

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Sebelumnya

Pengembangan media pembelajaran Aksara Jawa interaktif berbasis multimedia merupakan salah satu upaya untuk melestarikan warisan budaya Jawa yang semakin terlupakan. Penelitian ini dilakukan oleh Nugroho dan Wulandari pada tahun 2020 dengan menggunakan metode penelitian dan pengembangan (R&D). Penelitian ini menghasilkan sebuah produk berupa aplikasi berbasis Smart Apps Creator (SAC) yang memiliki fitur pembelajaran dan latihan aksara Jawa. Kelebihan produk ini adalah praktis dan menarik, namun kekurangannya adalah belum dilakukan uji coba secara luas dan belum memuat materi aksara Jawa yang lengkap.[1]

Penelitian yang berjudul "*An improved faster-RCNN model for handwritten character recognition*" oleh Saleh Albahli et al., diterbitkan pada tahun 2021, bertujuan untuk meningkatkan pengenalan karakter tulisan tangan dengan menggunakan model *Faster R-CNN* yang diperbaiki. Metode ini mengintegrasikan *DenseNet-41* untuk ekstraksi fitur yang lebih mendalam, menghasilkan peningkatan akurasi dalam mendeteksi dan mengklasifikasikan angka tulisan tangan. Hasil penelitian menunjukkan keunggulan metode ini dalam mengatasi variasi gaya tulisan dan distorsi. Kontribusi utama termasuk pengembangan anotasi yang lebih efisien dan penerapan lapisan regresi dan klasifikasi yang baru. Meskipun demikian, penelitian ini kurang membahas kekurangan metode yang diusulkan, yang membuka peluang untuk eksplorasi lebih lanjut.[5]

Deteksi huruf aksara Jawa menggunakan *Faster R-CNN* adalah sebuah model yang dapat mendeteksi dan mengenali aksara Jawa dengan menggunakan algoritma *Faster R-CNN*. Model ini diusulkan oleh Faishal, Sulistiyo, dan Ihsan pada tahun 2023 dengan tujuan untuk membantu orang yang tidak familiar dengan aksara Jawa. *Faster R-CNN* dipilih karena memiliki akurasi yang tinggi dan kemampuan untuk

mendeteksi objek kecil. Penelitian ini menggunakan total 3094 data, dengan pembagian 2498 data training dan testing sebesar 596. Kelebihan model ini adalah dapat mendeteksi beberapa objek secara simultan, namun kekurangannya adalah membutuhkan data yang banyak dan variatif untuk proses pelatihan.[3]

Pengenalan Aksara Jawa dengan Menggunakan Algoritma *Convolutional Neural Network* adalah sebuah aplikasi yang dibuat dengan menggunakan bahasa pemrograman Python. Aplikasi ini dibuat oleh Pradhana, Wisesty, dan Sthevanie pada tahun 2020 dengan menggunakan algoritma CNN untuk mengklasifikasikan Gambar aksara Jawa ke dalam 20 kelas. Aplikasi ini memiliki tingkat akurasi yang tinggi dalam mengenali Gambar aksara Jawa yang telah disiapkan. Kelebihan aplikasi ini adalah mudah digunakan dan dapat dijalankan di berbagai platform, namun kekurangannya adalah belum dapat mengenali Gambar aksara Jawa yang berasal dari sumber lain.[6]

Berdasarkan uraian penelitian sebelumnya, dapat disimpulkan bahwa penelitian-penelitian tersebut juga memiliki kekurangan, seperti membutuhkan data yang banyak dan variatif untuk proses pelatihan, serta belum dapat mengenali gambar aksara jawa. Oleh karena itu, penelitian "Model Pengenalan Karakter Aksara Jawa Dengan *Faster R-CNN*" dapat memperbaiki kekurangan-kekurangan tersebut dengan menggunakan 500 dataset yang bervariasi untuk mengembangkan model yang lebih akurat dan efisien dalam mengenali aksara jawa.

2.2 Aksara Jawa

Aksara jawa atau lebih dikenal dengan Hanacaraka adalah turunan aksara brahmani(berasal dari Hindustan) yang digunakan untuk penulisan berbahasa Jawa, bahasa Makasar, bahasa Madura, bahasa Melayu, bahasa Sunda, bahasa Bali dan bahasa Sasak. Di negeri Hindustan tersebut terdapat bermacam-macam aksara, salah satunya yaitu aksara Pallawa yang berasal dari India bagian Selatan. Supaya lebih mengenal tentang Aksara Jawa perhatikan Gambar 2.1 di bawah ini.



Gambar 2. 1 Aksara Jawa

Dinamakan aksara Pallawa karena berasal dari salah satu kerajaan yang ada di sana yaitu Kerajaan Pallawa. Aksara Pallawa itu digunakan sekitar abad ke-4 Masehi. Di Nusantara terdapat bukti sejarah berupa prasasti Yupa di Kutai, Kalimantan Timur, ditulis dengan menggunakan aksara Pallawa. Aksara Pallawa ini menjadi cikal bakal dari semua aksara yang ada di Nusantara, seperti aksara Hanacaraka, aksara Rencong (aksara Kaganga), aksara Batak, Aksara Makasar dan Aksara Baybayin (aksara di Filipina).[7]

2.3 Machine Learning

Pembelajaran mesin (*Machine Learning*) adalah studi ilmiah tentang algoritma dan model statistik yang digunakan oleh sistem komputer untuk melaksanakan tugas tertentu tanpa diprogram secara eksplisit. Algoritma pembelajaran ini digunakan dalam banyak aplikasi yang kita gunakan sehari-hari. Setiap kali mesin pencari web seperti Google digunakan untuk mencari di internet, salah satu alasan mengapa mesin tersebut bekerja dengan baik adalah karena algoritma pembelajaran yang telah belajar bagaimana merangking halaman web.

Algoritma-algoritma ini digunakan untuk berbagai tujuan seperti penambangan data, pemrosesan Gambar, analitik prediktif, dan lain-lain. Keuntungan utama dari menggunakan pembelajaran mesin adalah bahwa, setelah

sebuah algoritma belajar apa yang harus dilakukan dengan data, ia dapat melakukan tugasnya secara otomatis.

Machine Learning digunakan untuk melatih mesin bagaimana mengelola data dengan lebih efisien. Tujuan dari *Machine Learning* adalah untuk belajar dari data. Banyak penelitian telah dilakukan tentang cara membuat mesin belajar sendiri tanpa harus diprogram secara eksplisit. Banyak matematikawan dan programmer menerapkan beberapa pendekatan untuk menemukan solusi dari masalah ini yang melibatkan dataset yang besar.[8]

2.4 Pengolahan Citra

Pengolahan citra digital merupakan salah satu disiplin ilmu yang mempelajari hal-hal yang berkaitan dengan perbaikan kualitas Gambar (peningkatan kontras, transformasi, warna, restorasi citra), transformasi Gambar (rotasi, translasi, skala, transformasi, geometrik), melakukan pemulihan citra ciri (feature images) yang optimal untuk tujuan analisis, melakukan proses penarikan informasi deskripsi objek atau pengenalan objek yang terkandung pada citra, melakukan kompresi atau reduksi data untuk tujuan penyimpanan data, input pengolahan citra adalah citra, sedangkan outputnya adalah citra hasil pengolahan.[9] Terdapat beberapa tahap dalam pengolahan citra:

1. Import Gambar melalui tools akuisisi Gambar
2. Analitic dan manipulation citra
3. Hasil dari pengolahan tersebut bisa berupa citra yang telah di edit atau data dari analisis citra.

Pada umumnya analog dan digital merupakan tahapan yang sering digunakan dalam pengolahan citra. Analog sering di praktikan dalam pengopian dalam bentuk print out. Pengolahan citra digital banyak digunakan dengan menggunakan komputer.[10]

2.4.1 Segmentasi Citra

Teknik untuk membagi atau memisahkan citra ke dalam beberapa daerah (region) berdasarkan kemiripan atribut yang dimilikinya disebut dengan segmentasi. Segmentasi juga disebut sebuah proses yang membagi sebuah citra menjadi sejumlah bagian atau objek. Segmentasi bukanlah proses tunggal yang dilakukan dalam pengolahan citra digital. Namun segmentasi merupakan proses yang penting dalam pengolahan citra digital.

Pada proses segmentasi objek yang menjadi target akan diambil untuk proses selanjutnya. Ada dua karakteristik nilai derajat kecerahan citra pada Teknik segmentasi ini, yaitu discontinuity dan similarity. Pada discontinuity, citra dipisah-pisah berdasarkan perubahan yang mencolok dari derajat kecerahannya, biasanya diterapkan pada deteksi tepi, garis, area dan sisi citra. Sedangkan pada *similarity* citra akan dipisah berdasarkan *Thresholding*, *region growing* dan *region spiltting* dan merging, biasanya diterapkan pada citra yang statis dan dinamis.[11] Proses segmentasi berarti memisahkan dengan latar belakang. Proses ini dapat dilakukan dengan mengekstraksi fitur-fitur yang ada pada citra plat nomor, salah satunya adalah kontur citra.[12]

Dengan cara atau metode segmentasi, sebuah citra yang semulanya sulit diolah dengan indera penglihatan secara langsung, kemudian setelah disegmentasi bisa mempermudah mengolah informasi yang terkandung didalamnya. Proses segmentasi masing-masing objek pada citra dapat diambil secara individu sehingga dapat digunakan sebagai input bagi proses lain. Ada dua macam *segmentation* yaitu:

1. *Full segmentation* adalah pemisahan suatu objek secara individu dari background dan diberi label pada tiap-tiap segmen.
2. *Partial segmentation* adalah pemisahan sejumlah data dari *background* dimana data yang disimpan hanya data yang dipisahkan saja untuk mempercepat proses selanjutnya.[9]

2.4.2 Transformasi Hough

Metode transformasi hough merupakan sebuah metode yang memisahkan antara objek dan background dengan penentuan batas-batas objek tersebut dalam sebuah citra. Tujuan dari transformasi ini adalah untuk mendapatkan suatu fitur yang lebih spesifik karena metode ini dapat mendeteksi sebuah tepian dengan celah pada batas fitur dan tidak dipengaruhi oleh derau. Misalkan mendeteksi garis, lingkaran, elips dan parabola. Adapun rumus yang digunakan pada transformasi hough bergantung pada jenis objek yang dicari. Untuk mencari objek garis menggunakan rumus.[11]

2.4.3 Thresholding

Salah satu teknik segmentasi citra digital adalah *Thresholding*. *Thresholding* mengubah citra skala keabuan menjadi citra biner. Metode ini melakukan pengaturan nilai ambang sedemikian rupa sehingga citra dapat tersegmentasi. Dengan kata lain dilakukan pemanipulasian nilai kemudian meninjau hasil *Thresholding* sampai segmentasi yang memuaskan tercapai. Ambang batas dapat digunakan untuk mengelompokkan citra jika objek terang tertanam dalam citra yang lebih gelap. Poin utama dalam metode ini adalah ketepatan dalam memilih nilai ambang batas.[11]

2.5 Deep Learning

Deep Learning merupakan metode learning yang memanfaatkan artificial *Neural Network* yang berlapis-lapis (multilayer). *Artificial Neural Network* ini dibuat mirip otak manusia, dimana neuron-neuron terkoneksi satu sama lain sehingga membentuk sebuah jaringan neuron yang sangat rumit. *Deep Learning* atau *deep structured learning* atau *hirarchial learning* atau deep neural merupakan metode learning yang memanfaatkan multiple non-linier transformation, *Deep Learning* dapat dipandang sebagai gabungan *Machine Learning* dengan AI (*artificial Neural Network*).[13]

Selain itu, metode *Machine Learning* konvensional hanya mengandalkan CPU dan RAM dalam proses komputasi, sehingga spesifikasi CPU dan RAM menentukan kecepatan komputasi. Sedangkan metode *Deep Learning*, selain menggunakan CPU dan RAM dalam proses komputasi, metode ini juga memanfaatkan kemampuan GPU sehingga proses komputasi data yang besar dapat berlangsung lebih cepat.[14]

2.5.1 NN (*Neural Network*)

Neural Network atau biasa disebut juga jaringan syaraf tiruan adalah suatu metode komputasi yang meniru sistem jaringan syaraf biologi. Metode ini menggunakan elemen perhitungan non-linear dasar yang disebut neuron yang diorganisasikan sebagai jaringan yang saling berhubungan, sehingga mirip dengan jaringan saraf manusia. Jaringan saraf tiruan ini dibentuk untuk memecahkan suatu masalah tertentu seperti pengenalan pola atau klasifikasi karena proses pembelajaran. *Neural Network* terdiri dari input lapisan neuron (atau node, unit), satu atau dua (atau bahkan tiga) lapisan tersembunyi neuron, dan lapisan akhir output neuron.[15]

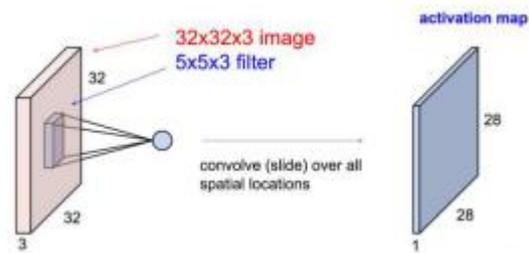
2.5.2 CNN (*Convolutional Neural Network*)

Convolutional Neural Network (CNN) adalah pendekatan pembelajaran mendalam yang banyak digunakan untuk memecahkan masalah yang kompleks. Pendekatan ini mengatasi keterbatasan pendekatan pembelajaran mesin tradisional. Motivasi dari penelitian ini adalah untuk memberikan pengetahuan dan pemahaman tentang berbagai aspek CNN.[16]

CNN digunakan untuk mengklasifikasi data yang terlabel dengan menggunakan metode *supervised learning*. Cara kerja dari *supervised learning* adalah terdapat data yang dilatih dan terdapat variabel yang ditargetkan sehingga tujuan dari metode ini adalah mengelompokkan suatu data ke data yang sudah ada.[14]

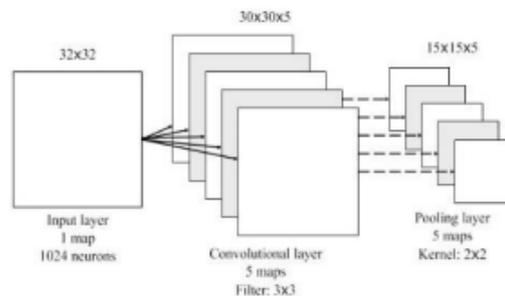
Arsitektur CNN terdiri atas satu lapisan (input layer), satu lapis keluaran (output layer), dan lapis tersembunyi (hidden layer). Lapis tersembunyi umumnya berisi *Convolutional Layer*, *Pooling Layer*, *normalization layer*, ReLu layer, *fully connected layer*, dan loss layer. Pada proses *training* terdapat tiga tahapan yaitu *Convolutional Layer*, *Pooling Layer*, dan *Fully connected layer*. Berikut penjelasan tiga tahapan tersebut.[17]

1. *Convolutional Layer*, terdiri dari neuron yang tersusun sedemikian rupa sehingga membentuk sebuah filter dengan tinggi dan panjang (pixels). Tiap filter memiliki nilai dan nilai dari filter inilah yang menjadi parameter yang akan di-update dalam proses learning. Proses tersebut bisa dilihat dalam Gambar 2.2 di bawah ini.



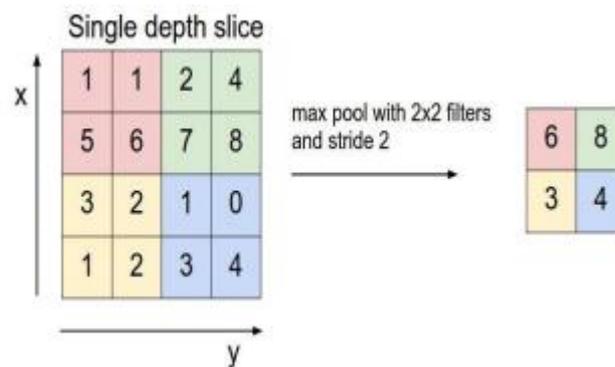
Gambar 2. 2 *Convolutional Layer*

2. *Pooling Layer*, *Pooling Layer* berfungsi menjaga ukuran data ketika convolution dilakukan, yaitu dengan melakukan reduksi sampel (down sampling) dengan pooling, kita dapat merepresentasikan data menjadi lebih kecil, mudah dikelola, dan mudah mengontrol overfitting. Perhatikan Gambar 2.3 berikut.



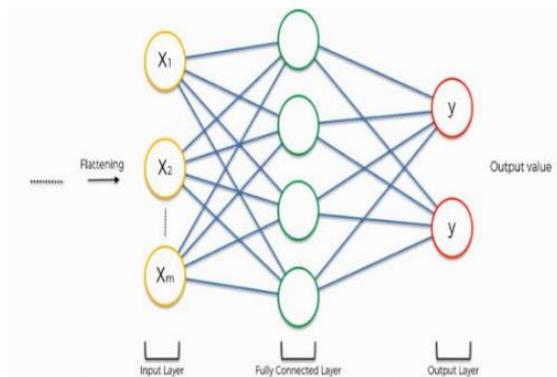
Gambar 2. 3 *Pooling Layer*

Pada umumnya proses pooling dilakukan menggunakan *Max Pooling* atau average pooling. Caranya adalah dengan memilih nilai maksimum atau nilai rata-rata dalam suatu area tertentu. Contohnya bisa dilihat Gambar 2.4 di bawah ini.



Gambar 2. 4 *Max Pooling*

3. *Fully connected layer*, pada lapisan yang terhubung secara penuh (*Fully connected layer*), setiap neurons memiliki koneksi penuh ke semua aktivasi dalam lapisan sebelumnya. Hal ini sama persis dengan MLP, yaitu komputasi menggunakan suatu perkalian matriks yang diikuti dengan bias offset. Contohnya dapat diperhatikan pada Gambar 2.5 berikut.



Gambar 2. 5 *Fully connected layer*

2.5.3 DNN (*Deep Neural Network*)

DNN menjadi metode yang paling penting dan kuat dalam pembelajaran mesin, yang secara luas diterapkan dalam visi komputer, pemrosesan bahasa alami, dan banyak bidang lainnya. Unit dasar DNN adalah perceptron, yang merupakan transformasi affine bersama dengan fungsi aktivasi. Non-linearitas dari fungsi aktivasi dan kedalamannya memberikan DNN kekuatan representasi yang besar.[18]

2.5.4 R-CNN

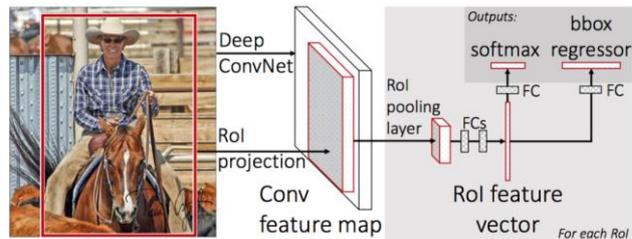
R-CNN bersekala sangat baik dengan jumlah kelas objek yang harus dideteksi karena hampir semua komputasi dibagi di antara semua kategori objek. Satu-satunya komputasi khusus kelas adalah produk matriks-vektor yang cukup kecil dan penekanan non-maksimum yang serakah. Meskipun komputasi ini berskala linier dengan jumlah kategori, faktor skalanya kecil. Diukur secara empiris, hanya perlu 30ms lebih lama untuk mendeteksi 200 kelas daripada 20 kelas pada CPU, tanpa perkiraan apa pun. Hal ini memungkinkan untuk mendeteksi puluhan ribu kategori objek dengan cepat tanpa modifikasi apa pun pada algoritma inti.[19]

2.5.5 *Fast R-CNN*

Fast R-CNN adalah model deteksi objek yang memperbaiki model sebelumnya yaitu R-CNN. Model ini menggunakan Gambar dan beberapa kotak yang menunjukkan kemungkinan terdapat objek. Model ini dapat melakukan hal berikut:

1. Mengolah Gambar dengan lapisan-lapisan conv dan *Max Pooling* untuk mendapatkan fitur-fitur penting dari Gambar.
2. Mengambil fitur-fitur dari setiap kotak dengan *RoI Pooling Layer* untuk mendapatkan vektor fitur yang sama panjangnya.

3. Mengklasifikasikan vektor fitur dengan lapisan fc dan softmax untuk menentukan jenis objek dan kelas “latar belakang” yang tidak ada objek.
4. Mengoptimalkan posisi kotak dengan lapisan fc lainnya untuk membuat kotak lebih pas dengan objek.



Gambar 2.6 *Fast R-CNN*

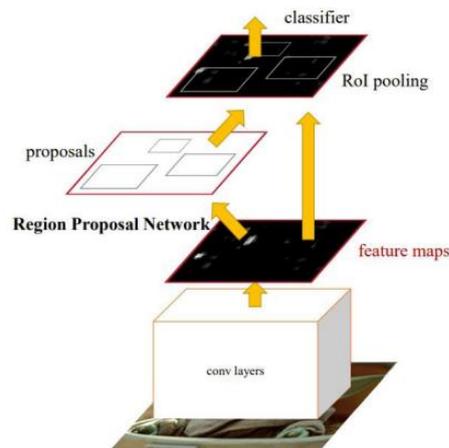
Pada Gambar 2.6 adalah arsitektur *Fast R-CNN*. Sebuah Gambar input dan beberapa daerah yang diminati (RoI) dimasukkan ke dalam jaringan konvolusi penuh. Setiap RoI digabungkan ke dalam peta fitur ukuran tetap dan kemudian dipetakan ke vektor fitur oleh lapisan yang terhubung penuh (FC). Jaringan ini memiliki dua vektor keluaran per RoI: probabilitas softmax dan offset regresi bounding-box per kelas. Arsitektur ini dilatih dari ujung ke ujung dengan kerugian multi-tugas.[20]

2.5.6 *Faster R-CNN*

Region based Convolutional Neural Network (Faster R-CNN). *Faster R-CNN* merupakan algoritma deteksi objek berbasis wilayah terbaru yang menunjukkan hasil luar biasa. Ada banyak penelitian yang sudah menggunakan *Faster R-CNN* pada bidang deteksi objek, seperti deteksi objek pada Gambar. Kinerja *Faster R-CNN* sangat bergantung pada kualitas data *training* yang besar.[4]

Faster R-CNN secara sederhana dapat dianggap sebagai sistem yang terdiri dari jaringan proposal regional dan *Regions Convolutional Neural Network (Faster R-CNN)*. Jaringan usulan regional digunakan sebagai pengganti algoritma Pencarian Selektif dari *Faster R-CNN*. Metode yang diusulkan berfokus pada

penyelesaian tiga masalah yaitu bagaimana merancang jaringan proposal regional, bagaimana membuat jaringan proposal wilayah, bagaimana membagi jaringan ekstraksi fitur.[21] Untuk pembahasan lebih jelasnya mengenai arsitektur *Faster R-CNN* dapat dilihat pada Gambar 2.7 di bawah ini.



Gambar 2. 7 Arsitektur *Faster R-CNN*

1. **Lapisan Konvolusi (conv layers):** Ini adalah lapisan dasar yang melakukan ekstraksi fitur dari Gambar.
2. **Peta Fitur (feature maps):** Menghasilkan peta fitur (feature maps) dari Gambar input.
3. **Region Proposal Network:** Mengolah peta fitur untuk mengusulkan potensi region objek.
4. **RoI Pooling:** Proposal region dari RPN diteruskan ke tahap RoI pooling. RoI pooling mengekstrak fitur tetap (fixed-size feature map) dari peta fitur untuk setiap proposal region.
5. **Klasifikasi (classifier):** Fitur yang diekstraksi oleh RoI pooling digunakan oleh classifier untuk melakukan klasifikasi dan regresi bounding box. Classifier mengklasifikasikan setiap proposal region ke dalam salah satu kelas objek atau sebagai latar belakang. Selain itu, classifier juga memprediksi penyesuaian koordinat untuk setiap bounding box agar lebih tepat mengelilingi objek.

Banyak metode dan arsitektur *Faster R-CNN* yang memperoleh kinerja yang tinggi, salah satunya adalah ResNet-50. Performa dari model ini dihitung dengan menggunakan *Confusion Matrix* yang di dalamnya menghitung *accuracy*, *precision*, *recall*, dan *f1-score* pada arsitektur ResNet-50 dimana berfokus pada nilai *accuracy* dan *F1-score*. [22]

2.6 *Confusion Matrix*

Confusion Matrix adalah pengukuran performa dari model klasifikasi *Machine Learning* dimana keluaran dapat berupa dua class atau lebih Terdapat empat istilah yang merupakan representasi kinerja proses klasifikasi pada *Confusion Matrix* yaitu true positive, true negative, false positive, dan false negative. Setelah ditemukan representasi kinerja proses langkah selanjutnya yaitu menghitung kinerja model seperti akurasi, presisi, *recall*, dan *F1-score* untuk kasus binary class. [22]

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \times 100\% \quad (2-1)$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (2-2)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (2-3)$$

$$F1\ Score = \frac{2 \times (Recall \times Precision)}{(Recall + Precision)} \times 100\% \quad (2-4)$$

Keterangan: **TP** (*True Positive*), **TN** (*True Negative*), **FP** (*False Positive*), **FN** (*False Negative*)