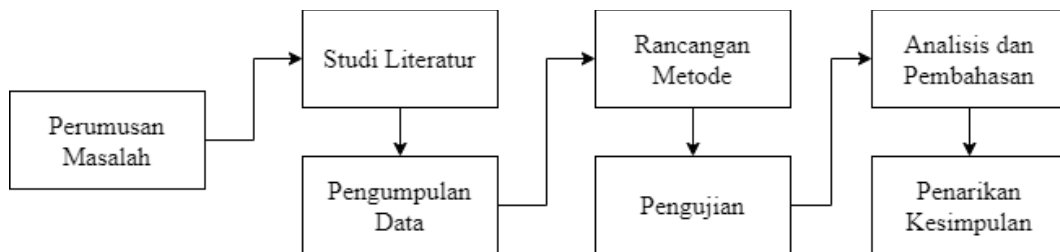


## BAB III METODE PENELITIAN

Pada bab ini akan dijelaskan mengenai detail dari penelitian yang akan dilakukan. Hal-hal yang akan dijelaskan pada bab ini mencakup tahapan penelitian dari studi literatur hingga penarikan kesimpulan dari penelitian yang dilakukan. Pada bab ini juga disebutkan lingkungan komputasi yang digunakan pada penelitian.

### 3.1 Tahapan Penelitian

Penelitian ini akan melalui beberapa tahap. Penelitian akan diawali dengan tahap studi literatur, lalu dilanjutkan dengan tahap pengumpulan data, rancangan metode, pengujian, analisis dan pembahasan, dan terakhir penarikan kesimpulan. Pada gambar 3.1 akan diperlihatkan rancangan tahapan yang akan dilewati pada penelitian ini.



Gambar 3. 1 Tahapan Penelitian

#### 3.1.1 Perumusan Masalah

Langkah pertama yang akan dilakukan dalam penelitian ini yaitu akan melakukan perumusan masalah dari penelitian yang akan dilakukan. Di mana dari perumusan masalah tersebut akan dicari jawabannya secara ilmiah dengan melakukan penelitian yang sudah dirancang. Pada tahap ini, penulis akan mencari beberapa permasalahan yang sesuai dengan penelitian yang dilakukan yaitu mengenai pencarian wajah pada video rekaman CCTV.

### 3.1.2 Studi Literatur

Studi literatur dilakukan dengan mencari referensi terkait dengan penelitian yang akan dilakukan. Referensi yang dicari dapat berasal dari buku, jurnal, maupun artikel yang terkait dengan permasalahan yang telah dirumuskan dan direncanakan untuk diteliti dan dicari solusinya. Studi literatur ini meliputi:

1. Pencarian Wajah
2. Rekaman Video CCTV
3. Python
4. Pengolahan Citra Digital
5. *Computer Vision*
6. Jaringan Saraf Tiruan
7. *Deep Learning*
8. *Convolutional Neural Network (CNN)*
9. *Exemplar Based*
10. *Triplet Loss Optimization*
11. *Normalized Cross-correlation Matching*
12. *Cross-correlation Matching*

### 3.1.3 Pengumpulan Data

Penelitian ini dilakukan dengan menggunakan dataset yang sudah tersedia, yaitu dengan menggunakan Chokeypoint dataset. Dataset tersebut bersifat *open source*. Dataset Chokeypoint dirancang untuk percobaan dalam penelitian *face recognition*, baik untuk identifikasi maupun verifikasi manusia pada video rekaman CCTV. Pada dataset tersebut, terdapat variasi dalam hal kondisi pencahayaan, pose, dan skala.

Terdapat 25 subjek yang terdiri dari 19 laki-laki dan 6 perempuan dalam dataset tersebut. Dataset memiliki *frame rate 30 fps* dan resolusi gambar 800x600 piksel, di mana setidaknya terdapat 64.204 gambar wajah dalam keseluruhan dataset.

Dataset tersebut dapat diaplikasikan untuk berbagai macam penelitian lainnya, seperti, *person re-identification*, *image set matching*, *face quality*

*measurement, face clustering, 3D face reconstruction, pedestrian* atau *face tracking, background estimation, dan subtraction*.

Chokepoint dataset dibagi menjadi 2 kelompok, (G1 dan G2). Di mana G1 digunakan untuk pengembangan, dan G2 digunakan untuk evaluasi. Untuk mempelajari verifikasi wajah manusia pada lingkungan yang berbeda dan interval waktu pada tiap perekaman, maka dapat menggunakan beberapa studi kasus yang telah disediakan.

### 1. Studi Kasus I

Pada studi kasus pertama yang disediakan pada dataset ini, terdapat video rekaman CCTV yang telah diubah ke dalam bentuk citra per-*frame* dengan *scene* dengan keadaan dalam ruangan. Selain itu, interval waktu yang terdapat pada studi kasus ini merupakan interval waktu yang lebih pendek. Untuk folder-folder rekaman video yang digunakan pada studi kasus pertama ini dapat dilihat di Tabel 3.1.

Tabel 3. 1 Studi Kasus I

G1	P1E_S1_C1	P1E_S2_C2
	P1L_S1_C1	P1L_S2_C2
G2	P1E_S3_C3	P1E_S4_C1
	P1L_S3_C3	P1L_S4_C1

### 2. Studi Kasus II

Pada studi kasus kedua yang disediakan pada dataset ini, terdapat video rekaman CCTV yang telah diubah ke dalam bentuk citra per-*frame* dengan *scene* dengan keadaan dalam ruangan serta luar ruangan. Seperti studi kasus pertama, interval waktu yang digunakan pada studi kasus kedua ini menggunakan interval waktu yang lebih pendek. Untuk folder-folder rekaman video yang digunakan pada studi kasus kedua ini dapat dilihat di tabel 3.2.

Tabel 3. 2 Studi Kasus II

G1	P2E_S2_C2	P2E_S1_C3
	P2L_S2_C2	P2L_S1_C1
G2	P2E_S4_C2	P2E_S3_C1
	P2L_S4_C2	P2L_S3_C3

### 3. Studi Kasus III

Pada studi kasus ketiga ini, data yang citra yang disediakan merupakan citra yang didapat dari video rekaman CCTV dengan *scene* dalam dan luar ruangan. Berbeda dengan studi kasus pertama dan kedua, interval waktu yang digunakan pada studi kasus ini, merupakan interval waktu yang lebih panjang. Selain itu untuk menggunakan studi kasus ini, setiap pasangannya harus dicocokkan terlebih dahulu dengan setiap urutan set satu dan set dua. Untuk folder-folder rekaman video yang digunakan pada studi kasus ketiga ini dapat dilihat di Tabel 3.3.

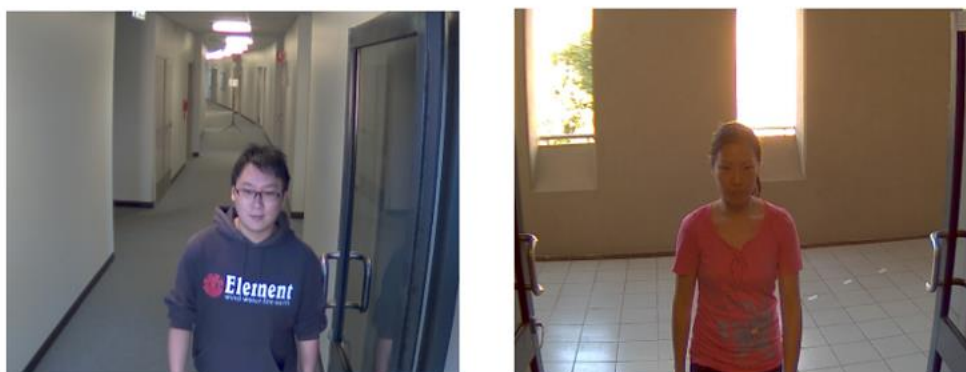
Tabel 3. 3 Studi Kasus III

	Set 1	Set 2
G1	P1E_S1_C1	P2E_S2_C2
	P1E_S2_C2	P2E_S1_C3
	P1L_S1_C1	P2L_S2_C2
	P1L_S2_C2	P2L_S1_C1
G2	P1E_S3_C3	P2E_S4_C2
	P1E_S4_C1	P2E_S3_C1
	P1L_S3_C3	P2L_S4_C2
	P1L_S4_C1	P2L_S3_C3

Setiap folder dinamai sesuai dengan kondisi perekaman, di mana P, S, C merupakan singkatan untuk portal, *sequence*, dan *camera*. E dan L menunjukkan

subjek memasuki (*entering*) atau meninggalkan ruangan (*leaving*). Misalnya, P2L\_S1\_C3 menunjukkan bahwa perekaman dilakukan di Portal 2, di mana rekaman berisi orang-orang yang meninggalkan ruangan, dan ditangkap oleh kamera 3 dalam *sequence* pertama.

Pada rekaman portal kedua, terdapat dua jenis *scene* dengan *background* yang berbeda. Folder rekaman video dengan kode P2E, merupakan rekaman video CCTV dengan jenis *scene*, luar ruangan. Sedangkan pada folder rekaman video dengan kode P2L, merupakan rekaman video CCTV dengan jenis *scene*, dalam ruangan. Contoh salah satu citra yang diambil dari rekaman video CCTV dapat dilihat pada Gambar 3.2.



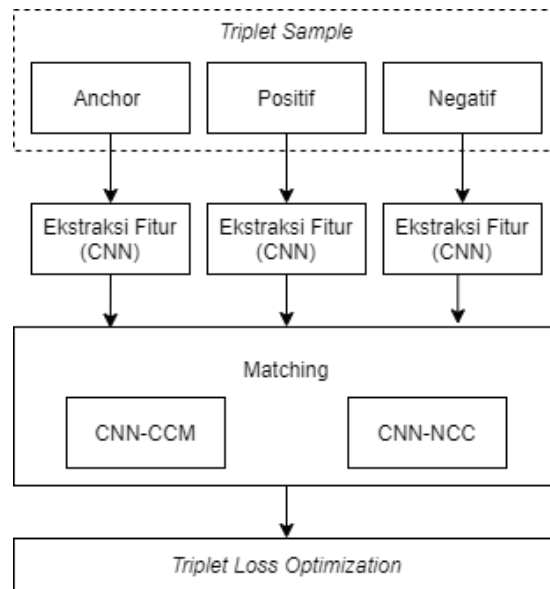
P2L – Indoor Scene

P2E – Outdoor Scene

Gambar 3. 2 Contoh Citra dari Tiap Tipe Scene

### 3.1.4 Rancangan Metode

Pada tahap ini akan dilakukan rancangan terhadap metode yang akan dibangun dalam pencarian wajah pada rekaman video CCTV sebagai masukannya seperti yang dapat dilihat pada Gambar 3.3. Dalam metode yang dirancang terdapat dua metode berbeda pada proses *matching*, yaitu dengan menggunakan metode *Cross-correlation Matching* (CCM) dan *Normalized Cross-correlation Matching* (NCC). Hal ini dilakukan untuk mengetahui perbandingan dari kinerja yang dilakukan oleh kedua metode tersebut.



Gambar 3. 3 Rancangan Metode

Tahap pertama dalam eksperimen adalah membangun arsitektur dari metode yang akan digunakan untuk pencarian wajah pada rekaman video CCTV. Tahap ini dilakukan dengan menganalisis metode-metode yang akan digunakan serta kebutuhan yang diperlukan sistem. Rancangan metode yang akan dilakukan digunakan untuk melakukan proses *training*. Pada tiap metodenya akan dibuat dua arsitektur berbeda, sehingga terdapat 4 arsitektur yang digunakan untuk melakukan proses *training*. Dalam proses *training* akan digunakan *triplet sample*, di mana *triplet loss optimization* akan digunakan sebagai *loss function* dari penelitian ini.

a. *Triplet Sample (Anchor, Positif, Negatif)*

*Training* dilakukan dengan menggunakan *triplet sample*. Di mana terdapat *anchor*, positif, dan negatif sampel yang dibutuhkan sebagai *input*. *Anchor* dan positif merupakan citra wajah dari subjek yang sama. Sedangkan negatif merupakan citra wajah dari subjek yang berbeda dari *anchor* dan positif. *Anchor* memiliki kualitas citra dengan resolusi yang lebih tinggi, sedangkan positif dan negatif merupakan citra dengan resolusi yang lebih rendah. Hal ini dikarenakan, sampel positif dan negatif didapati dari *frame* yang terdapat dalam video rekaman CCTV.

b. Ekstraksi Fitur

Proses ini bertujuan untuk mendapatkan fitur-fitur wajah yang akan dicari dari sampel yang telah dimasukkan. Untuk mendapatkan fitur-fitur wajah tersebut

digunakan metode *Convolutional Neural Network* (CNN). Masing-masing input (*anchor*, positif, negatif) akan diekstraksi fiturnya secara paralel dengan menggunakan bobot yang sama dari tiap *layer* yang akan digunakan.

c. *Matching*

Pada proses ini, setiap sampel akan dipasangkan. *Anchor* dengan positif, *anchor* dengan negatif, dan negatif dengan positif. Hasil ini bertujuan untuk menangkap kesamaan dari tiap pasangan sampel.

Proses ini akan dilakukan dengan 2 metode berbeda, yaitu CCM dan NCC. Pada metode CCM, akan digunakan perkalian matriks dengan menggunakan *Matrix Hadamard Product*. Yang selanjutnya hasil perkalian tersebut akan masuk kedalam *Fully Connected layer* dan *Softmax layer*. Pada persamaan yang digunakan pada untuk melakukan *matching* dapat dilihat pada persamaan (1) dan (2).

$$Sim(t_i, p_i) = (t_i \odot p_i) \quad (1)$$

$$wSim(t_i, p_i) = w_m \cdot RELU(w_n \cdot Sim(t_i, p_i) + b_n) + b_m \quad (2)$$

Metode CCM menggunakan formula yang berdasarkan pada penelitian yang dilakukan oleh Pachami dan kawan-kawan (Pachami dkk., 2017). Di mana  $w_m, w_n, b_n, b_m$  merupakan bobot dan bias yang didapati dari perkalian dengan menggunakan *Matrix Hadamard Product*.

Sedangkan pada metode NCC, proses perkalian akan diganti dengan menggunakan formula dari NCC. Formula yang digunakan berdasarkan pada penelitian yang dilakukan oleh Subramnian dan Chatterjee (Subramaniam & Chatterjee, 2016). Pada formula tersebut, masukan yang diterima akan dinormalisasikan sebelum diberikan filter. Persamaan yang digunakan untuk melakukan *matching* dengan menggunakan metode NCC dapat dilihat pada persamaan (3).

$$Sim(t_i \bar{*} p_i) = \frac{\sum_{i=1}^N (t_i - \mu_t)(p_i - \mu_p)}{(N-1) \cdot \sigma_t \cdot \sigma_p} \quad (3)$$

Di mana  $\bar{*}$  merupakan symbol yang mempresentasikan *normalized cross-correlation matching*. Sedangkan  $N$  merupakan total piksel  $t$  dan  $p$ .  $\mu_t$  dan  $\mu_p$  adalah rata-rata nilai piksel.  $\sigma_t$  dan  $\sigma_p$  merupakan standar deviasi dari nilai piksel  $t$  dan  $p$ .

$$wSim(t_i \bar{*} p_i) = w_m \cdot RELU(w_n \cdot Sim(t_i \bar{*} p_i) + b_n) + b_m \quad (4)$$

Pada persamaan (4),  $w_m, w_n, b_n, b_m$  merupakan bobot dan bias yang didapati dengan menggunakan NCC. Hasil dari NCC, selanjutnya akan masuk kedalam *Fully Connected layer* dan *Softmax layer*.

d. *Triplet Loss Optimization*

Persamaan yang digunakan dalam *Triplet loss function* pada penelitian ini dapat dilihat pada persamaan (5).

$$\text{Triplet Loss} = \frac{1}{L} \sum_{R_i \in B} \sqrt{(1 - S_{tipi})^2 + S_{tini}^2 + S_{nipi}^2} \quad (5)$$

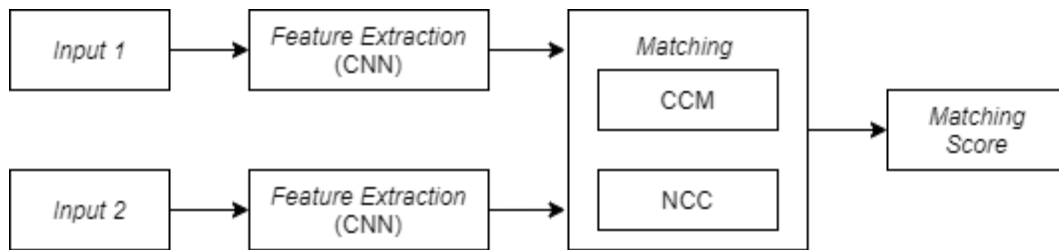
Di mana  $S_{tipi}$ ,  $S_{tini}$ , dan  $S_{nipi}$  merupakan nilai dari kesamaan yang didapatkan dari *cross correlation matching* dan *normalized cross correlation matching*. Formula yang digunakan untuk *loss optimization*, merupakan formula yang berdasarkan penelitian yang dilakukan oleh Parchami dan kawan-kawan (Parchami dkk., 2017).

### 3.1.5 Pengujian

Pengujian akan dilakukan dengan menggunakan *duplet sample*. Kedua sampel akan melewati proses ekstraksi fitur dan juga proses *matching* dengan model yang telah dilatih sebelumnya. Sehingga setiap sampel akan melewati *layer* yang sama dengan proses *training*.

Setelah melewati proses ekstraksi fitur dan juga proses *matching*, akan didapatkan nilai kesamaan. Nilai kesamaan ini akan digunakan untuk melihat apakah kedua citra wajah yang dimasukkan merupakan citra wajah dari subjek yang sama atau subjek yang berbeda seperti yang dapat dilihat pada Gambar 3.4.





Gambar 3. 4 Alur Untuk Mendapatkan Nilai Kesamaan.

Untuk mengetahui hasil dari kinerja pada masing-masing metode, aspek yang digunakan untuk mengevaluasi tiap metodenya dalam proses pengujian, yaitu dengan menggunakan akurasi (6).

$$\text{Akurasi} = \frac{(tp+tn)}{(tp+tn+fn+fp)} \quad (6)$$

Di mana  $tp$  (*true positive*) merupakan *item* relevan yang dihasilkan oleh sistem,  $tn$  (*true negative*) merupakan *item* relevan yang tidak dihasilkan sistem,  $fp$  (*false positive*) merupakan semua *item* yang dihasilkan oleh sistem, dan  $fn$  (*false negative*) merupakan semua *item* relevan yang tidak dihasilkan oleh sistem (Suka Parwita & Winarko, 2015). Hasil evaluasi disajikan dalam bentuk tabel untuk mempermudah proses analisis hasil yang dilakukan.

Dalam penelitian ini terdapat beberapa skenario pengujian yang akan dilakukan untuk mendapatkan hasil kinerja dari kedua metode yang telah dibuat. Pengujian ini bertujuan untuk mengetahui keakuratan dari sistem pengenalan wajah pada video rekaman CCTV.

Model dari kedua metode yang telah dibuat akan dilatih dengan menggunakan *triplet network*. Keakuratan akan didapat dengan menggunakan *sklearn metrics accuracy*, di mana label-label yang didapatkan dari hasil prediksi dengan menggunakan model *embedding* akan dibandingkan dengan label-label yang sesuai dengan setiap pasangan citra wajah yang sesungguhnya. Terdapat tiga skenario

pengujian yang dilakukan dengan menggunakan data yang berbeda, berikut skenario pengujian yang dilakukan:

1. Pengujian dengan Data *Testing*

Dalam pengujian ini akan menggunakan data *testing*. Data tersebut berjumlah 1000 pasang citra wajah. Sebagian pasang citra wajah merupakan pasangan citra dari wajah yang sama, sedangkan sebagian lainnya berasal dari citra wajah yang berbeda. Data pengujian ini merupakan citra wajah dari identitas yang berbeda dengan data *training* dan data *validasi*. Hal ini dikarenakan penelitian yang dilakukan merupakan penelitian dengan menggunakan *exemplar based*, di mana tidak bisa dilakukannya *training* pada data-data yang akan diuji diawal proses. Pada Tabel 3.4 diperlihatkan tiap skenario yang digunakan untuk pengujian dengan data *testing*.

Tabel 3. 4 Skenario Pengujian dengan Data *Testing*

No.	Metode	Arsitektur	<i>Threshold</i>
1.	CNN-CCM	Arsitektur A (Softmax)	0.9999998
			0.9999999
			1.0
			1.0000001
			1.0000002
		Arsitektur B (Sigmoid)	0.627
			0.628
			0.629
			0.630
			0.631
2.	CNN-NCC	Arsitektur A (2 <i>units</i> )	0.9999994
			0.9999998
			0.9999998
			1.0
			1.0000001

		Arsitektur B (1 unit)	0.2
			0.3
			0.4
			0.5
			0.6
			0.7

## 2. Pengujian dengan Studi Kasus

Selain dengan menggunakan data *testing*, pengujian juga akan dilakukan menggunakan salah satu studi kasus yang terdapat pada Chokepoint *face dataset*. Dari tiga studi kasus yang ada, pada penelitian ini akan dilakukan dengan menggunakan studi kasus kedua. Pada studi kasus tersebut terdapat dua tipe *scene* dari rekaman video CCTV. Tipe *scene* tersebut yaitu *scene* luar ruangan, serta *scene* dalam ruangan dengan interval waktu yang lebih pendek.

Langkah-langkah yang akan dilakukan pada pengujian dengan menggunakan studi kasus ini dapat dilihat pada tabel 3.5. Pada kolom nama folder dapat dilihat nama-nama folder dari rekaman video CCTV yang telah diubah ke dalam gambar dari setiap gambar tiap *frame*-nya.

Tabel 3. 5 Skenario Pengujian dengan Studi Kasus

No.	Nama Folder	Total <i>Frame</i>	Jumlah <i>Frame</i> yang Harus Ditemukan	Arsitektur
<b>Metode CNN-CCM</b>				
1.	P2E_S4_C2	803	100	A (Softmax)
2.	P2E_S3_C1	873	106	
3.	P2L_S3_C3	716	111	
4.	P2L_S4_C2	853	98	
5.	P2E_S4_C2	803	100	B

6.	P2E_S3_C1	873	106	(Sigmoid)
7.	P2L_S3_C3	716	111	
8.	P2L_S4_C2	853	98	
<b>Metode CNN-NCC</b>				
9.	P2E_S4_C2	803	100	A (2 units)
10.	P2E_S3_C1	873	106	
11.	P2L_S3_C3	716	111	
12.	P2L_S4_C2	853	98	
13.	P2E_S4_C2	803	100	B (1 unit)
14.	P2E_S3_C1	873	106	
15.	P2L_S3_C3	716	111	
16.	P2L_S4_C2	853	98	

Pengujian dengan studi kasus akan dilakukan dengan mengambil salah satu *frame* yang terdapat wajah target untuk dijadikan contoh dari citra wajah yang akan dicari. Gambar target tersebut akan dibandingkan dengan citra wajah dari seluruh *frame* yang ada untuk didapati nilai kesamaannya. Selain itu, pada pengujian ini, akan diambil model dengan akurasi yang tertinggi pada masing-masing metode dengan menggunakan data *testing*. Pada pengujian ini akan dilakukan dengan dua *threshold* yang memiliki akurasi tertinggi pada tiap metodenya.

### 3. Pengujian pada Kerumunan Orang

Pada pengujian dengan menggunakan studi kasus, video CCTV yang ditampilkan hanya terdapat satu wajah pada setiap *frame*-nya. Untuk melihat performa metode yang dibuat pada video rekaman CCTV yang menampilkan kerumunan orang maka akan dilakukan pengujian dengan menggunakan video rekaman CCTV yang menampilkan setidaknya lebih dari dua orang pada setiap *frame*. Pada pengujian ini, data juga didapatkan dari dataset yang sama, yaitu *Chokeypoint Face Dataset*.



Gambar 3. 5 Contoh *Frame* pada Video Rekaman CCTV dengan Keruman Orang

Pengujian pada kerumunan orang ini juga dilakukan dengan memastikan bahwa wajah yang dicari berbeda dengan wajah yang sudah di-*training* sebelumnya. Langkah-langkah yang akan dilakukan dalam pengujian ini dapat dilihat pada tabel 3.6.

Tabel 3. 6 Skenario Pengujian pada Kerumunan Orang

No.	Metode	Arsitektur	<i>Threshold</i>
1.	CNN-CCM	Arsitektur A (Softmax)	0.9999994
			0.9999998
			0.9999998
			1.0
			1.0000001
		Arsitektur B (Sigmoid)	0.632
			0.633
			0.634
			0.635
2.	CNN-NCC	Arsitektur A (2 <i>units</i> )	0.9999994
			0.9999998
			0.9999998
			1.0
			1.0000001

		Arsitektur B (1 unit)	0.1
			0.2
			0.3
			0.4

### 3.1.6 Analisis dan Pembahasan

Setelah tahap pengujian selesai, pada hasil yang didapati dilakukan analisis serta pembahasan. Pada tahap ini akan dijelaskan mengenai hasil yang didapati oleh masing-masing metode, yang selanjutnya akan dilakukan perbandingan terhadap hasil tersebut. Selain itu, pada tahap ini juga akan menjelaskan kelebihan serta kekurangan dari masing-masing metode yang telah diuji.

### 3.1.7 Penarikan Kesimpulan

Tahap ini merupakan tahap akhir dari penelitian. Penarikan kesimpulan akan diambil dari unjuk kerja yang telah dicapai pada penelitian. Sehingga apa yang telah dihasilkan oleh penelitian ini akan diketahui melalui kesimpulan.

## 3.2 Lingkungan Komputasi

Penelitian ini akan didukung dengan seperangkat komputer serta beberapa perangkat lunak yang akan digunakan. Berikut spesifikasi dari komputer yang akan digunakan dalam penelitian ini:

1. RAM 8 GB
2. Processor intel i7 8<sup>th</sup> Gen
3. Mouse dan Keyboard

Adapun perangkat lunak yang akan digunakan adalah:

1. Sistem Operasi Microsoft Windows 10
2. Visual Studio
3. Google Colab

Bahasa pemrograman yang akan digunakan pada penelitian ini adalah Bahasa pemrograman Python.