93.01% dan 95.6% hasil ini diungguli oleh GAN + ConvNet dengan *random split*.



Tabel 2. 1 Tabel Penelitian Terdahulu

No	Ref	Tujuan	Metode	Kelas	Hasil
1.	[3]	Untuk klasifikasi tumor otak menjadi kanker dan non kanker	CNN dengan 8 lapisan digabungkan dengan augmentasi data dan deteksi tepi	2 Kelas: Kanker (155) Non Kanker (98)	Hasil uji coba menunjukkan bahwa model yang diusulkan dengan kumpulan data yang kecil mendapatkan akurasi yang sempurna. Model yang diusulkan mendapatkan akurasi 100%. VGG-16 mendapatkan akurasi 96% ResNet-50 mendapatkan akurasi 89% Inception-V3 mendapatkan akurasi 75%
2.	[11]	Klasifikasi tumor otak menggunakan CNN	Menggunakan CNN yang digabungkan dengan augmentasi data dengan normalisasi min-max	3 Kelas: Meningioma (708) Glioma (1426) Pituitary (930)	dengan normalisasi min-max dan diikuti dengan filter <i>Gaussian</i> dan <i>Laplacian</i> . Mendapatkan akurasi 98.78%. Metode EfficientNet yang diusulkan: Akurasi = 98.78% ResNet50: Akurasi = 96.33% MobileNet: Akurasi = 96.94% MobileNetV2: Akurasi = 94.80%
3.	[12]	Untuk mengklasifikasikan tumor otak dengan membandingkan beberapa arsitektur yang sudah dilatih sebelumnya	CNN: VGG19 InceptionV3 MobileNetV2	2 Kelas: Tumor otak (2513) Tidak ada tumor otak (2087)	Membandingkan beberapa arsitektur yang sudah dilatih model dengan akurasi terbaik yaitu MobileNetV2 Akurasi: MobileNetV2 = 92% InceptionV3 = 91% VGG19 = 88.22%
4.	[14]	Untuk mengklasifikasikan tumor otak dengan teknik optimasi hyperparameter yang efisien	CNN + <i>Bayesian Optimization</i>	3 Kelas: Meningioma (708) Glioma (1426) Pituitary (930)	Model yang diusulkan merupakan arsitektur CNN dasar yang ditambahkan <i>Bayesian optimization</i> untuk hyperparameter yang efisien. Model ini mendapatkan akurasi yang lebih tinggi dibandingkan beberapa model yang sudah dilatih sebelumnya.

					<p>Akurasi: CNN yang dioptimalkan = 98.70% VGG16 = 97.08% VGG19 = 96.43% DenseNet201 = 94.81% InceptionV3 = 92.86% ResNet50 = 89.29</p>
5.	[15]	Untuk mengklasifikasikan citra MRI otak menjadi normal atau abnormal yang ditambahkan jenis kelamin dan usia sebagai atribut tambahan untuk klasifikasi yang lebih akurat	Menggabungkan CNN + DNN dan mencoba beberapa arsitektur. Menambahkan atribut jenis kelamin dan usia.	2 Kelas: Normal (tidak ada tumor) dan Abnormal (tumor)	Klasifikasi otak bekerja lebih baik untuk otak dari berbagai usia dan jenis kelamin daripada otak dari jenis kelamin yang sama menggunakan LIM, CNN-DNN, dan empat metodologi lainnya. Akurasi keseluruhan menggunakan usia dan jenis kelamin sebagai atribut yaitu: SVM = 82 % AlexNet = 64 % ResNet = 44 % LeNet = 87 % LeNet Inspired Model = 88% CNN-DNN = 81%
6.	[16]	Untuk menghasilkan citra tumor otak yang tampak realistis untuk meningkatkan hasil klasifikasi secara signifikan	CNN berbasis ResNet dan menggabungkan dua model generatif yaitu VAEs dan GANs	3 Kelas: Meningioma (708) Glioma (1426) Pituitary (930)	Dengan metode yang diusulkan yaitu menggunakan dua model generatif meningkatkan akurasi rata-rata dari 72.63% menjadi 96.25% Model yang diusulkan mendapatkan akurasi 92.3% tanpa augmentasi klasik. Model yang diusulkan dengan augmentasi klasik mendapatkan akurasi 96.25%
7.	[17]	Membuat metode yang efisien untuk deteksi dan klasifikasi tumor otak	Deep Wavelet Auto Encoder	2 Kelas: Normal Abnormal	Dengan menggunakan DWAE yang diusulkan mendapatkan akurasi lebih baik dibandingkan beberapa penelitian terdahulu Akurasi yang didapatkan adalah 99.3%

8.	[18]	Mengembangkan dan menganalisis <i>Isolated CNN</i> terhadap berbagai dataset citra MRI otak.	Melatih <i>Isolated CNN</i> (19, 22, dan 25 layer) lalu dikembangkan dengan metode <i>transfer learning</i> .	Dataset I: 2 kelas yaitu tumor dan tidak tumor Dataset II: Otak tumor glioma, tumor meningioma, tidak ada tumor, dan tumor hipofisis Dataset III: Tumor glioma Tumor meningioma, Tumor hipofisis	Dengan metode <i>Isolated CNN</i> yang dibangun dengan 22layer mendapatkan akurasi tertinggi. Model dengan dataset I mendapatkan akurasi 99.33%. Dataset II model memperoleh akurasi sebesar 95.75%. Untuk dataset III model mendapatkan akurasi 96.6%.
9.	[19]	Untuk menguji kemajuan pendekatan baru untuk klasifikasi MRI tumor otak	CNN dengan arsitektur VGG19 ditambah data augmentasi berbasis PGGAN.	3 Kelas: Meningioma (708) Glioma (1426) Pituitary (930)	Dengan metode yang diusulkan yaitu augmentasi PGGAN untuk menghasilkan citra MRI tumor otak yang realistis. Membandingkan tiga model yang diusulkan yaitu VGG19+CNN, VGG19+GRU, VGG19-Bi-GRU semuanya menggunakan PGGAN. Model VGG19+CNN mendapatkan akurasi yang tertinggi yaitu 98.54%
10.	[20]	Membuat arsitektur CNN baru untuk klasifikasi tumor otak dari tiga jenis tumor.	CNN dengan total 22 lapisan sederhana dibandingkan dengan model <i>pre-trained</i> . Menggunakan augmentasi data.	3 Kelas: Meningioma (708) Glioma (1426) Pituitary (930)	Hasil klasifikasi yang dilakukan dengan menggunakan basis data citra MRI dengan kontras ditingkatkan dengan pembobotan T1 mendapatkan akurasi sebesar 96.56%.
11.	[21]	Menghasilkan citra MRI otak berdasarkan Vanilla GAN dan DCGAN	Vanilla GAN dan DCGAN untuk augmentasi data.	2 Kelas: No tumor Tumor	Dengan kombinasi DCGAN + ResNet152V2 mendapatkan akurasi tertinggi dibandingkan model lainnya dengan akurasi sebesar 99.09%.

12.	[22]	Mengatasi pembatasan ukuran dataset dengan menggunakan DCGAN	DCGAN + ConvNet	3 Kelas: Meningioma (708) Glioma (1426) Pituitary (930)	Hasil penelitian dengan menggunakan <i>random split</i> dapat mengungguli model satunya. Dengan akurasi sebesar 95.6%
-----	------	--	-----------------	--	---



2.1.2 Penelitian Sekarang

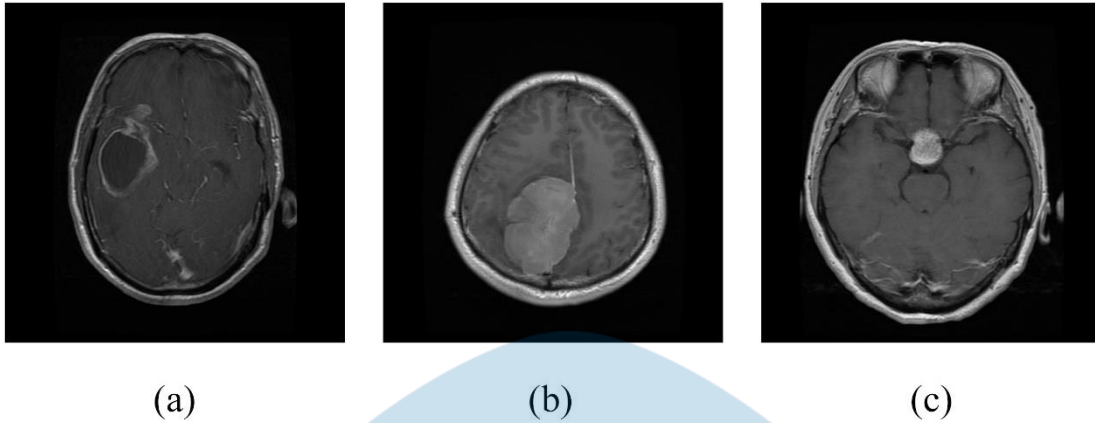
Berdasarkan penelitian terdahulu, beberapa peneliti menggunakan model CNN *pre-trained* [16], [12], [19], [21] [22] tetapi ada beberapa peneliti membangun sebuah model dari awal [3], [14], [15], [18], [20]. Kurangnya ukuran sampel yang memadai dalam suatu kumpulan data akan membuat kinerja dari model CNN kurang optimal. Oleh karena itu peneliti cenderung menggunakan teknik augmentasi data untuk memperbanyak kumpulan data. Augmentasi data yaitu penambahan dataset tanpa memerlukan citra baru dengan cara memanipulasi citra itu sendiri. Teknik augmentasi yang digunakan rata-rata masih sederhana seperti memutar citra, menggeser citra, dan membalikkan citra [3], [12], [15], [20]. Melihat adanya *gaps* penelitian terdahulu maka dapat menerapkan teknik augmentasi yang berbeda. Seperti DCGAN untuk menambah citra MRI seperti yang sudah diterapkan oleh peneliti [21], [22]. Tetapi masih sedikit penelitian yang menggunakan DCGAN untuk augmentasi data dalam citra MRI tumor otak dengan tujuan mengklasifikasikan tumor otak. Maka penelitian sekarang mencoba melakukan klasifikasi tumor otak dengan menggunakan metode CNN dan ditambahkan metode DCGAN sebagai augmentasi data. Citra yang dihasilkan secara artifisial nantinya dapat mengatasi keterbatasan kumpulan data medis yang kecil hingga batas yang wajar dan membantu model CNN untuk mengeksplorasi varian dalam kumpulan data.

2.2 Landasan Teori

2.2.1 Magnetic Resonance Imaging

Tumor otak termasuk jenis tumor yang paling mengancam di suruh dunia. *Glioma*, *meningioma*, dan *pituitary* (ditunjukkan pada Gambar 2. 1) merupakan tumor otak yang paling umum. *Magnetic Resonance Imaging* (MRI) adalah suatu teknik pencitraan medis dalam pemeriksaan diagnostik radiologi, yang menghasilkan rekaman citra potongan penampang tubuh atau organ manusia dengan menggunakan medan magnet dan resonansi getaran terhadap inti atom hidrogen. MRI mengukur sinyal frekuensi tinggi yang dipancarkan oleh atom hidrogen setelah menerapkan gelombang elektromagnetik (frekuensi radio), melokalisasi sinyal menggunakan gradien magnetik yang bervariasi secara spasial [23].

Meskipun ada banyak modalitas pencitraan medis yang dapat digunakan untuk membedakan fitur tumor otak, MRI adalah modalitas pencitraan medis yang paling umum digunakan karena keunggulannya yaitu menawarkan informasi lebih rinci tentang struktur otak dibandingkan CT atau USG.



Gambar 2. 1 Contoh Citra MRI (a) *Glioma*, (b) *Meningioma*, dan (c) *Pituitary*

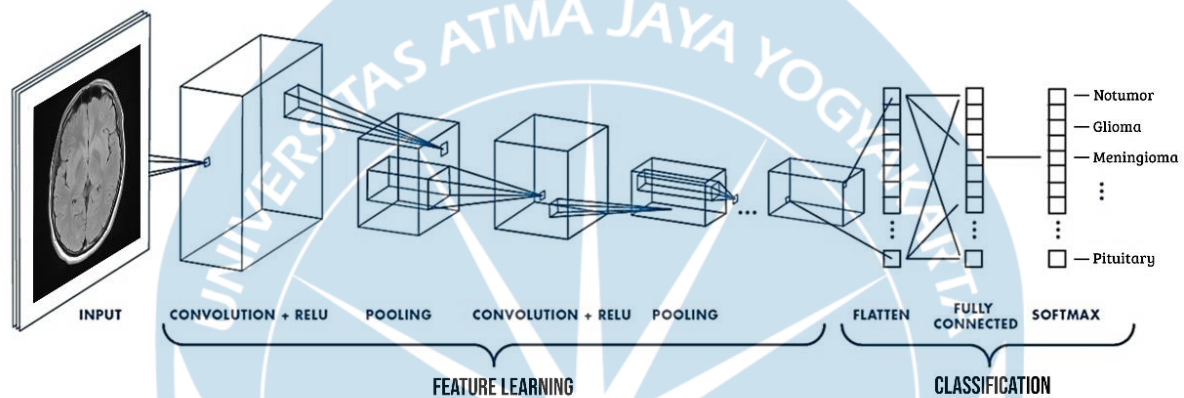
MRI berisi informasi berharga mengenai jenis, ukuran, bentuk, dan posisi tumor otak tanpa membuat pasien terkena radiasi ion yang berbahaya [22]. Dengan bantuan MRI, tumor otak dapat dibedakan menjadi beberapa tingkat. MRI dapat dianggap salah satu teknik yang paling maju untuk mendiagnosis dan mengevaluasi karakteristik tumor otak.

2.2.2 Image Classification

Citra dapat diklasifikasikan sebagai kelas tertentu berdasarkan pada data masa lalu dengan cara mengekstrak fitur. Dengan munculnya CNN sebagai salah satu metode klasifikasi citra dianggap sebagai teknologi terobosan dan CNN telah mencapai kinerja yang sangat baik pada bidang visual sekala besar [24], [25]. Untuk mengevaluasi algoritma klasifikasi citra ada banyak kumpulan data yang dapat digunakan seperti CIFAR-10/100 [26], Caltech-101/256 [27], dan ImageNet [28]. ImageNet sendiri terdiri dari 15.000.000 lebih citra dengan resolusi tinggi yang dibagi menjadi 22.000 kategori. Dari tahun 2010 hingga sekarang, kompetisi klasifikasi citra *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)* merupakan acara yang digelar setiap tahun untuk mengevaluasi berbagai macam algoritma untuk klasifikasi citra. Pada tahun 2018 pemenang ILSVRC Liu et al. [29] menunjukkan bahwa metode yang diusulkan lima kali lebih efisien dan delapan kali lebih cepat dalam hal komputasi. Keberhasilan CNN dikaitkan dengan pembelajaran fitur yang kuat. Metode klasifikasi CNN merupakan proses pembelajaran *end-to-end*, hanya citra asli yang dimasukkan, proses pelatihan dan prediksi dilakukan dalam jaringan, dan akhirnya menghasilkan output.

2.2.3 Convolution Neural Network

Convolutional Neural Network (CNN) adalah salah satu metode DL yang paling populer untuk pengenalan pola [30], [31] dan sering digunakan dalam bidang *computer vision* [32], seperti klasifikasi [33], deteksi [34], dan segmentasi [35]. Sebuah arsitektur CNN terdiri dari neuron, di mana setiap neuron memiliki bobot dan bias yang dapat dipelajari. CNN memiliki lapisan *input*, lapisan *output*, dan beberapa lapisan tersembunyi, di mana lapisan tersembunyi terdiri dari *convolutional layer*, *pooling layer*, dan *fully connected layer*. CNN mengekstrak fitur dari gambar melalui pembelajaran berulang, dan kemudian menghasilkan *feature map* [36]. Arsitektur CNN secara umum dapat dilihat pada Gambar 2. 2 di bawah.



Gambar 2. 2 Arsitektur CNN

Pada Gambar 2. 2 CNN dikelompokkan menjadi dua bagian besar yaitu *feature learning* dan *classification*. *Feature learning* adalah bagian untuk mengekstrak fitur penting dari sebuah *input* sedangkan *classification* yang berhubungan dengan *output* setelah melakukan proses *learning*.

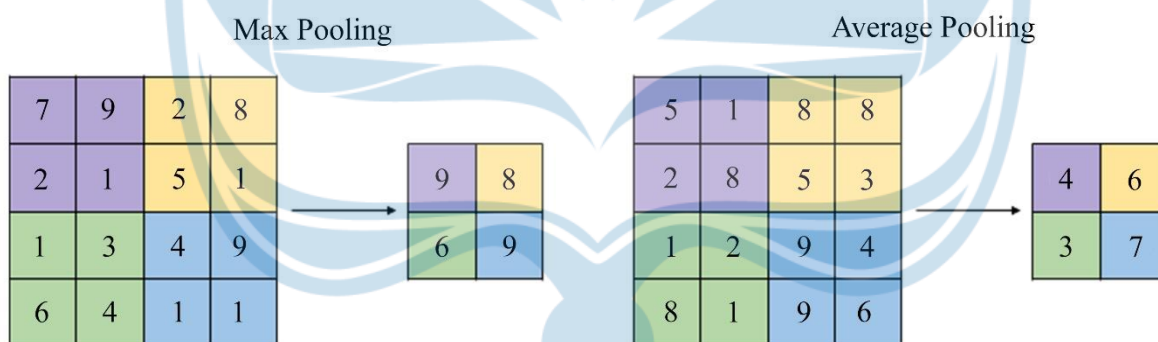
2.2.3.1 Convolutional Layer

Convolutional layer adalah lapisan utama dari CNN yang menentukan *output* dari *input* yang terkait. Pada proses ini memanfaatkan sebuah kernel, di mana ukuran matriks kernel jauh lebih kecil daripada ukuran matriks *input* [37]. Pada proses awal pelatihan kernel awalnya di inialisasi dengan nilai acak yang ditetapkan sebagai bobot kernel. Selanjutnya bobot akan disesuaikan untuk setiap periode pelatihan sehingga kernel dapat mengekstrak fitur-fitur penting. Nantinya kernel akan bergeser ke seluruh bagian *input* dan melakukan operasi konvolusi. Hasil tersebut berupa *feature maps* yang digunakan untuk *input layer* berikutnya. *Output* yang dihasilkan dapat dipengaruhi oleh sebuah hyperparameter. Hyperparameter yang mempengaruhi *output* adalah jumlah kernel pada satu lapisan, *stride*, dan *zero padding*.

Stride merupakan sebuah parameter yang mengontrol berapa jumlah pergeseran sebuah kernel. Jika nilai *stride* adalah dua, kernel akan bergeser sebanyak dua piksel secara horizontal dan kemudian secara vertikal. *Zero padding* merupakan hyperparameter yang menentukan jumlah piksel yang berisi angka nol yang ditambahkan di setiap sisi dari *input*. Bertujuan untuk mengatur dimensi *output* agar tetap sama seperti dimensi *input*. Sehingga ekstraksi *feature* bisa lebih mendapatkan informasi penting [38].

2.2.3.2 Pooling Layer

Pooling layer adalah proses mereduksi *feature maps* secara spasial atau mengurangi jumlah parameter dengan operasi *down sampling* tetapi menyimpan informasi yang paling penting [5]. *Pooling layer* memiliki fungsi yang dapat mempercepat operasi jaringan dan meningkatkan kinerja dari jaringan convolutional [30]. Beberapa metode *pooling* yang dapat digunakan seperti *tree pooling*, *gated pooling*, *average pooling*, *min pooling*, *max pooling*, *global average pooling* (GAP), dan *global max pooling*. Metode yang sering digunakan adalah *max pooling*, *average pooling*, dan *min pooling* [39]. Gambar 2. 3 merupakan ilustrasi dari *max pooling* dan *average pooling*.



Gambar 2. 3 Ilustrasi Max Pooling dan Average Pooling

Seperti pada Gambar 2. 3 contoh penggunaan *max pooling* adalah mengambil nilai terbesar dari proses konvolusi sedangkan untuk *average pooling* menjumlahkan nilai rata-rata dari hasil poses konvolusi.

2.2.3.3 Activation Function

Fungsi aktivasi diwakili oleh persamaan matematika yang mendefinisikan *output* dari neuron berdasarkan *input* yang diberikan. Fungsi aktivasi dianggap sebagai salah satu parameter penting dalam arsitektur pembelajaran mendalam yang menentukan output, akurasi,

dan efisiensi model yang dilatih [30]. Contoh fungsi aktivasi yang dapat digunakan seperti, Sigmoid, Tanh, ReLU, Leaky ReLU, Exponential Linear Unit, Swish, dan SoftMax [40].

Sigmoid: Input dari fungsi aktivasi Sigmoid adalah bilang *real*, dimana output akan dibatasi antara nol dan satu. Biasanya digunakan untuk klasifikasi biner.

Tanh: Fungsi aktivasi Tanh mirip dengan fungsi aktivasi Sigmoid, hanya saja output akan dibatasi antara -1 dan 1 .

ReLU: Merupakan fungsi aktivasi yang paling umum digunakan dalam CNN. Fungsi ini melakukan *thresholding* dengan nilai nol terhadap nilai piksel pada *input* citra. Fungsi ini akan mengubah seluruh nilai input ke bilangan positif dengan begitu beban komputasi menjadi lebih rendah.

Leaky ReLU: Merupakan jenis fungsi aktivasi berdasarkan ReLU, tetapi memiliki kemiringan kecil untuk nilai negatif. Leaky ReLU sering digunakan dalam melatih GAN.

2.2.3.4 Fully Connected Layer

Fully connected layer merupakan sebuah lapisan di mana setiap neuron terhubung ke semua neuron dari lapisan sebelumnya. Lapisan ini biasanya terletak di akhir setiap arsitektur CNN karena tujuannya adalah untuk meringkas semua fitur yang telah diekstraksi oleh lapisan sebelumnya untuk menghasilkan *output* [31]. Lapisan ini menggunakan fungsi aktivasi *SoftMax* di lapisan *output* untuk tujuan klasifikasi. Fungsi *SoftMax* menghitung probabilitas setiap kelas menggunakan fungsi eksponensial yang dinormalisasi dan mengeluarkan kelas dengan probabilitas tertinggi sebagai hasil klasifikasi.

2.2.3.5 CNN Architecture

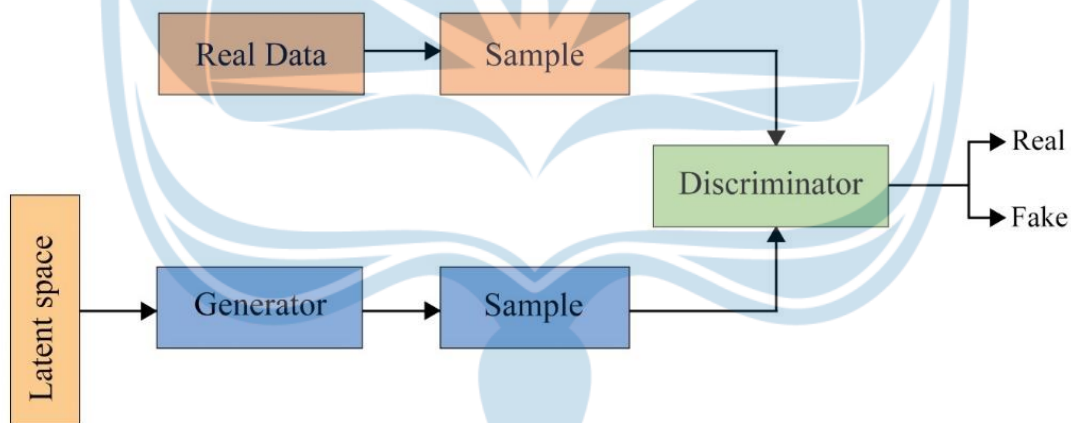
CNN memiliki beberapa arsitektur yang sudah dilatih sebelumnya salah satunya adalah MobileNet yang dikembangkan oleh Howard et al. [41]. Tidak seperti arsitektur lain yang memerlukan biaya komputasi yang tinggi dan menghasilkan model dengan ukuran yang sangat besar, berbeda dengan arsitektur MobileNet dengan arsitektur ini dapat mengurangi ukuran model dan mengurangi biaya komputasi. Dengan menggunakan konvolusi yang dapat dipisahkan secara mendalam untuk membangun jaringan yang dalam. Dua hyperparameter global diperkenalkan untuk mengembangkan model dengan penyesuaian akurasi dan latensi.

Arsitektur MobileNet terdapat versi 2 yaitu arsitektur MobileNetV2 yang dikembangkan oleh Sandler et al. [42] dengan basis yang sama seperti arsitektur MobileNetV1, MobileNetV2 menyertakan dua fitur baru, yaitu pertama beberapa batasan dapat muncul secara

linier antar lapisan, kedua adalah pengembangan *shortcut* antara *bottlenecks* [43]. Arsitektur MobileNetV1 konvolusi *pointwise* harus menggendakan jumlah *channel* atau mempertahankannya tetap sama. Untuk di MobileNetV2 konvolusi *pointwise* melakukan sebaliknya, konvolusi *pointwise* membuat jumlah *channel* lebih kecil. Lapisan ini sering dikenal sebagai lapisan proyeksi, karena lapisan ini memproyeksikan data dengan jumlah dimensi (*channels*) yang tinggi ke dalam tensor sambil menurunkan dimensi [44].

2.2.4 GAN

Generative Adversarial Nets diusulkan oleh Goodfellow et al. [45] dengan tujuan sebagai metode augmentasi lain untuk memperbaiki ketidakseimbangan dalam sebuah dataset yang jumlahnya sedikit. GAN telah terbukti sangat efektif untuk pembuatan citra dalam tugas *computer vision* yang berbeda [46]. Arsitektur GAN mempelajari fitur tanpa pengawasan dengan proses pembelajaran yang kompetitif dan dapat belajar dari variasi data yang sangat kecil. Gambar 2. 4 adalah arsitektur GAN secara umum.



Gambar 2. 4 Arsitektur GAN

Berdasarkan Gambar 2. 4 GAN adalah arsitektur pembelajaran mendalam yang terdiri dari dua jaringan, *generator* (G) dan *discriminator* (D), yang bersaing satu sama lain dan dilatih saat yang sama tanpa pengawasan. *Generator* mencoba menghasilkan sampel dari *latent space* yang mirip dengan data pelatihan sedangkan *discriminator* bertujuan untuk membedakan antara sampel asli dan yang dihasilkan oleh *generator* [47].

2.2.4.1 GENERATOR

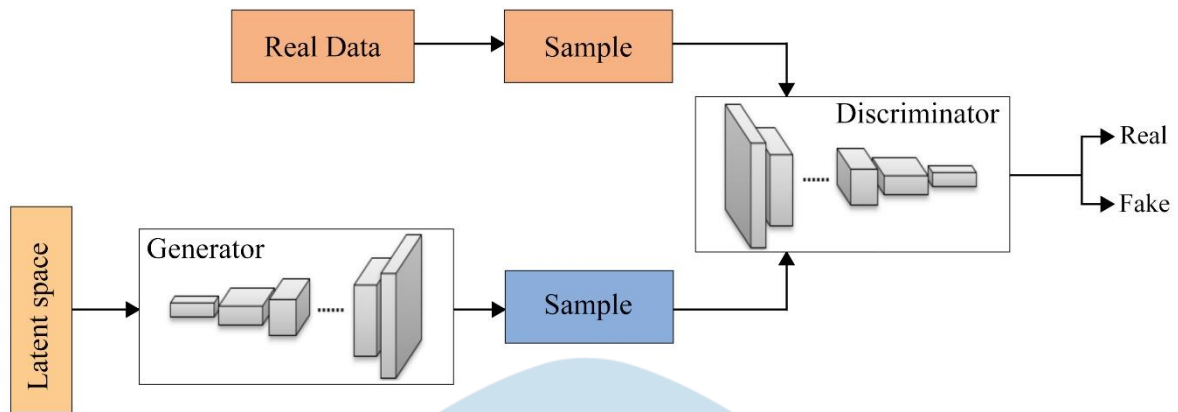
Generator (G) berfungsi untuk membuat data palsu yang akan dimasukkan ke dalam *discriminator* (D). Tujuan dari *generator* adalah berusaha agar citra yang dihasilkan dapat diklasifikasikan sebagai data asli oleh *discriminator*. Dalam proses pelatihan *generator* akan mengambil *noise* secara acak untuk dijadikan *input* awal pelatihan. Selanjutnya *generator* akan mengubah *noise* tersebut menjadi data palsu. Berikutnya *discriminator* akan mengklasifikasikan antara data asli dan data palsu yang telah dibuat.

2.2.4.2 DISCRIMINATOR

Discriminator (D) mempunyai tujuan untuk membedakan citra asli dengan citra yang dibuat oleh *generator*. Untuk melatih *discriminator*, dibutuhkan dua jenis data, yang pertama dibutuhkannya data asli yang akan mau dihasilkan sebagai citra sintesis. *Discriminator* akan menganggap dataset ini sebagai contoh positif dalam membedakan citra. Kedua adalah data palsu yang dihasilkan oleh *generator* nantinya *discriminator* akan menganggap dataset ini sebagai contoh negatif dalam membedakan citra.

2.2.4.3 DCGAN

Deep Convolutional Generative Adversarial Network (DCGAN) [48] merupakan peningkatan untuk pertama kalinya dari arsitektur GAN. DCGAN dapat menghasilkan citra yang lebih baik dan lebih stabil dalam proses pelatihan [49]. DCGAN memiliki prinsip yang sama dengan GAN, tetapi arsitektur DCGAN menggunakan lapisan *convolutional transpose* pada *generator* (G) dan lapisan *convolutional* pada *discriminator* (D). DCGAN memiliki dua fitur penting, yaitu *BatchNorm* untuk mengatur skala fitur yang diekstraksi, dan *LeakyReLU* untuk mencegah hilangnya gradien. DCGAN mengganti semua *max pooling* dengan *convolutional stride* dan menggunakan *transposed convolutional* untuk *up sampling*. Selain itu DCGAN juga menggantikan *fully connected layer* dengan *batch normalization*. DCGAN juga menggunakan *Rectified Linear Unit* (ReLU) di *generator* kecuali untuk *output* menggunakan *Tanh* sedangkan untuk *discriminator* menggunakan *LeakyReLU*. Gambar 2. 5 di bawah merupakan arsitektur DCGAN.



Gambar 2. 5 Arsitektur DCGAN

Generator (G) memproses *input noise* melalui serangkaian lapisan *convolutional transpose* dan *up sampling* untuk menghasilkan sampel palsu, *discriminator* (D) akan menilai apakah gambar *input* adalah sebenarnya melalui serangkaian *sub sample*. Jika sampel tersebut dinyatakan benar maka probabilitas keluarannya adalah 1, tetapi jika sampel tersebut dinyatakan salah maka probabilitas keluarannya adalah 0 [50].

2.2.5 Evaluation Matric

Evaluation matric adalah sebuah gambar atau tabel yang digunakan untuk menggambarkan kinerja dari sebuah model. *Evaluation matric* digunakan untuk mengoptimalkan algoritma klasifikasi selama tahap pelatihan. *Evaluation matric* juga dapat digunakan untuk tahap pengujian model menggunakan data baru yang tidak termasuk dalam proses pelatihan. *Confusion matrix* adalah ringkasan hasil prediksi pada masalah klasifikasi dengan menghitung jumlah prediksi yang benar dan salah pada masing-masing kelas. Gambar 2. 6 di bawah merupakan *confusion matrix*.

		Actual	
		Positive	Negative
Predicted	Positive	TP (True Positive)	FP (False Positive)
	Negative	FN (False Negative)	TN (True Negative)

Gambar 2. 6 Confusion Matrix

Gambar 3 menunjukkan *confusion matrix* dengan empat istilah yaitu TP, TN, FP, dan FN. TP merupakan data positif yang diprediksi benar, TN adalah data negatif yang diprediksi benar, FP adalah data negatif namun diprediksi sebagai data positif, dan FN adalah data positif namun diprediksi sebagai data negatif. Untuk menentukan baik atau tidaknya suatu model dapat dilihat dari parameter seperti *accuracy*, *sensitivity*, *precision*, *specificity*, dan *F1 score*.

1. *Accuracy*: Merupakan rasio prediksi benar yang benar dengan jumlah total sample yang dievaluasi.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2 - 1)$$

2. *Recall*: Merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif.

$$Recall = \frac{TP}{TP + FN} \quad (2 - 2)$$

3. *Precision*: Merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif.

$$Precision = \frac{TP}{TP + FP} \quad (2 - 3)$$

4. *F1 Score*: Menghitung rata-rata antara tingkat *recall* dan *precision*.

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2 - 4)$$