

ABSTRACT

Proper facial skin care requires an understanding of skin types, such as dry, normal, and oily skin. However, manual identification of skin types is still often influenced by subjective assessments, resulting in inconsistent accuracy levels. This study develops a facial skin type classification system based on digital image processing using the Convolutional Neural Network (CNN) method with the MobileNetV3 architecture. The model was developed by applying a transfer learning approach to extract facial image features effectively, even with limited training data. The training process was conducted using facial images with a resolution of 224×224 pixels and three output classes, namely dry, oily, and normal skin. Based on the testing results, the developed model achieved an accuracy of 99.2% with a low loss value, indicating good classification performance. Furthermore, the model was implemented in a web-based application using the flask framework, enabling users to detect facial skin types and obtain skincare recommendations according to the classification results. Based on the obtained results, the CNN-based facial skin type classification model can be effectively utilized as a supporting system for determining appropriate facial skin care.

Keyword : CNN, MobileNetV3, facial skin type classification

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada era digital saat ini, perkembangan teknologi informasi dan kecerdasan buatan (*Artificial Intelligence*) telah memberikan dampak yang sangat besar terhadap berbagai bidang kehidupan, termasuk bidang kecantikan dan perawatan kulit. Analisis kulit wajah yang sebelumnya dilakukan secara manual kini dapat dibantu oleh sistem berbasis pengolahan citra digital, yang memungkinkan proses identifikasi jenis kulit dilakukan secara objektif, cepat, dan akurat. Hal ini menunjukkan bahwa teknologi dapat berperan penting dalam meningkatkan efisiensi serta keakuratan proses analisis kulit wajah (Dewi & Prasetyo, 2022).

Dalam bidang pengolahan citra digital dan pembelajaran mesin, analisis jenis kulit wajah umumnya dilakukan dengan menggunakan pendekatan berbasis ekstraksi fitur manual. Beberapa metode yang sering diterapkan antara lain *Gray Level Co-occurrence Matrix* (GLCM), *Histogram of Oriented Gradient* (HOG), serta *Support Vector Machine* (SVM) yang digunakan untuk mengenali perbedaan pola warna dan tekstur pada citra kulit wajah. Meskipun pendekatan tersebut mampu memberikan hasil klasifikasi yang cukup baik, performanya masih sangat bergantung pada kondisi pencahayaan, posisi wajah, serta memerlukan proses ekstraksi fitur yang kompleks dan memakan waktu cukup lama. Kondisi ini menunjukkan perlunya pengembangan metode yang lebih adaptif dan efisien dalam mengenali pola citra secara otomatis tanpa ketergantungan tinggi terhadap faktor lingkungan (Sari & Rahmawati, 2019).

Seiring dengan perkembangan teknologi *deep learning*, muncul pendekatan baru dalam bidang klasifikasi citra wajah, yaitu *Convolutional Neural Network* (CNN). CNN memiliki kemampuan unggul dalam mengenali pola visual seperti tekstur, warna, dan kontur wajah secara otomatis melalui proses *feature extraction* tanpa harus dilakukan secara manual. Menurut Chen, Zhang, dan Xu (2020), penggunaan CNN dalam klasifikasi jenis kulit wajah mampu

meningkatkan akurasi hingga lebih dari 95% pada *dataset* dengan kondisi pencahayaan yang beragam. Berdasarkan hasil penelitian tersebut, penggunaan CNN menjadi pilihan yang tepat dalam pengembangan sistem klasifikasi jenis kulit wajah karena kemampuannya dalam mengenali fitur citra secara mendalam dan menyeluruh.

Sehingga pada penelitian ini menerapkan metode CNN untuk melakukan klasifikasi jenis kulit wajah berdasarkan citra digital. Tahapan yang dilakukan meliputi pengumpulan dan pelabelan *dataset* kulit wajah sesuai kategori umum normal, kering, dan berminyak, pelatihan model CNN, serta evaluasi kinerja model berdasarkan tingkat akurasi. CNN dinilai mampu mengekstraksi pola visual kompleks dari citra wajah sehingga menghasilkan klasifikasi yang lebih konsisten pada berbagai kondisi pencahayaan (Chen, 2020). Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi jenis kulit wajah yang akurat dan adaptif, serta menjadi dasar pengembangan rekomendasi produk *skincare* berbasis kecerdasan buatan yang sesuai dengan kondisi kulit pengguna (Dewi & Prasetyo, 2022).

1.2 Tujuan Penelitian

Tujuan dari penelitian ini yaitu dapat membangun model klasifikasi jenis kulit wajah menggunakan *Convolutional Neural Network* (CNN).

1.3 Manfaat Penelitian

Adapun manfaat dari penelitian ini, yaitu:

1. Memberikan kontribusi dalam pengembangan sistem klasifikasi citra berbasis *deep learning* di bidang kecantikan dan kesehatan kulit.
2. Membantu mengenali jenis kulit dan mengurangi risiko kesalahan dalam memilih produk *skincare* karena salah mengenali jenis kulit.

BAB II

DASAR TEORI

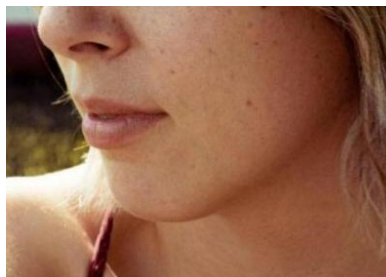
2.1 Jenis - Jenis Kulit Wajah

Kulit adalah lapisan paling luar pada tubuh yang berfungsi melindungi jaringan di bawahnya dari berbagai pengaruh lingkungan, termasuk paparan sinar ultraviolet (UV). Paparan sinar UV secara terus-menerus dapat memicu berbagai masalah pada kulit, seperti munculnya kerutan, flek hitam, hingga risiko kanker kulit. Selain berfungsi sebagai pelindung tubuh, kulit wajah juga memiliki nilai estetika dan cenderung lebih sensitif dibandingkan bagian kulit lainnya, sehingga perawatannya perlu diperhatikan dengan baik (Sulastri & Chaerunisaa, 2016).

Perawatan kulit wajah dapat dilakukan secara rutin, salah satunya melalui penggunaan produk kosmetik yang sesuai. Dalam memilih produk perawatan, kondisi dan jenis kulit pengguna menjadi aspek penting yang harus dipertimbangkan. Beberapa faktor seperti usia, pola makan, gaya hidup, serta keseimbangan hormon dapat memengaruhi karakteristik kulit seseorang, baik kulit kering, normal, maupun berminyak. Oleh karena itu, pemilihan produk kosmetik sebaiknya disesuaikan dengan kebutuhan kulit masing-masing individu agar perawatan yang dilakukan lebih optimal (Fadila, 2022).

1. Kulit Kering

Kulit kering merupakan kondisi kulit yang memiliki kadar minyak alami atau sebum dalam jumlah rendah. Produksi sebum menyebabkan kelembapan kulit lebih cepat berkurang sehingga kulit terasa kering dan kasar. Ciri-ciri kulit kering antara lain permukaan kulit tidak berminyak, tekstur kulit relatif tipis, serta lebih mudah mengalami kekusaman dan iritasi (Shinta, 2023).



Gambar 2.1 Kulit Kering

2. Kulit Berminyak

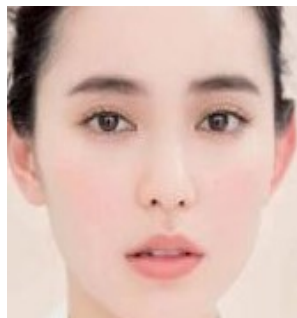
Kulit berminyak merupakan kondisi kulit yang terjadi akibat produksi minyak alami yang berlebihan oleh kelenjar *sebaceous*. Kondisi ini umumnya mulai terlihat saat masa pubertas karena adanya pengaruh hormon androgen yang meningkatkan aktivitas kelenjar minyak. Ciri-ciri kulit berminyak dapat dikenali dari tampilan wajah yang mengilap, pori-pori yang terlihat lebih besar terutama pada area hidung, pipi, dan dagu, serta kulit yang lebih mudah mengalami jerawat atau komedo (Shinta, 2023).



Gambar 2.2 Kulit Berminyak

3. Kulit Normal

Kulit normal adalah jenis kulit yang memiliki keseimbangan kadar minyak dan kelembapan sehingga kondisi kulit cenderung stabil dan mudah dirawat. Pada jenis kulit ini, kelenjar minyak bekerja secara optimal dengan menghasilkan sebum dalam jumlah yang cukup, tidak terlalu banyak maupun terlalu sedikit. Karakteristik kulit normal antara lain permukaan kulit terasa halus dan segar, tidak tampak terlalu berminyak, jarang mengalami jerawat, serta memiliki pori-pori yang relatif kecil. Selain itu, kulit normal umumnya lebih mudah beradaptasi dengan berbagai produk perawatan kulit (Shinta, 2023).



Gambar 2.3 Kulit Normal

2.2 Deep Learning

Artificial Intelligence atau kecerdasan buatan merupakan bidang teknologi yang telah banyak dikembangkan untuk menciptakan sistem yang mampu meniru kemampuan berpikir manusia. Perkembangan AI mendorong terciptanya berbagai metode yang dapat membantu komputer memahami, mempelajari, dan menyelesaikan suatu permasalahan secara otomatis (Prahartiningsyah & Kurniawan, 2021). Salah satu cabang penting dalam kecerdasan buatan adalah *deep learning*, yaitu metode pembelajaran yang berfokus pada pemanfaatan data dan algoritma untuk membangun proses belajar menyerupai cara kerja otak manusia (Anggeli, 2021).

Deep learning bekerja dengan memanfaatkan jaringan saraf tiruan yang memiliki banyak lapisan pemrosesan data. Setiap lapisan bertugas melakukan ekstraksi dan transformasi informasi sehingga sistem dapat mengenali pola tertentu pada data yang dipelajari. Hubungan antar-neuron dalam jaringan tersebut memungkinkan model melakukan proses klasifikasi, analisis, dan pengambilan keputusan secara lebih efektif pada berbagai permasalahan pembelajaran mesin. Dalam bidang pengolahan citra, penggunaan *deep learning* telah memberikan perkembangan yang signifikan, terutama pada proses klasifikasi gambar berbasis CNN. Metode ini mampu mempelajari representasi data secara otomatis sehingga dapat membantu penyelesaian masalah yang lebih kompleks dan meningkatkan performa sistem pengenalan citra (Arifin, 2021).

2.3 Pengolahan Citra Digital

Pengolahan citra digital merupakan proses pengolahan data visual dalam bentuk citra menggunakan sistem komputer. Citra digital dapat dipahami sebagai representasi gambar dua dimensi yang tersusun atas sejumlah piksel dan disimpan dalam bentuk nilai numerik tertentu (Wibowo & Nugroho, 2021). Agar sebuah citra dapat diproses secara komputasional, gambar perlu dikonversi menjadi data digital diskrit melalui proses digitalisasi. Salah satu tantangan yang banyak diteliti dalam bidang ini adalah deteksi objek otomatis, yakni proses mengidentifikasi serta menentukan posisi objek penting pada sebuah gambar maupun video.

Kemampuan tersebut memiliki peranan penting dalam berbagai penerapan berbasis *computer vision*, seperti sistem pengawasan keamanan (CCTV), kendaraan otonom, hingga teknologi pendukung diagnosis medis berbasis komputer (Rahmawati 2022).

2.3.1 Piksel

Piksel biasanya disebut juga sebagai *picture element* merupakan unit atau elemen terkecil yang menyusun sebuah citra digital. Piksel dapat dipahami sebagai titik-titik atau kotak-kotak kecil yang membentuk suatu gambar. Jumlah dan kualitas piksel pada citra sangat berpengaruh terhadap tingkat ketajaman serta detail informasi yang dapat ditampilkan oleh citra tersebut. Jumlah piksel yang terdapat pada citra berkaitan dengan resolusi gambar, di mana semakin tinggi resolusi suatu citra maka semakin banyak piksel yang dimiliki dan semakin besar pula informasi visual yang dapat ditampilkan. Selain resolusi, kualitas citra juga dipengaruhi oleh *color depth*, yaitu jumlah variasi warna yang dapat direpresentasikan oleh setiap piksel. Oleh karena itu, ukuran dan kualitas suatu citra digital ditentukan oleh resolusi serta kedalaman warna yang digunakan.

Citra digital dapat direpresentasikan dalam suatu sistem koordinat untuk memudahkan proses pengolahan dan pemetaan piksel. Pada citra dengan ukuran $M \times N$, sistem koordinat umumnya dimulai dari sudut kiri atas yang dinyatakan sebagai titik (0,0). Setiap piksel pada koordinat tertentu memiliki nilai yang menunjukkan informasi warna atau intensitas cahaya. Pada citra *grayscale*, nilai piksel merepresentasikan tingkat keabuan dari hitam hingga putih. Sementara itu, pada *color image*, setiap piksel terdiri atas beberapa komponen warna, yang umumnya menggunakan model RGB (*Red, Green, Blue*) untuk menghasilkan berbagai kombinasi warna pada citra.

2.3.2 Jenis Citra Digital

Menurut Kadir dan Susanto (2013), citra digital berdasarkan nilai pikselnya dapat dikelompokkan menjadi tiga jenis, yaitu citra biner, citra grayscale, dan citra berwarna.

1. Citra biner

Citra biner merupakan citra yang hanya memiliki dua kemungkinan nilai piksel, yaitu 0 dan 1. Nilai 0 biasanya merepresentasikan warna hitam, sedangkan nilai 1 merepresentasikan warna putih. Karena hanya memiliki dua tingkat warna, citra biner cukup direpresentasikan menggunakan 1 bit untuk setiap piksel. Dalam bidang pengolahan citra, citra biner banyak digunakan pada berbagai proses, seperti segmentasi dan operasi morfologi, karena mampu memisahkan objek dari latar belakang secara sederhana. Sebagai contoh, citra *grayscale* dapat dikonversi menjadi citra biner sehingga setiap piksel hanya memiliki nilai hitam atau putih (Susanto, 2013).



Gambar 2.4 Contoh Citra Biner (Pamungkas, 2023)

2. Citra grayscale

Citra grayscale merupakan jenis citra digital yang hanya memiliki satu kanal intensitas warna dengan rentang nilai dari hitam hingga putih. Setiap piksel pada citra ini merepresentasikan tingkat keabuan tertentu, di mana nilai yang lebih rendah menunjukkan warna yang lebih gelap dan nilai yang lebih tinggi menunjukkan warna yang lebih terang. Oleh karena itu, citra grayscale tidak mengandung informasi warna seperti pada citra RGB, melainkan hanya menyimpan informasi intensitas cahaya. Citra jenis ini banyak digunakan dalam pengolahan citra karena mampu menyederhanakan data tanpa menghilangkan informasi penting yang diperlukan untuk proses analisis. Tingkat keabuan yang dihasilkan membentuk gradasi warna dari hitam, abu-abu, hingga putih (Susanto, 2013).



Gambar 2.5 Contoh Citra *Grayscale* (Pamungkas, 2023)

3. Citra berwarna

Citra berwarna atau *color image* adalah citra digital yang tersusun atas piksel-piksel dengan informasi warna yang direpresentasikan oleh tiga komponen utama, yaitu merah (*Red*), hijau (*Green*), dan biru (*Blue*). Masing-masing komponen memiliki rentang intensitas nilai dari 0 hingga 255. Kombinasi ketiga komponen tersebut memungkinkan terbentuknya sekitar 16,58 juta variasi warna (255^3). Jumlah bit yang digunakan untuk menyatakan warna pada setiap piksel dikenal sebagai *color depth*. Secara umum, citra berwarna terdiri atas tiga lapisan matriks, yaitu lapisan merah (*R-layer*), lapisan hijau (*G-layer*), dan lapisan biru (*B-layer*), yang secara bersama-sama membentuk warna pada citra (Susanto, 2013).



Gambar 2.6 Contoh Citra RGB (Pamungkas, 2023)

2.4 Convolutional Neural Network (CNN)

CNN merupakan salah satu jenis jaringan saraf tiruan yang dirancang khusus untuk melakukan pengolahan dan analisis citra digital. Metode ini banyak digunakan dalam tugas klasifikasi, pengenalan, maupun pendeteksian objek pada gambar karena dapat mengenali pola visual secara efektif. CNN bekerja dengan menerapkan operasi konvolusi menggunakan kernel tertentu pada data citra sehingga sistem dapat mempelajari fitur-fitur penting dari gambar. Struktur CNN dirancang menyerupai mekanisme pengolahan visual pada manusia, sehingga model mampu memahami detail citra dengan lebih baik. Oleh karena itu, CNN umumnya diterapkan pada data yang memiliki bentuk dua dimensi, seperti citra digital (Suartika, 2016).

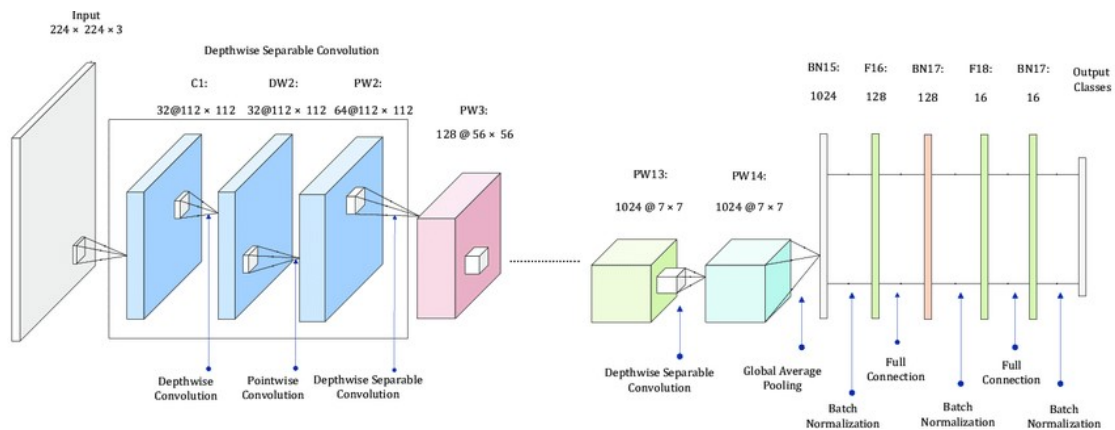
Dalam perkembangan *deep learning*, CNN menjadi salah satu metode yang paling populer karena memiliki kemampuan ekstraksi fitur otomatis dari citra masukan. Proses tersebut memungkinkan dimensi gambar diperkecil tanpa menghilangkan karakteristik utama dari objek pada gambar (Omori & Shima, 2020). Arsitektur CNN tersusun atas sejumlah neuron yang memiliki parameter berupa bobot dan bias. Setiap neuron menerima data masukan yang kemudian diproses menggunakan operasi matematika seperti *dot product* untuk menghasilkan keluaran pada jaringan (Zein Ersyad, 2020).

2.4.1 Arsitektur CNN

Kemampuan CNN untuk mengekstraksi fitur-fitur mendalam dari citra menjadikannya efektif dalam membedakan berbagai jenis kulit seperti normal, kering, dan berminyak. Hal ini sangat bermanfaat dalam pengembangan sistem rekomendasi produk perawatan kulit yang lebih personal dan akurat. Namun, tantangan utama dalam penerapan CNN terletak pada upaya mencapai akurasi tinggi dengan efisiensi komputasi yang optimal, terutama ketika model dijalankan pada perangkat dengan sumber daya terbatas. Dalam konteks tersebut MobileNetV3, sebagai pengembangan arsitektur CNN yang dirancang untuk efisiensi tinggi, menjadi solusi potensial. MobileNetV3 merupakan arsitektur model yang dikenal efisien karena memiliki ukuran yang relatif kecil dan

penggunaan memori yang hemat. Model ini memanfaatkan teknik *channel-wise filtering* untuk mengoptimalkan proses komputasi, sehingga menghasilkan keluaran yang ringan (Edel & Kapustin, 2022).

Arsitektur ini menggabungkan teknik *depthwise separable convolution* serta mekanisme penyesuaian otomatis melalui pengaturan *hyperparameter*, sehingga mampu mempercepat proses pelatihan dan inferensi tanpa mengorbankan akurasi. Dengan demikian, penelitian yang berfokus pada pengembangan dan penerapan MobileNetV3 ini dapat memberikan kontribusi penting dalam membangun sistem rekomendasi *skincare* yang efisien, cepat, dan memiliki tingkat akurasi tinggi (Hindarto, 2023).



Gambar 2.7 Arsitektur CNN MobileNetV3 (Pamungkas, 2023)

2.4.2 Komponen CNN

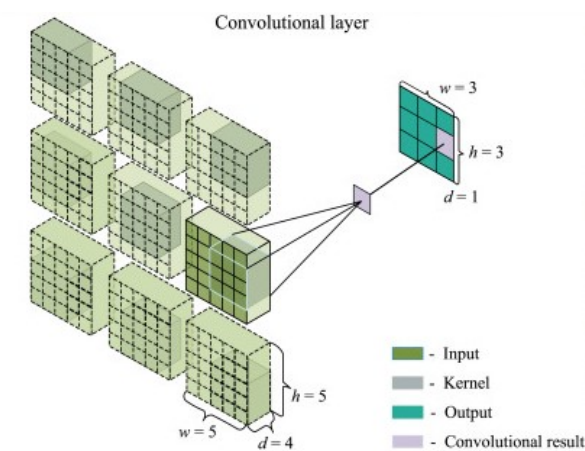
Arsitektur CNN terdiri dari beberapa *layer* yang saling terhubung dan bekerja secara berurutan untuk mengekstraksi fitur dari citra masukan. Setiap lapisan memiliki fungsi spesifik dalam proses pengolahan citra, yang secara keseluruhan memungkinkan model mengenali pola visual dengan tingkat kompleksitas yang meningkat (Purwono, 2022).

1. *Input Layer*

Lapisan *input* berfungsi menerima data citra dalam bentuk matriks piksel. Umumnya, citra direpresentasikan dalam format RGB (*Red, Green, Blue*) dengan ukuran tetap, misalnya 224×224 piksel, agar seluruh data masukan memiliki dimensi yang seragam sebelum diproses lebih lanjut (Weng, 2022).

2. Convolutional Layer

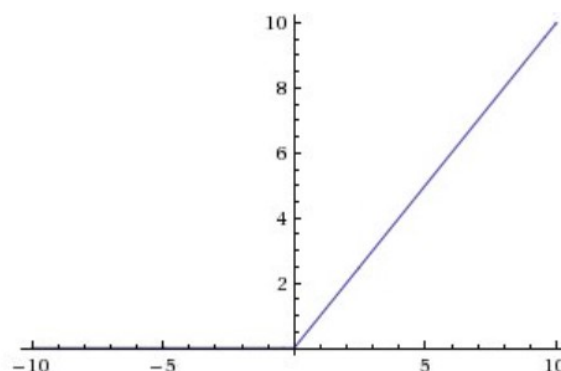
Lapisan ini menjadi bagian utama dalam arsitektur CNN yang bertugas melakukan proses ekstraksi fitur dari citra *input* melalui penerapan *kernel* atau filter tertentu. Setiap filter digeser melintasi citra dan melakukan operasi *dot product* dengan *patch* citra, menghasilkan *feature map* yang menggambarkan respons terhadap pola-pola seperti tepi, garis atau tekstur. Semakin banyak lapisan konvolusi yang ditumpuk, semakin kompleks fitur yang dapat dihasilkan, dari fitur rendah hingga fitur tinggi (Weng, 2022).



Gambar 2.8 Convolutional Layer (Changdong, 2023)

3. ReLU

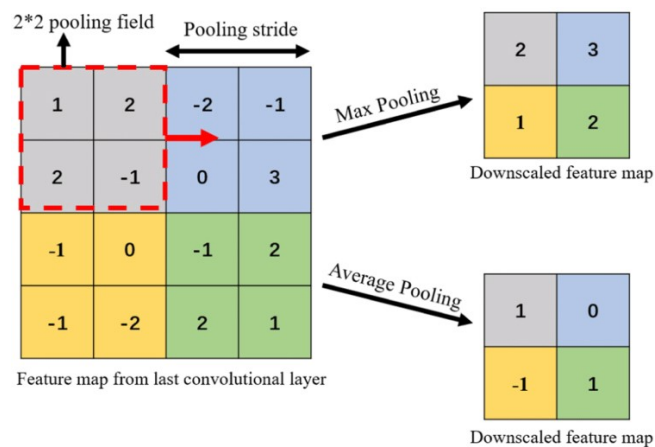
Setelah operasi konvolusi, hasilnya diteruskan ke fungsi aktivasi seperti *Rectified Linear Unit* (ReLU) yang mengubah nilai negatif menjadi nol dan mempertahankan nilai positif, sehingga menambahkan sifat non-linearitas ke dalam jaringan. Fungsi aktivasi memastikan bahwa jaringan dapat mempelajari pola yang tidak hanya linier, tetapi juga kompleks dan abstrak (Purwono, 2022).



Gambar 2.9 ReLU (Kartika, 2020)

4. Pooling Layer

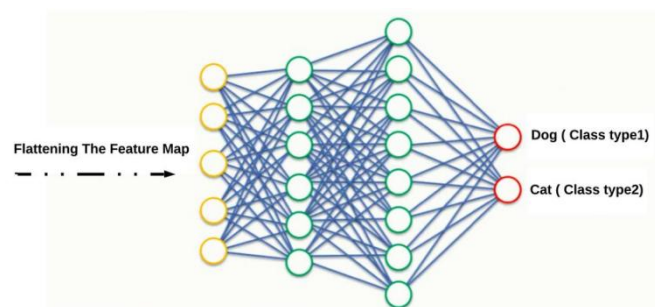
Pooling layer digunakan untuk memperkecil ukuran *feature map* sehingga proses komputasi menjadi lebih ringan dan kinerja model lebih efisien. Beberapa metode *pooling* yang sering diterapkan pada CNN antara lain *max pooling* dan *average pooling*. Pengurangan dimensi juga membantu model menjadi lebih *robust* terhadap perubahan translasi, rotasi, atau skala pada citra (Weng, 2022).



Gambar 2.10 Pooling Layer (Minfei, 2022)

5. Fully Connected Layer

Setelah fitur diekstraksi dan diperkecil dimensinya, fitur-fitur tersebut kemudian diratakan dan masuk ke lapisan *fully connected* yang menghubungkan semua neuron dari lapisan sebelumnya. Di sinilah keputusan klasifikasi diambil berdasarkan bobot-bobot yang telah dipelajari selama pelatihan. Lapisan ini sangat penting karena mengubah representasi fitur dalam bentuk vektor menjadi *output* kelas (Purwono, 2022).



Gambar 2.11 Fully Connected Layer (Farhana, 2021)

6. *Output Layer*

Lapisan terakhir menghasilkan keluaran dalam bentuk probabilitas untuk setiap kelas target. Umumnya menggunakan fungsi aktivasi seperti *softmax* untuk memastikan bahwa jumlah total probabilitas dari semua kelas adalah satu. Sebagai contoh, untuk klasifikasi jenis kulit wajah, keluaran bisa berupa probabilitas tiap kelas normal, kering, dan berminyak (Purwono, 2022).

2.5 Tensor Flow

Saat ini, kemajuan di bidang *deep learning* menjadi lebih mudah berkat ketersediaan banyak *library* dan API (*Application Program Interface*). Salah satu *library* yang dipakai adalah TensorFlow. TensorFlow berfungsi sebagai antarmuka yang memungkinkan perumusan algoritma *machine learning* dan eksekusi perintah. Ini dilakukan dengan memanfaatkan informasi yang ada mengenai suatu objek atau target yang dikenali, serta memungkinkan pembedaan antara satu objek dengan objek lainnya. Meskipun TensorFlow mendukung pelatihan model menggunakan baik CPU maupun GPU, implementasi yang dimaksud dalam konteks ini akan fokus pada pelatihan model dengan fitur CPU (Shukla, 2018).

TensorFlow memiliki fitur utama yang mencakup dukungan komprehensif untuk pemrograman jaringan saraf dalam dan teknik *machine learning* lainnya. *Library* ini unggul dalam mendefinisikan, mengoptimalkan, dan melakukan perhitungan matematis pada tensor. Selain itu, TensorFlow menawarkan efisiensi tinggi dalam penggunaan GPU, secara otomatis mengelola memori dan menentukan bagian komputasi yang optimal untuk dijalankan di GPU, dengan fleksibilitas menjalankan kode yang sama pada CPU atau GPU. TensorFlow menyediakan skalabilitas komputasi yang tinggi, menjadikannya ideal untuk mengelola kumpulan data besar di berbagai mesin (Taufiq, 2018).

2.6 Metode Evaluasi

Dalam penelitian yang melibatkan model *deep learning* seperti CNN, diperlukan metode evaluasi yang bertujuan untuk menilai sejauh mana model mampu melakukan klasifikasi secara akurat terhadap data uji. Metode evaluasi pada umumnya dilakukan dengan membandingkan hasil prediksi model dengan label sebenarnya menggunakan data validasi atau data uji. Melalui proses ini, diperoleh informasi mengenai performa model dalam hal akurasi, kesalahan prediksi, serta tingkat generalisasi. Evaluasi yang baik membantu peneliti menentukan apakah model perlu dilakukan *finetuning* atau pengaturan ulang parameter (Weng, 2022).

Untuk mendukung proses evaluasi tersebut, digunakan metrik evaluasi yang menjadi indikator utama dalam menilai kualitas model CNN. Metrik evaluasi merupakan alat ukur kuantitatif esensial yang dipakai untuk menilai efikasi dan kualitas sistem *Natural Language Processing*. Alat ukur ini memungkinkan penilaian seberapa baik suatu sistem NLP menjalankan tugasnya, seperti translasi otomatis, analisis opini, atau identifikasi entitas tertentu. Dalam dunia *Natural Language Processing*, metrik evaluasi sangat krusial karena menentukan keandalan, ketepatan, dan mutu luaran yang dihasilkan oleh model atau algoritma. Metrik ini memberikan standar baku untuk mengkomparasi performa berbagai sistem dan metode NLP, sehingga memandu peneliti dan praktisi dalam memilih solusi yang paling cocok untuk aplikasi yang berbeda. Beberapa metrik evaluasi yang sering digunakan dalam *Natural Language Processing* mencakup *precision*, *recall*, *F1 score*, dan *accuracy*.

1. *Precision*

Presisi mengukur seberapa banyak hasil prediksi positif yang benar-benar relevan. Presisi tinggi menunjukkan bahwa model jarang salah mengklasifikasikan data yang seharusnya negatif menjadi positif.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2.1)$$

dengan *TP* adalah *True Positive*, *TN* adalah *True Negative*, *FP* adalah *False Positive*, dan *FN* adalah *False Negative* (Purwono et al., 2022).