

BAB II TINJAUAN PUSTAKA

2.1 Studi Literatur

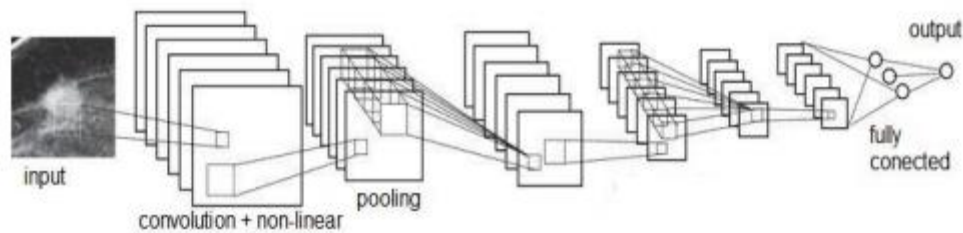
Pada sub bagian ini akan menjelaskan beberapa rincian dari penelitian sebelumnya yang digunakan sebagai pedoman pada penelitian yang sedang dilakukan. Adapun uraian hasil dari beberapa penelitian sebelumnya terlihat seperti pada tabel dibawah ini.

Tabel 2.1 Tabel studi literatur penelitian terdahulu

No	Penulis (tahun)	Judul	Dataset	Metode	Hasil akurasi
1.	M. Naji, S. Filali, K. Aarika er al. (2021)	Machine Learning Algorithms for Breast Cancer Prediction and Diagnosis	Wisconsin	SVM	97.2%
2.	A. Algarni, B. Aldahri, H. Alghamdi (2021)	Convolutional neural networks for breast tumor classification using structured features	Wisconsin	CNN	98%
3.	T. Assegie (2021)	An optimized K-Nearest neighbor based breast cancer detection	Wisconsin	KNN	94.35%

2.2 Convolutional Neural Network

Convolutional neural network (CNN) adalah sebagian dari pendekatan pembelajaran mendalam paling menonjol, membentuk banyak lapisan dengan cara yang ampuh[16]. Convolutional neural network (CNN) menyertakan sebuah ekstraktor fitur dalam prosesnya daripada mendesain secara manual, ekstraktor fitur sendiri terdiri dari jaringan saraf khusus[17]. Dengan begitu CNN memiliki keunggulan utama yakni merubah ekstraksi fitur manual menjadi otomatis. Dalam model yang diusulkan ini menggunakan hyperparameter

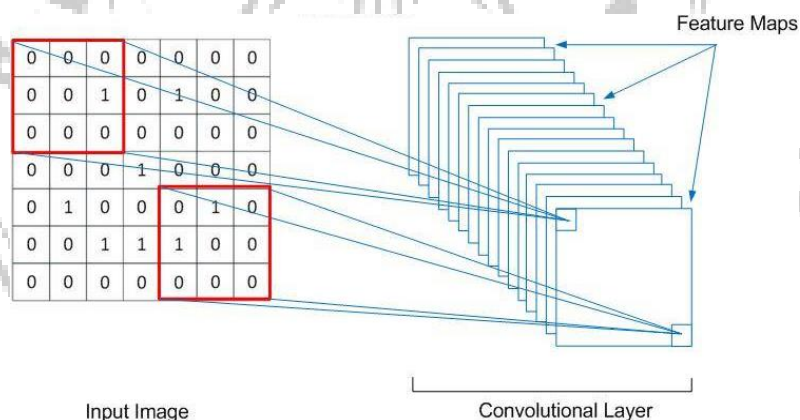


Gambar 2.1 struktur arsitektur CNN

berikut: tingkat dropout 0,5, epoch 75, batch size 32. CNN terdiri dari beberapa lapisan yakni: convolutional layer, pooling layer dan fully layer.

2.2.1 Convolutional layer

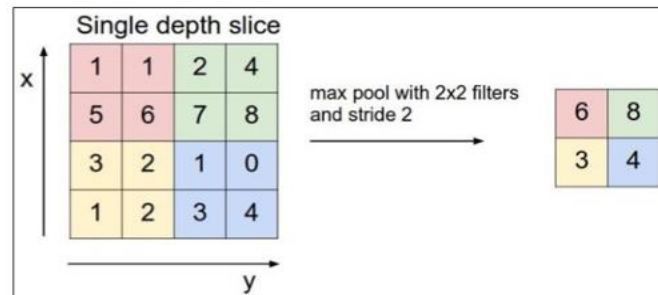
Convolutional layer akan membuahkan hasil yang telah dilakukan proses konvolusi, banyaknya hasil dari lapisan ini sama dengan jumlah pada proses konvolusi[18].



Gambar 2.2 Proses Convolutional layer

2.2.2 Pooling layer

Pooling layer merupakan proses pengumpulan dan pengurangan ukuran gambar karena penggabungan pixel-pixel terdekat dari area tertentu menjadi satu nilai sehingga mencegah overfitting[19].



Gambar 2.3 Proses Pooling Layer

2.2.3 Fully connected layer

Fully connected layer merupakan lapisan terakhir dari CNN yang sistematis menjumlahkan segala hasil dari fitur-fitur sebelumnya. Dalam lapisan ini seluruh elemen dari lapisan fitur sebelumnya akan diaplikasikan dalam perhitungan setiap bagian elemen dari setiap fitur output[20].

2.3 MobileNet

Desain dari mobilenet adalah kecil dan memiliki latensi rendah dan sangat cocok untuk system dan aplikasi yang mungkin memerlukan model yang lebih kecil, ringan dan cepat. Sama dengan istilahnya *mobile*, para peneliti Google membuat sebuah arsitektur CNN yang mungkin digunakan untuk kebutuhan *Mobile*. Mobilenet merupakan arsitektur yang telah ditingkatkan untuk peningkatan sebuah aplikasi terhadap perangkat mobile atau perangkat lain dimana mempunyai keterbatasan sumber daya perangkat keras dengan mengoptimalkan ukurannya [21].

2.4 Tensorflow

Tensorflow adalah *opensource library* untuk *machine learning* yang ditingkatkan oleh tim google yang telah dirilis pertama kali pada tau 2015. Memiliki beberapa komponen utama dan kelebihan yang dapat mendukung dalam sebuah penelitian[22].

2.5 Breast Cancer Wisconsin Dataset

Breast cancer Wisconsin dataset merupakan kumpulan data yang dapat diakses secara umum untuk digunakan sebuah penelitian[23]. Dataset ini dikembang oleh Universitas

Wisconsin. Dataset dapat diakses melalui laman UCI *repository* yang merupakan sebuah *database* publik yang memiliki beberapa deretan dataset yang bisa diakses publik.

Memiliki beberapa kategori dataset seperti *data mining*, statistika, bisnis, kesehatan, teknologi, lingkungan dan masih ada beberapa lagi.

2.6 Pengukuran performa

2.6.1 Confusion Matrix

confusion matrix untuk mengklasifikasikan apakah mendapat hasil kategori yang baik/buruk[24]. Memiliki empat istilah penafsiran hasil proses pengkategorian confusion matrix yaitu *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)* dan *False Negative (FN)* Terdapat beberapa istilah pada *confusion matrix* seperti yang terlihat pada **table 2.2** dibawah ini

Tabel 2.2 confusion Matrix

<i>Actual class</i>	<i>Predicted class</i>		Total
	yes	No	
Yes	TP	FN	P
No	FP	TN	N
Total	P'	N'	P+N

2.6.2 Accuracy

Matrix yang paling umum dalam melakukan evaluasi klasifikasi. Cara kerjanya dengan menghitung nilai probabilitas berdasarkan nilai yang benar.

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

Gambar 2.4 Accuracy

2.6.3 Precision

Nilai kebenaran dari hasil yang telah diklasifikasikan

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

Gambar 2.5 Precision

2.6.4 Recall

Nilai keakuratan dari nilai positif yang ada

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

Gambar 2.6 Recall

2.6.5 F-1 Score

F1-Score berhubungan dengan perbandingan nilai rata-rata dari precision dan recall. Perhitungan *F1-Score* dapat dihitung menggunakan persamaan :

$$F1 - Score = 2 \times \frac{recall \times precision}{recall + precision}$$

Gambar 2.7 F-1 Score

