

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Pustaka

Penelitian yang menerapkan *Convolutional Neural Network* (CNN) untuk melakukan deteksi dan klasifikasi pada objek gambar beberapa waktu belakangan ini banyak dilakukan dan dimanfaatkan di berbagai bidang teknologi untuk melakukan penyelesaian masalah yang ada. Berdasarkan penelitian terdahulu menunjukkan bahwa *Convolutional Neural Network* (CNN) merupakan suatu perkembangan teknologi yang dalam penerapannya menggunakan konsep meniru otak manusia untuk menyelesaikan permasalahan dalam mendeteksi sebuah gambar dan klasifikasi objek gambar.

Pada penelitian ini penulis telah melakukan studi literatur terhadap beberapa jurnal dan penelitian sebelumnya. jurnal dan penelitian yang dipilih berdasarkan topik yang sesuai dan berkaitan dengan tema penelitian yang diteliti yaitu Klasifikasi Gerakan Yoga Dengan Model *Convolutional Neural Network* Menggunakan Streamlit. Pada tabel 2.1 adalah penjelasan dari penelitian terdahulu.

Tabel 2.1 Penelitian Terdahulu

No.	Judul	Tujuan	Dataset	Metode	Hasil
1	Analisis Performa dan Pengembangan Sistem Deteksi Ras Anjing pada Gambar dengan Menggunakan <i>Pre-Trained</i> CNN Model [13].	Penelitian tersebut memiliki tujuan untuk membandingkan dan menganalisa hasil performansi citra ras anjing yang didapat dari tiga model yang ada pada <i>Convolutional Neural Network</i> (CNN) yaitu ada ResNet 50, Xception, dan VGG16 dan mengembangkan user interface yang berupa <i>website</i> dengan menggunakan <i>framework</i> flask agar memudahkan user serta menampilkan hasil prediksi dari ketiga model diatas.	Dataset yang digunakan berjumlah 79 citra anjing dari berbagai ras yang sudah diberi label.	<i>Convolutional Neural Networks</i>	Hasil percobaan yang digunakan dari tiga model yaitu model Xception memiliki tingkat akurasi <i>Dog/Not</i> sebesar 98.7 % dan <i>Dog Race</i> sebesar 67.1%, model ResNet50 memiliki tingkat akurasi <i>Dog/Not</i> sebesar 97.5% dan <i>Dog Race</i> sebesar 62.0%, dan model VGG16 memiliki tingkat akurasi <i>Dog/Not</i> sebesar 97.5% dan <i>Dog Race</i> sebesar 58.2%. Dengan hasil tersebut maka model Xception memiliki akurasi yang paling tinggi.

No.	Judul	Tujuan	Dataset	Metode	Hasil
2	Model CNN LeNet Dalam Pengenalan Jenis Golongan Kendaraan Pada Jalan Tol [14].	Penelitian tersebut memiliki tujuan untuk melakukan perancangan klasifikasi citra jenis kendaraan yang ada di jalan tol menggunakan arsitektur CNN LeNet.	Dataset yang digunakan merupakan hasil foto dari gerbang jalan tol. Masing – masing golongan diambil 100 foto.	LeNet	Hasil ujicoba dapat disimpulkan bahwa model LeNet dapat dibuktikan bahwa nilai dari akurasi yang didapat tergantung dari besarnya <i>epochs</i> yang digunakan antara 25 - 100 dan dengan <i>epochs</i> 100 nilai akurasi yang didapat bisa mencapai 95% dengan nilai loss 10%.
3	Implementasi <i>Convolutional Neural Network</i> (CNN) Untuk Klasifikasi Ekspresi Citra	Penelitian tersebut memiliki tujuan untuk mengetahui performa dari model CNN dalam mengidentifikasi ekspresi wajah seseorang yang sedang bahagia dan sedih dalam kondisi data yang tidak ideal.	Dataset yang digunakan adalah <i>Facial Expression Recognition</i> (FER) 2013. Terdapat 2 jenis label yaitu bahagia, sedih.	<i>Convolutional Neural Networks</i>	Dari hasil pengujian yang dilakukan menggunakan beberapa <i>optimizer</i> untuk melihat performa dari model CNN. Yang pertama menggunakan Adam <i>optimizer</i> memiliki nilai akurasi 64%

No.	Judul	Tujuan	Dataset	Metode	Hasil
	Wajah Pada FER-2013 Dataset [15].				yang cenderung lebih mengenali ekspresi bahagia daripada ekspresi sedih. Yang kedua Adamax <i>optimizer</i> memiliki nilai akurasi 66% yang memiliki porsi pengenalan ekspresi yang berimbang. Yang ketiga menggunakan N-Adam <i>optimizer</i> memiliki nilai akurasi 52% yang cenderung lebih mengenali ekspresi sedih daripada ekspresi bahagia. Yang keempat menggunakan SGD <i>optimizer</i> memiliki nilai akurasi 48% yang tidak

No.	Judul	Tujuan	Dataset	Metode	Hasil
					memiliki kecenderungan terhadap dua ekspresi tersebut.
4	Klasifikasi Batik Riau Dengan Menggunakan Convolutional Neural Networks (CNN) [8].	Penelitian tersebut memiliki tujuan untuk merancang model CNN digunakan untuk klasifikasi gambar yang mana merupakan batik riau dan yang bukan termasuk batik riau.	Dataset yang digunakan berjumlah 168 citra diantaranya 68 citra batik riau dan 100 citra bukan batik riau	<i>Convolutional Neural Networks</i>	Hasil dari penelitian ini menghasilkan akurasi sebesar 65% dikarenakan pengaruh dari banyak motif yang sama antara batik riau dan bukan termasuk batik riau.
5	Penilaian Tingkat Keparahan Penyakit COVID-19 Menggunakan Model CNN [16].	Penelitian tersebut memiliki tujuan untuk merancang model CNN untuk multi klasifikasi keparahan penyakit COVID-19 dengan penggunaan model CNN yang sepenuhnya otomatis dan mengetahui hasil performansi dari model arsitektur CNN.	Dataset yang digunakan merupakan X – Ray dari pasien COVID-19 yang berjumlah 3260. Ada beberapa klasifikasi tingkatan yaitu <i>mild</i>	<i>Convolutional Neural Networks</i>	Dari pengujian yang dilakukan diperoleh hasil akurasi dari model CNN yang telah dibuat sebesar 95.52% sedangkan penelitian lain yang menggunakan ResNet-101 mendapatkan akurasi 79.87%, ResNet-34 mendapatkan

No.	Judul	Tujuan	Dataset	Metode	Hasil
			sebanyak 1000 citra, <i>moderate</i> sebanyak 950 citra, <i>critical</i> sebanyak 600 citra, <i>severe</i> sebanyak 710 citra.		akurasi 75.76%, dan VGG16 mendapatkan akurasi 88.34%.
6	Klasifikasi Kapal Berdasarkan Convolutional Neural Networks [17].	Penelitian ini memiliki tujuan untuk merancang sistem yang menggunakan model CNN untuk mengklasifikasi gambar yang termasuk <i>coarse-grained ships</i> (grup 1) diantaranya ada <i>aircraft carriers</i> , <i>warships</i> , <i>civil ships</i> dan juga yang termasuk <i>fine-grained warships</i> (grup 2) diantaranya ada <i>coastal combat ships</i> , <i>shipyard transport ships</i> , <i>amphibious assault ships</i> ,	Dataset yang digunakan dibagi menjadi 2 class yaitu <i>coarse-grained ships</i> dan memiliki <i>subclass</i> yaitu <i>aircraft carriers</i> , <i>warships</i> , <i>civil ships</i> . Class satu lagi yaitu <i>fine-grained warships</i>	AlexNet, GoogLeNet	Hasil penelitian tersebut ada beberapa hasil yaitu model AlexNet yang dilatih dari awal tanpa menggunakan <i>transfer learning</i> sebesar 83,3% untuk <i>coarse-grained ships</i> dan 74.5% untuk <i>fine-grained warships</i> . Lalu model AlexNet yang telah dilakukan <i>transfer learning</i> sebesar 94,7% untuk <i>coarse-grained ships</i> dan 91%

No.	Judul	Tujuan	Dataset	Metode	Hasil
		<p><i>submarines, destroyer.</i> Dan melakukan analisa dari hasil masing – masing arsitektur CNN yang digunakan.</p>	<p>dan memiliki <i>subclass</i> yaitu <i>coastal combat ships, shipyard transport ships, amphibious assault ships, submarines, destroyer.</i></p>		<p>untuk <i>fine-grained warships</i>. Model GoogLeNet yang dilatih dari awal tanpa menggunakan <i>transfer learning</i> sebesar 86,7% untuk <i>coarse-grained ships</i> dan 77% untuk <i>fine-grained warships</i>. Model GoogLeNet yang telah dilakukan <i>transfer learning</i> sebesar 98% untuk <i>coarse-grained ships</i> dan 95% untuk <i>fine-grained warships</i>.</p>
7	<p>Pengenalan Gerakan Olahraga Berbasis (<i>Long Short-Term Memory</i>)</p>	<p>Penelitian ini memiliki tujuan untuk merancang pengenalan gerakan olahraga yoga dan dapat ditampilkan secara visual menggunakan camera.</p>	<p>Dataset yang digunakan adalah <i>keypoints</i> dari gerakan yoga yaitu</p>	<p><i>Long Short-Term Memory</i></p>	<p>Hasil penelitian tersebut bahwa akurasi <i>training</i> menggunakan LSTM sebesar 91%. Parameter validasi menghasilkan nilai akurasi sebesar 89%, dan nilai</p>

No.	Judul	Tujuan	Dataset	Metode	Hasil
	Menggunakan Mediapipe [3].		<i>warrior2, t-pose, tree pose</i>		<i>loss</i> sebesar 0.29 dan validasi <i>loss</i> sebesar 0.35 dengan penggunaan 400 <i>epochs</i> . Untuk klasifikasi gerakan <i>t-pose</i> mendapatkan hasil 100% dalam pengujian. Klasifikasi gerakan <i>warrior2</i> mendapatkan hasil sebesar 85%, sedangkan <i>tree pose</i> mendapatkan nilai akurasi sebesar 80%
8	Deteksi Gerakan Yoga Menggunakan Teknik <i>Machine Learning</i> [6].	Penelitian ini memiliki tujuan untuk merancang pengenalan gerakan yoga menggunakan bantuan <i>keypoints</i> .	Dataset yang digunakan adalah <i>keypoints</i> dari gerakan <i>vrikashasana, bhujangasana,</i>	<i>Logistic Regression, Support Vector Machine, Random Forest, K-Nearest</i>	Hasil penelitian tersebut menunjukkan bahwa <i>Logistic Regression</i> memperoleh hasil akurasi sebesar 94%, <i>SVM</i> memperoleh hasil akurasi sebesar 93%, <i>Random Forest</i>

No.	Judul	Tujuan	Dataset	Metode	Hasil
			<i>tadasana, padamasana, trikonasana, shavasana</i>	<i>Neighbour, Naïve Bayes</i>	memperoleh akurasi 89%, <i>K-Nearest Neighbour</i> memperoleh akurasi sebesar 93%, <i>Naïve-Bayes</i> memperoleh akurasi sebesar 91%.
9	Implementasi Arsitektur InceptionResNet-v2 dan (Studi Kasus Klasifikasi Pose Yoga) Squared Hinge Loss [7]	Penelitian ini memiliki tujuan untuk mendapatkan hasil terbaik dari pose gerakan yoga berdasarkan tingkat akurasi, <i>recall, precision, f1-score</i> .	Dataset yang digunakan adalah citra dari gerakan yoga <i>bridge, child, downwarddog, mountain, plank, seatedforwardbend, trianglepose, warrior1, warrior2</i>	InceptionResnet-V2	Hasil penelitian tersebut menunjukkan akurasi tertinggi model sebesar 89.98% dengan nilai loss 0.9494 dengan arsitektur InceptionResnet-v2 dan Squared Hinge Loss dengan epoch 200, learning rate 0.0001

2.2 Dasar Teori

2.2.1 Yoga

Yoga berasal dari ajaran *sad darsana*, kata yoga sendiri asal muasal nya berasal dari sebuah kata yaitu “*yu*” yang artinya mengikat bersama, menyeimbangkan, dan mengembangkan kekuatan fisik dari manusia yang melakukan yoga.

Yoga merupakan sebuah kesatuan mistik dari jiwa manusia yang terhubung dengan roh universal. Yoga juga memiliki sebuah artian yang disebut metode yang melatih fisik dan mental dengan tujuan agar mencapai kesatuan tersebut [18].

Keunggulan yoga yaitu memiliki sifat statis dalam artian gerakan dalam satu posisi yang dipertahankan beberapa saat, efek dari yoga memiliki banyak pengaruh pada sistem fisiologi, yoga memiliki sifat anabolic dalam artian gerakan terhubung dengan pernapasan yang memiliki efek suplai oksigen ke tubuh selalu terpenuhi, yoga juga memiliki sifat subjektif yang dapat membuat cenderung menguasai diri dan menjernihkan pikiran [19].

Gerakan yoga memiliki banyak aneka ragam yang disesuaikan dengan cabang dari yoga dan kebutuhan yang diinginkan oleh para pelaku yoga. Berikut merupakan beberapa gerakan yoga:

1. Gerakan *Goddess*

Gerakan yang biasa disebut dengan half squat ini memiliki kegunaan yang sangat mempengaruhi bagi tubuh yaitu memperkuat bagian kaki, otot pada bokong, dan otot bagian perut [20].

2. Gerakan *Downdog*

Gerakan *downdog* atau *downward dog* ini memiliki kegunaan yang dapat membantu tulang lengan yang biasanya rentan terkena penyakit tulang yaitu osteoporosis [4].

3. Gerakan *Plank*

Gerakan *plank* memiliki kegunaan untuk memperkuat otot inti dari tubuh, membantu pembentukan otot perut, dan dapat memulihkan postur tubuh [21].

4. Gerakan *Tree*

Gerakan *tree* ini memiliki kegunaan untuk melatih keseimbangan dan gerakan ini sangat berguna untuk ibu hamil terutama pada janin yang ada di dalam perut mendapatkan asupan oksigen yang sangat baik [22].

5. Gerakan *Warrior 2*

Gerakan *warrior 2* atau *virabhadrasana II* memiliki kegunaan untuk memperkuat otot kaki, perut, lengan dan dapat memperkuat daya tahan serta meningkatkan energi yang ada dalam tubuh [4].

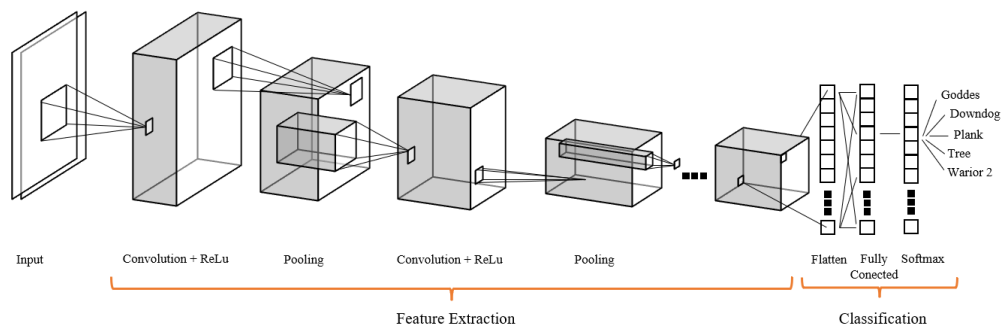
2.2.1 *Deep Learning*

Deep learning adalah metode *learning* yang memanfaatkan *artificial neural network* yang memiliki *multi layer*. *Artificial Neural Network* ini dibuat mirip seperti otak yang ada pada manusia, dimana *neuron – neuron* terkoneksi satu sama lain sehingga membentuk sebuah jaringan *neuron* yang sangat kompleks. *Deep learning* juga merupakan cabang ilmu dari *machine learning* berbasis jaringan saraf tiruan yang telah dikembangkan. Dalam *deep learning* komputer belajar mengklasifikasi secara langsung dari sebuah citra atau suara. *Convolutional Neural Network* adalah salah satu cabang dari *deep learning* yang merupakan perkembangan dari *Multi Layer Perceptron* (MLP) [23].

2.2.2 *Convolutional Neural Network*

Convolutional Neural Network (CNN) adalah bagian dari yang namanya *Deep Neural Network* yang bisa digunakan untuk klasifikasi gambar. Metode CNN adalah bentuk dari pengembangan *Multi Layer Perceptron* (MLP). Perbedaannya dengan MLP hanya menerima inputan satu dimensi sedangkan CNN dapat menerima data dua dimensi diantaranya yaitu citra. CNN merupakan jenis dari model *deep learning* yang bisa mengolah data dalam bentuk pola *grid*. CNN terdiri dari tiga jenis layer yaitu *convolutional layer*, *pooling layer*, *fully connected layer*. Masing – masing dari layer tersebut memiliki peranan yang berbeda yaitu *convolution layer* dan *pooling layer* memiliki kegunaan untuk ekstraksi fitur, sedangkan *fully connected layer* kegunaannya untuk memetakan fitur dari hasil ekstraksi yang telah dilakukan dengan output klasifikasi.

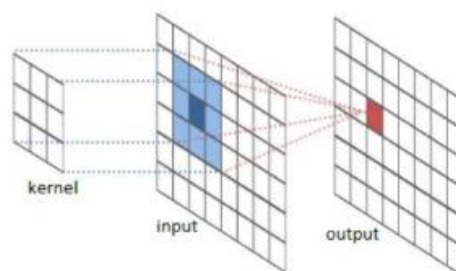
CNN memiliki kelebihan yang dimana tidak membutuhkan metode untuk ekstraksi dari ciri – ciri khusus, didalam proses metode ini memerlukan beberapa tahap yang nantinya dapat menghasilkan dari atribut yang dibutuhkan pada saat klasifikasi. Akan tetapi metode CNN membutuhkan jumlah data training yang banyak, sehingga perangkat yang digunakan pada unit pemrosesan dibutuhkan biaya yang cukup mahal untuk melakukan pelatihan sebuah model [24]. Dibawah ini merupakan contoh dari arsitektur metode CNN.



Gambar 2.1 Arsitektur Convolutional Neural Network

2.2.2.1 Convolution Layer

Convolution layer adalah bagian penting dari CNN dikarenakan hampir ke seluruh proses komputasi CNN akan dilakukan di lapis *convolutional layer*. Operasinya sama seperti operasi dari konvolusi yang dilakukan pada saat pengolahan citra yang dimana memiliki sebuah kernel dan *sub* citra. Pada umumnya kernel yang dipakai berukuran 3×3 , untuk setiap *sub* citra memiliki ukuran yang sama dengan kernel yang dilakukan pada operasi konvolusi [15].



Gambar 2.2 Convolutional Layer

2.2.2.2 ReLu (Rectified Linear Unit)

Rectified linear unit adalah bagian dari fungsi aktivasi. Kegunaan dari ReLu adalah menghilangkan sebuah nilai yang ada pada citra. Fungsi aktivasi ReLu cara

kerjanya yaitu dengan mengganti nilai negatif yang terdapat pada sebuah citra atau *feature maps* menjadi nilai 0 [25].

2.2.2.3 Pooling Layer

Pooling layer berfungsi untuk meringankan keluaran yang didapat dari convolutional layer yang mengakibatkan datanya kecil yang dapat dengan mudah untuk dikelola. Pada *pooling layer* masukan akan menjadi *grid* dan setiap *grid* memiliki *neuron*, lalu operasi *pooling* akan dipakai pada setiap sel yang ada di *grid*. *Pooling layer* memiliki tipe yang salah satunya adalah *max pooling*. *Max pooling* adalah operasi *pooling* yang cara memiliki cara kerja dimana *neural* yang mempunyai nilai aktivasi paling besar itulah yang akan digunakan [26].

2.2.2.4 Batch Normalization

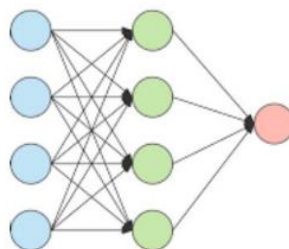
Batch normalization adalah teknik untuk melatih *deep neural network* membakukan input setiap lapisan mini-batch. *Batch normalization* memiliki efek menstabilkan proses *training* dan secara dramatis mengurangi jumlah periode pelatihan yang diperlukan untuk melatih jaringan dalam [27].

2.2.2.5 Flattening

Flattening adalah sebuah proses perubahan data *pooling* yang mempunyai array 2D lalu diubah ke data 1D *single vector*. *Flattening* bisa didefinisikan sebagai *reshape feature map* ke *vector* yang nantinya dapat digunakan untuk inputan dari *fully connected layer* [28].

2.2.2.6 Fully Connected Layer

Pada lapisan *fully connected layer* semua *neuron* akan disambungkan secara *fully connected* atau keseluruhan sama dengan *Multi Layer Perceptron* (MLP). Pada lapisan *fully connected layer* memiliki bobot dan bias yang dipakai oleh MLP untuk memodelkan data [15].



Gambar 2.3 Fully Connected Layer

2.2.2.7 Dropout

Dropout memiliki sebuah kinerja yang menghilangkan sementara sebuah *neuron* yang berupa *hidden layer* ataupun *visible layer* yang berbeda didalam sebuah jaringan. *Dropout* sendiri adalah salah satu bentuk usaha untuk mencegah terjadinya *overfitting* dan juga mempercepat proses pembelajaran model [11].

2.2.2.8 Softmax

Aktivasi *softmax* adalah bentuk lain dari algoritma *logistic regression* yang berguna untuk klasifikasi lebih dari dua kelas. Pada umumnya klasifikasi yang dilakukan dengan algoritma *logistic regression* tugasnya yaitu untuk *binary classification*. Persamaan bentuk yang ada pada softmax yaitu :

$$f_j(Z) = \frac{e^{z_j}}{\sum_k e^{z_k}} \quad (1)$$

Notasi dari f_j merupakan hasil fungsi untuk elemen j yang terdapat di vector keluaran kelas. Argumen z merupakan hipotesis supaya bisa mengklasifikasi yang dilakukan oleh fungsi *softmax*. *Softmax* bisa menghasilkan hasil yang lebih baik dibandingkan dengan algoritma klasifikasi lain dengan lebih intuitif dan interpretasi probabilitas. *Softmax* dapat mengkalkulasi probabilitas pada semua label. Dari label-label yang tersedia maka dilakukan pengambilan sebuah vektor dengan nilai riil dan dikonversikan menjadi vektor-vektor yang memiliki nilai dari 0 sampai satu yang jika dijumlahkan semuanya akan bernilai satu [23].

2.2.3 Confusion Matrix

Pengukuran terhadap kinerja sebuah sistem klasifikasi adalah hal yang penting. Kinerja sistem klasifikasi menggambarkan seberapa baik sistem dalam mengklasifikasikan sebuah data [29]. *Confusion matrix* merupakan table untuk menggambarkan kinerja model klasifikasi pada set data pengujian yang nilai nya telah diketahui [30]. Ini memungkinkan identifikasi yang mudah dari kebingungan antar suatu kelas, contohnya satu kelas umumnya salah diberi label sebagai yang lain. Jumlah prediksi yang benar dan salah dijadikan rangkuman dengan nilai – nilai hitung dan dipecah oleh setiap kelasnya. *Confusion matrix* juga memvisualisasikan akurasi klasifikasi dengan membandingkan kelas aktual dan prediksi [31].

Pada pengukuran menggunakan *confusion matrix* menghasilkan 4 hasil yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), *False Negative* (FN). Nilai *True Positive* adalah jumlah data positif yang terdeteksi benar. Nilai *True Negative* (TN) adalah jumlah data negatif yang terdeteksi benar. Nilai *False Positive* (FP) adalah jumlah data negatif namun terdeteksi sebagai data positif. Sedangkan *False Negative* (FN) adalah jumlah data positif namun terdeteksi sebagai data negatif [29].

<i>Confusion Matrix</i>		<i>Modeled Values: x_m</i>	
		<i>True</i>	<i>False</i>
<i>Actual Values: x</i>	<i>True</i>	TP	FN (<i>Type II error</i>)
	<i>False</i>	FP (<i>Type I error</i>)	TN

Gambar 2.4 *Confusion Matrix*

Pada gambar 2.4 merupakan penjelasan dari *confusion matrix* dan dari *confusion matrix* tersebut bisa digunakan untuk membuat metrics pengukuran nilai *testing* dan prediksi model yang hasilnya yaitu nilai Accuracy, Recall, F-1 Score, dan Precision.

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

$$Recall = \frac{TP}{(TP+FN)} \quad (3)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (4)$$

$$F - 1 \text{ Score} = 2 \frac{(Recall*Precision)}{(Recall+Precision)} \quad (5)$$

2.2.4 Streamlit

Streamlit adalah *framework* dari python yang sifatnya *open source*. Streamlit juga memiliki kemampuan yang dapat mengubah *script* menjadi sebuah aplikasi *website* yang bisa ditampilkan dalam hitungan menit. Hal tersebut membuat sangat mudah dalam pembuatan aplikasi berbasis web terutama bagi yang memiliki sedikit pengetahuan tentang desain *front – end web development* [32].