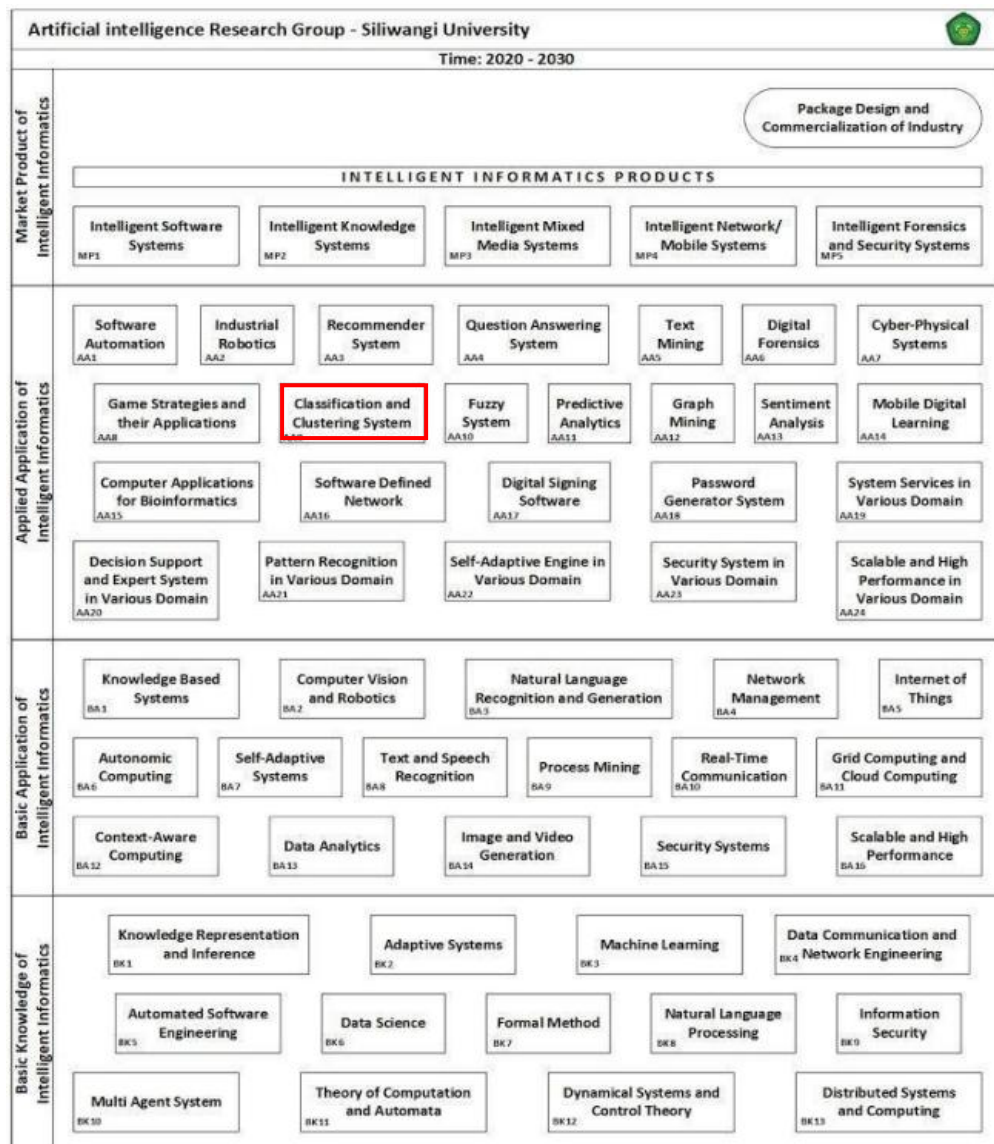


BAB III

METODOLOGI PENELITIAN

3.1 Roadmap Penelitian

Rencana penelitian ini sejalan dengan *roadmap* yang ada pada Universitas Siliwangi pada *classification and clustering system* pada ranah *applied application of intelligence informatics*. *Roadmap* penelitian dapat dilihat pada gambar 3.1.

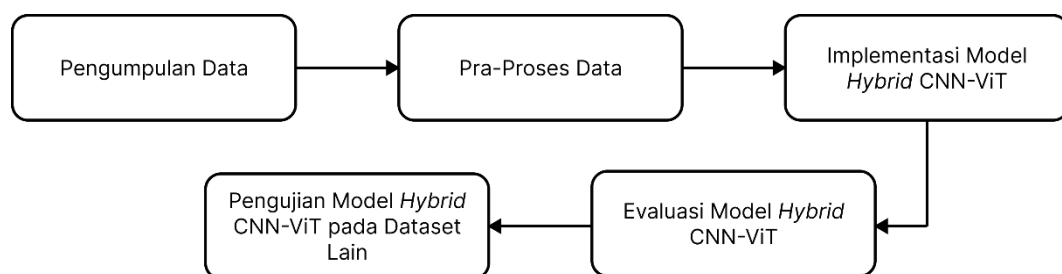


Gambar 3.1 Roadmap Penelitian (AIS Univesitas Siliwangi, 2020)

Pada gambar 3.1, pemilihan topik *classification and clustering system* pada ranah *applied application of intelligence informatics* sehingga tujuan dari penelitian ini yaitu mengimplementasikan pemodelan identifikasi. Penelitian ini melakukan pemodelan yang berorientasi pada penyakit pada daun pisang.

3.2 Metode Penelitian

Metode penelitian untuk identifikasi penyakit pada daun pisang dengan mengimplementasikan metode *hybrid CNN-ViT*. Metode penelitian yang dilakukan melibatkan beberapa langkah seperti pada gambar 3.2.



Gambar 3.2 Tahapan Penelitian

Pada gambar 3.2, merupakan ilustrasi dari beberapa tahapan utama yang dilakukan dalam proses penelitian. Proses pengimplementasian model untuk memprediksi penyakit pada daun pisang membutuhkan *dataset* berupa penyakit-penyakit pada daun pisang, sehingga pada tahap pengumpulan data ini menjadi langkah awal yang krusial dalam penelitian dan pemilihan desisi karena pada tahap ini memberikan wawasan yang berharga (Habsy dkk, 2024). Pada tahap pra-proses data dilakukan untuk meningkatkan kinerja dari model sehingga memastikan representasi dan identifikasi data mentah yang lebih efektif (Deng, 2024). Pada tahap implementasi model *hybrid CNN-ViT* yaitu diimplementasikannya arsitektur dari masing-masing model CNN dan ViT yang kemudian digabungkan menjadi

suatu lapisan *fully connected*. Pada tahap evaluasi model yaitu mengevaluasi performa model yang telah diimplementasi. Pada tahap pengujian model pada data lain yaitu dilakukannya pengujian model untuk mengidentifikasi penyakit pada daun pisang menggunakan data atau *dataset* yang belum pernah digunakan seperti *dataset* lain. Terdapat penjelasan detail dari setiap tahapannya yaitu sebagai berikut:

3.2.1 Pengumpulan Data

Pada proses pengumpulan data, dilakukannya pencarian informasi pada studi literatur untuk mendapatkan rumusan masalah terkait mengapa penelitian ini penting untuk dilakukan seperti yang sudah dilakukan pada BAB 1, setelah itu dilakukan pengambilan *dataset* yang akan digunakan dalam penelitian. *Dataset* diambil dari *platform* Kaggle.com yang bernama *Banana Leaf Spot Diseases Dataset* dan pada folder *OriginalSet* yang dimana folder *OriginalSet* ini memiliki satu kelas yang dominan daripada kelas lainnya, berbeda dengan folder *AugmentedSet* yang dimana jumlah gambar setiap kelas memiliki 400 gambar dan penambahan gambar diisi dengan gambar yang telah diaugmentasi oleh pemilik *dataset*, digunakannya folder *OriginalSet* pada penelitian ini agar lebih mudah untuk mengatur data agar sesuai pada model yang diimplementasi, *dataset* ini berfokus pada data penyakit pada daun pisang. *Dataset* ini terdiri dari empat kelas yaitu *cordana*, *sigatoka*, *pestalotiopsis*, dan *healthy*. Ukuran setiap citra pada *dataset* yaitu sebesar 224x224 piksel dengan format warna *Red, Green, Blue* (RGB) atau memiliki tiga saluran warna. Karena terdapat ketidakseimbangan kelas atau *imbalanced class*, maka dibutuhkannya sebuah konfigurasi agar model tetap dapat

mengidentifikasi gambar secara akurat walau memiliki *imbalanced class* yang dapat menimbulkan bias model.

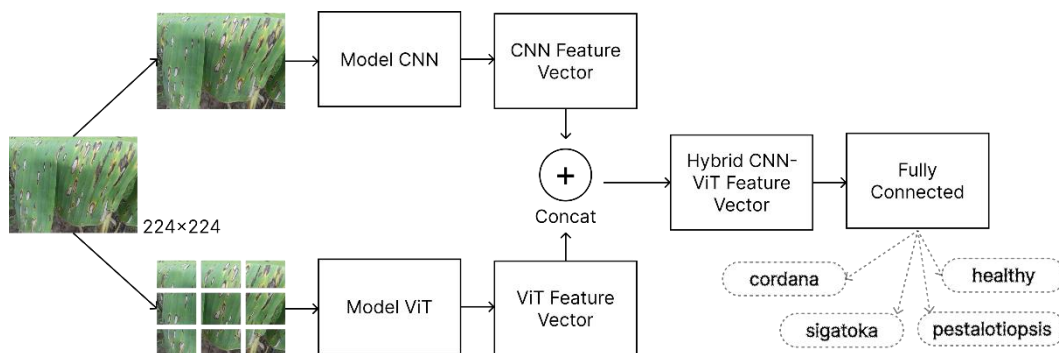
3.2.2 Pra-proses Data

Tahap pra-proses data merupakan tahap yang penting karena data merupakan bahan yang akan digunakan untuk pembelajaran model, karenanya data harus sesuai dengan keperluan model yang digunakan, setelah data diambil dari platform Kaggle maka dilakukannya beberapa pengolahan data untuk mengatasi beberapa masalah seperti *overfitting*, *underfitting*, dan *imbalanced class*, adapun pembagian data *training* dan data validasi dengan 80% data pelatihan dan 20% data validasi, pembagian data ini penting untuk mengestimasi performa generalisasi model, metode ini juga dapat menghindari *overfitting* dan memastikan model dapat bekerja baik pada data yang belum pernah dilatih sehingga performa model menjadi akurat (Xu dan Goodacre, 2018). Metode augmentasi data dilakukan untuk memperluas dan memperkaya *dataset* melalui variasi seperti *rotation*, *zoom*, *shear*, *horizontal flip*, dan *vertical flip* dengan *fill mode nearest*, sehingga dapat meningkatkan kinerja model menjadi lebih akurat dan presisi (Amrulloh dkk, 2024). Metode yang digunakan untuk mengatasi *imbalanced class* yaitu dengan menggunakan fungsi *class weight*,

3.2.3 Implementasi Model *Hybrid* CNN-ViT

Pada tahap ini dilakukannya proses penggabungan model CNN-ViT untuk mengidentifikasi penyakit pada daun pisang. Penggunaan model CNN digunakan karena menurut (Zahra dkk, 2023), CNN unggul dalam menangkap fitur lokal karena struktur CNN yang hierarkis dan memungkinkannya mendeteksi tepi,

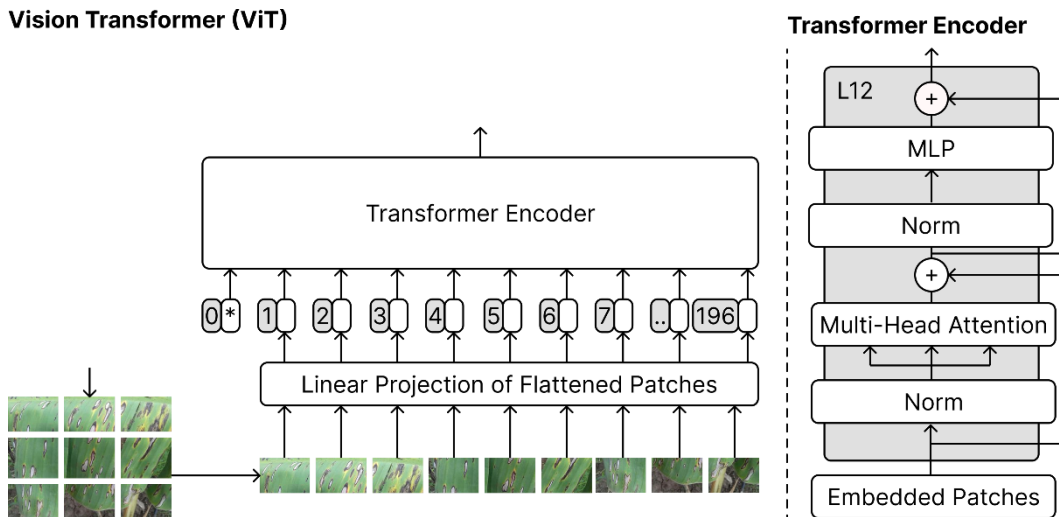
tekstur, dan pola secara efektif sehingga dapat berkontribusi secara signifikan dalam pemrosesan gambar. Sementara itu model ViT digunakan untuk mengekstraksi fitur global karena menurut (Wang dkk, 2025), model ViT unggul dalam menangkap fitur global karena arsitekturnya dapat memproses seluruh gambar sebagai urutan *patch*, sehingga persepsi global ini dapat meningkatkan kemampuan model dalam mengenali pola dan hubungan di seluruh gambar. Oleh karena itu, dengan menggunakan metode *hybrid* CNN-ViT dapat dilakukannya penggabungan kemampuan ekstraksi fitur lokal dari CNN dengan pemahaman fitur global dari ViT, sehingga dapat memaksimalkan kemampuan ekstraksi fitur jaringan pada model (Zhao dkk, 2024).



Gambar 3.3 Arsitektur Usulan *Hybrid* CNN-ViT

Terdapat usulan model *hybrid* CNN-ViT yaitu seperti pada gambar 3.3, yang dimana terdapat satu *input* gambar dengan ukuran 224x224 piksel yang kemudian gambar tersebut akan diproses kedalam model CNN dan model ViT secara paralel, adapun proses yang lebih detail untuk model ViT-B16 dapat dilihat pada gambar 3.4.

Vision Transformer (ViT)



Gambar 3.4 Proses pada Model ViT

Pada gambar 3.4, model menerima gambar dengan notasi $x \in \mathbb{R}^{224 \times 224 \times 3}$ karena gambar yang diterima merupakan gambar dengan piksel 224×224 dengan 3 saluran atau RGB yang akan dibagi menjadi beberapa *patch* berukuran 16×16 karena varian ViT yang digunakan adalah ViT-B16, sehingga total *patch* yang didapatkan yaitu $\frac{224 \times 224}{16^2} = 196$ *patch* dan dimensi yang didapatkan yaitu $16 \times 16 \times 3 = 768$ dimensi sehingga *flatten patch* dapat dinotasikan sebagai $x_p \in \mathbb{R}^{196 \times 768}$, setelah itu *patch* akan diproyeksikan ke dimensi D atau dimensi 768, setelah itu token klasifikasi dan *positional embedding* ditambahkan sehingga menjadi $x_p \in \mathbb{R}^{197 \times 768}$ dan didapatkan input awal seperti persamaan 2.1, setelah itu akan dimasukkan kedalam *transformer encoder* sebanyak 12 kali karena terdapat 12 lapisan pada ViT-B16 yang dimana dilakukan beberapa tahapan seperti *Multi-Head Self-Attention* (MSA) yang dimana menggunakan notasi seperti persamaan 2.2, setelah itu dilakukan tahapan *Multi-Layer Perceptron* (MLP) yang menggunakan notasi seperti persamaan 2.3, adapun *output* dari *transformer*

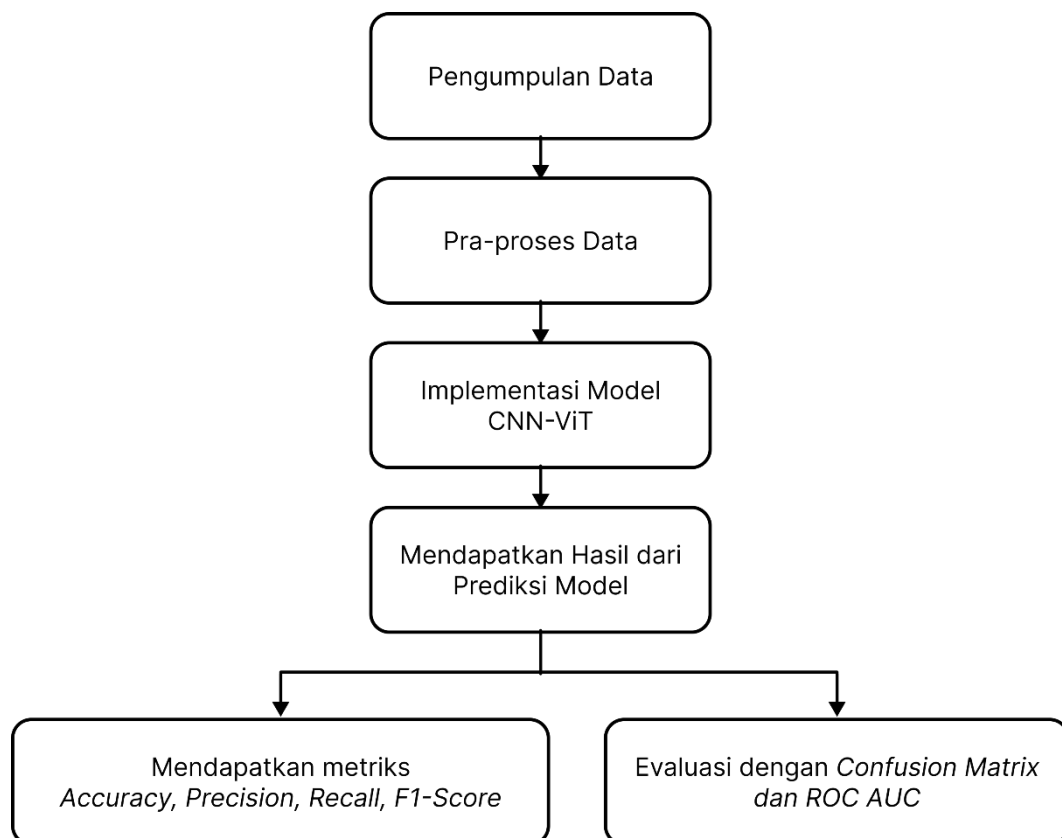
encoder yaitu berupa ViT *feature vector* dengan notasi $y = \text{LN}(z_L^0)$ yang dimana merupakan token [CLS] pada posisi ke-0 setelah melewati seluruh 12 lapisan pada *transformer encoder*.

Hasil dari pemrosesan model CNN dan ViT secara paralel tersebut akan menghasilkan CNN *feature vector* dan ViT *feature vector*, lalu masing-masing *feature vector* dari model CNN dan ViT tersebut akan digabungkan dengan menggunakan fungsi *concatenate* menjadi *hybrid CNN-ViT feature vector*, setelah itu *hybrid CNN-ViT feature vector* akan dimasukkan kedalam lapisan *fully connected* yang akan menjadi *classifier* yang dapat mengidentifikasi citra yang telah dimasukkan pada awal tahap termasuk kedalam kelas *cordana*, *healthy*, *pestalotiopsis*, ataupun *sigatoka*.

3.2.4 Evaluasi Model *Hybrid CNN-ViT*

Pada tahap ini dilakukannya evaluasi pada model seperti didapaknya performa matrix seperti *accuracy*, *precision*, *recall*, dan *F1-score*, lalu digunakan juga *confusion matrix* dan *Area Under the Receiver Operating Characteristic Curve* (ROC AUC). Evaluasi pada model dilakukan untuk memahami kualitas model yang telah diimplementasi, kualitas dapat dikonseptualisasikan sebagai representasi yang akurat dan komprehensif (Parker, 2024). Evaluasi pada model juga penting untuk menilai metrik kinerja seperti akurasi pelatihan dan validasi, evaluasi juga dapat mengidentifikasi jika model mengalami *overfitting* ataupun *underfitting*, serta dapat menentukan pengaturan *epoch* yang optimal, sehingga dapat diketahui kelebihan dan kekurangan dari model yang telah diimplementasi guna mempermudah dilakukannya penyempurnaan pada model (Jaafar dkk, 2024).

Pada gambar 3.5, terdapat tahapan metode evaluasi model yaitu pengumpulan data, pra-proses data, dan implementasi model *hybrid* CNN-ViT, setelah itu model akan memprediksi data validasi sehingga didapatkannya metrik performa model seperti *accuracy*, *precision*, *recall*, dan *F1-score*, dan terdapat juga metode evaluasi seperti *confusion matrix* dan ROC AUC, sehingga dengan dilakukannya tahapan evaluasi model, maka dapat ditentukan jika model sudah dapat mengidentifikasi penyakit pada daun pisang secara akurat tanpa terdapat masalah seperti *overfitting* ataupun *underfitting*.



Gambar 3.5 Tahapan Evaluasi Model

Performa dari model *hybrid* CNN-ViT yang telah diimplementasi akan dilakukan komparasi terhadap model-model terdahulu yang terdapat pada BAB 2

dan model usulan yang hanya menggunakan metode tunggal seperti hanya CNN atau ViT saja, komparasi dilakukan untuk menunjukkan bahwa model usulan *hybrid* CNN-ViT yang telah diimplementasi dapat lebih baik dari model-model sebelumnya dan sesuai pada tujuan penelitian ini yaitu mengukur pengaruh implementasi dari model *hybrid* CNN-ViT terhadap model-model lainnya atau model tunggal saja.

Pada metriks performa yang digunakan seperti *accuracy*, *precision*, *recall*, dan *F1-score* yaitu jika memiliki nilai yang mendekati 1 atau 100%, menunjukkan bahwa model memiliki performa yang bagus dibandingkan dengan hasil nilai yang kecil atau mendekati 0. Pada metode evaluasi *confusion matrix* yaitu menunjukkan model memiliki performa yang bagus jika terdapat banyak nilai pada diagonal kiri atas ke kanan bawah dibandingkan dengan nilai yang berada diluar diagonal kiri atas ke kanan bawah yang dimana menunjukkan bahwa model berhasil memprediksi sesuai dengan kelas aslinya. Pada metode evaluasi ROC AUC yaitu menunjukkan model memiliki performa yang bagus jika nilai ROC pada setiap kelas mendekati 1 atau 100%, dalam grafik ROC AUC posisi nilai setiap kelas berada pada pojok kiri atas menunjukkan performa model yang bagus dibandingkan dengan nilai kelas yang berada pada pojok kanan atas.

3.2.5 Pengujian Model *Hybrid* CNN-ViT pada Dataset Lain

Pada tahap pengujian model *hybrid* CNN-ViT pada data lain yaitu dilakukannya pengujian identifikasi penyakit pada daun pisang pada model *hybrid* CNN-ViT dengan varian ViT-B16 yang telah diimplementasikan terhadap data baru atau data lain sehingga dapat memastikan bahwa model yang telah diimplementasikan

mampu mengidentifikasi penyakit pada daun pisang dengan data lain atau data baru. Pengujian pada data lain juga dapat untuk menghindari bias yang dimana model hanya memiliki akurasi yang tinggi pada *dataset* yang digunakan untuk pembelajaran model saja tetapi pada data lain memiliki akurasi yang rendah, data uji yang digunakan untuk kelas *cordana*, *healthy*, dan *sigatoka* yaitu didapatkan dari *platform* Kaggle dengan nama *dataset* yaitu *banana leaf disease dataset v4* yang dibuat oleh Rayhan Arlistya, lalu data uji untuk kelas *pestalotiopsis* yaitu didapatkan dari *platform* Mendeley Data dengan nama *dataset* yaitu *Banana Leaf Disease Dataset Bangladesh* yang dibuat oleh Shahrin, adapun data lain yang digunakan pada tahap ini yaitu berjumlah 40 citra dengan masing-masing kelas berjumlah 10 citra sehingga terdapat 10 citra *cordana*, 10 citra *healthy*, 10 citra *pestalotiopsis*, dan 10 citra *sigatoka*.