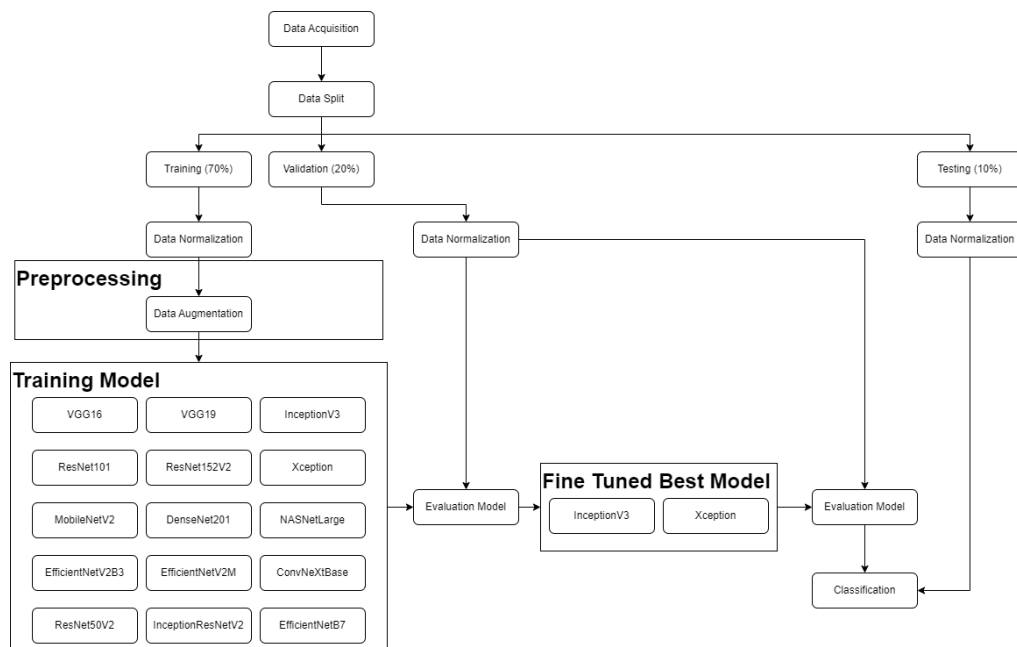


## Evaluasi

Model terbaik digunakan untuk klasifikasi akhir citra Pap smear, menghasilkan prediksi untuk kondisi HSIL, LSIL, Koilocyt, dan Normal berdasarkan hierarki kelas yang telah ditentukan.

Alur Penelitian dapat dilihat pada Gambar 1 berikut:



Gambar 1. Alur Penelitian

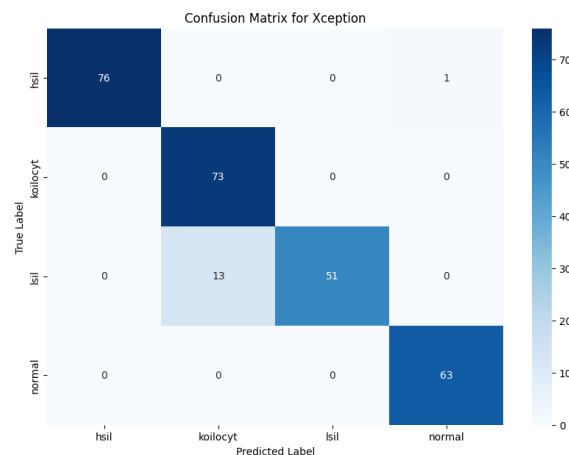
Gambar 1 ini menggambarkan alur kerja dari pengumpulan data, preprocessing, pelatihan dan evaluasi model, hingga fine-tuning dan klasifikasi akhir. Proses ini memastikan bahwa model yang dihasilkan adalah yang paling optimal dan mampu memberikan prediksi yang akurat berdasarkan data yang ada.

## Xception

Hasil evaluasi yang dilakukan dapat dilihat seperti berikut:

Classification Report for Xception:

	precision	recall	f1-score	support
hsil	1.00	0.99	0.99	77
koilocyt	0.85	1.00	0.92	73
lsil	1.00	0.80	0.89	64
normal	0.98	1.00	0.99	63
accuracy			0.95	277
macro avg	0.96	0.95	0.95	277
weighted avg	0.96	0.95	0.95	277



Pada Xception, dapat terlihat bahwa:

**Precision:** Kemampuan model untuk tidak mengklasifikasikan sampel positif sebagai negatif. Model Xception menunjukkan presisi yang sangat tinggi untuk kelas HSIL (1.00) dan LSIL (1.00), serta cukup baik untuk kelas Koilocyt (0.85).

**Recall:** Kemampuan model untuk menemukan semua sampel positif. Model ini memiliki recall tertinggi untuk kelas Normal (1.00) dan Koilocyt (1.00).

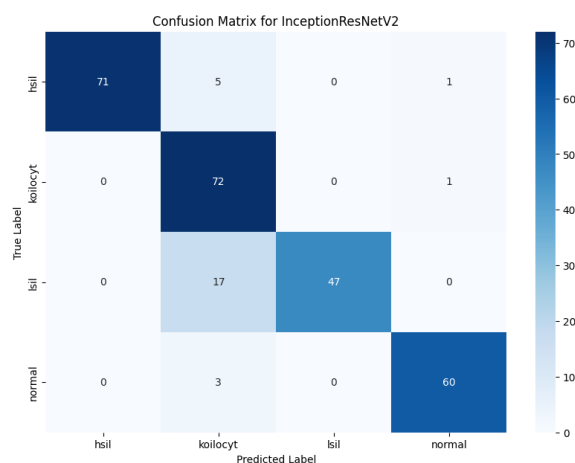
**F1-Score:** Rata-rata nilai dari precision dan recall. Model Xception memiliki F1-Score tertinggi untuk kelas HSIL (0.99) dan Normal (0.99).

**Akurasi:** 0.95 menunjukkan bahwa model ini mampu mengklasifikasikan dengan benar 95% dari semua sampel pengujian.

## InceptionResNetv2

Classification Report for InceptionResNetV2:

	precision	recall	f1-score	support
hsil	1.00	0.92	0.96	77
koilocyt	0.74	0.99	0.85	73
lsil	1.00	0.73	0.85	64
normal	0.97	0.95	0.96	63
accuracy			0.90	277
macro avg	0.93	0.90	0.90	277
weighted avg	0.92	0.90	0.90	277



Pada InceptionResNetv2 terlihat bahwa:

**Precision:** Model InceptionResNetV2 menunjukkan presisi yang sangat tinggi untuk kelas HSIL (1.00) dan LSIL (1.00), namun lebih rendah untuk kelas Koilocyt (0.74).

**Recall:** Recall tertinggi untuk kelas Koilocyt (0.99) dan Normal (0.95), menunjukkan kemampuan model untuk mendeteksi hampir semua sampel positif dalam kelas tersebut.

**F1-Score:** F1-Score tertinggi untuk kelas HSIL (0.96) dan Normal (0.96).

**Akurasi:** 0.90 menunjukkan bahwa model ini mampu mengklasifikasikan dengan benar 90% dari semua sampel pengujian.