Ahmad Noval Muhyiddin

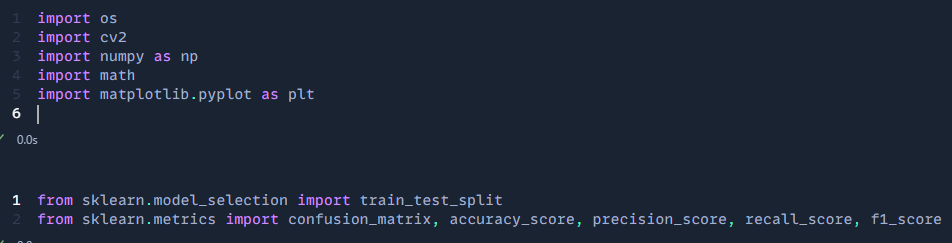
220411100154

Eksperiment

Tahapan

1. Library

Untuk library utama saya menggunakan cv2 untuk read image, library untuk proses data sperti numpy dan pandas, sklearn split train test untuk bagi data, dan library sklearn untuk evaluasi lainnya

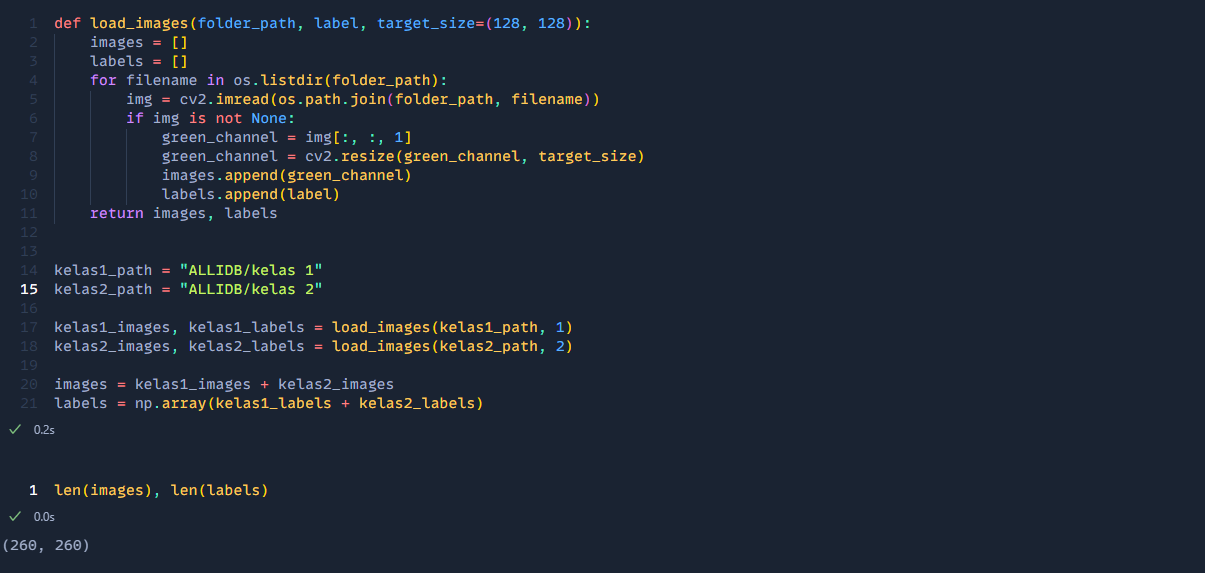


1. Understanding data

Pertama saya coba dulu untuk mengecek salah satu image dari dataset dengan opencv, dan saya coba ekstrak untuk channel greenya, hasilnya berikut



Setelah itu saya mulai menginput semua dataset ke dalam sebuah list image dan memberi kelas berdasarkan struktur folder dataset



Karena saya tidak tau kelas mana yang leukimia dan tidak, dan tidak ada keterangan juga di dataset, jadi saya memberikan label saja untuk kelas 1 dan kelas 2 sesuai data



Untuk image awal saya lihat 103 pixel, namun untuk input saya setarakan saja semua menjadi 128 pixel dengan channel green 8 bit, dan saya beri label sesuai input folder kelas

Maka toy datanya seharusnya seperti ini

|  |  |
| --- | --- |
| Images | Label |
| … opencv image | Kelas 1 |
| … image | Kelas 2 |

Bisa dilihat banyak data image juga sudah sesuai dataset, yaitu 260 image dari kelas 1 dan kelas 2

1. Prepocessing

* Segmentasi.

Untuk segmentasi saya memilih referensi dari google tentang jenis segmentasi dari google dan memilih metode treshold segmentation dengan otsu

Referensinya bisa dilihat disini

<https://blog.bigbox.co.id/image-segmentation-segmentasi-gambar-definisi-jenis-kelebihan-dan-penerapannya/>

<https://www.baeldung.com/cs/otsu-segmentation>

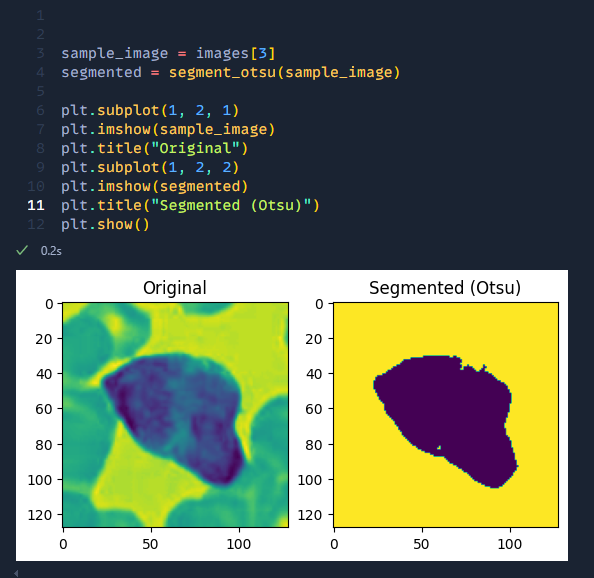
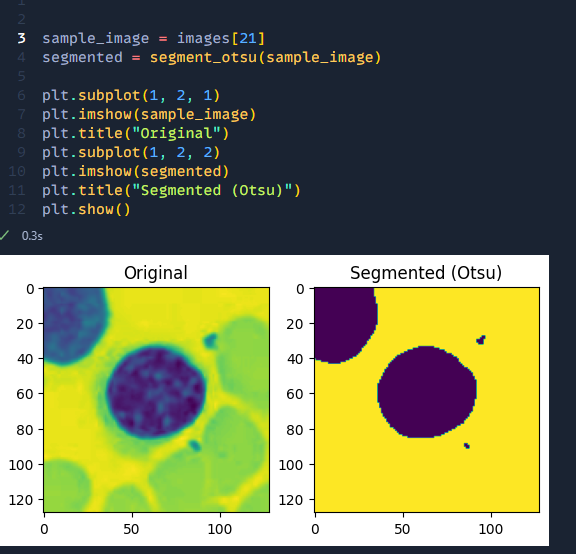
dan saya buat kode komputasinya untuk segmentasi citra gambar seperti berikut, kode segmentasi ini saya cari berdasarkan referensi dari link ini <https://www.baeldung.com/cs/otsu-segmentation>



Dari referensi untuk segmentasi otsu stepnya sebagai berikut

* Menghitung Histogram Skala Abu-abu
* Menghitung Fungsi Distribusi Kumulatif
* Menghitung Nilai Intensitas Skala Abu-abu Rata-rata Gambar
* Menghitung Varians Antar Kelas untuk Setiap Nilai Ambang yang Mungkin
* Menemukan Nilai Ambang Batas yang Memaksimalkan Varians Antar Kelas

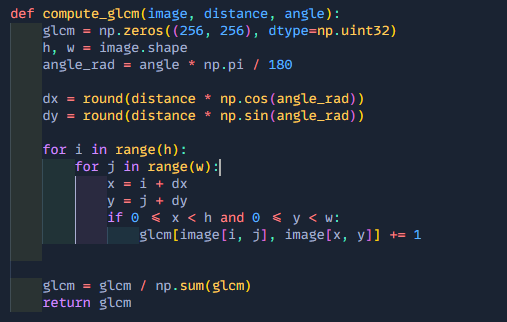
Setelah di test pada beberapa gambar di dataset, hasilnya seperti ini

* Ekstraksi Fitur GLCM

Untuk ekstraks saya mengekstrak 4 fitur dari image dengan distance 1 dan angle 45 derajat yaitu contrast, energy, homogeneity, entropy

1. Pertama untuk mengkalkulasi 4 fitur tadi saya ekstrak siapkan untuk array glcm sebagai matriks co-occurrence



Karena channel yang dipakai 8 bit green channel, saya siapkan array 256x256

Lalu saya menentukan offset pergeseran pixel dengan dx itu distancenya , dan dy itu angle nya

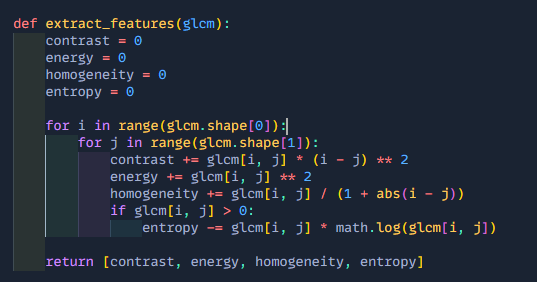
Setelah itu dengan perulangan untuk ukuran h dan w untuk mengisi setiap matrix dengan I dan J

Untuk setiap piksel (i, j), cari pasangannya di lokasi (x, y) yang digeser oleh (dx, dy).

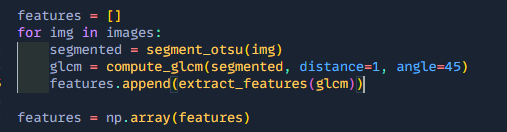
Jika pasangan berada dalam batas gambar (0 ≤ x < h dan 0 ≤ y < w), increment nilai GLCM pada indeks [intensitas (i, j), intensitas (x, y)].

Lalu saya normalisasi

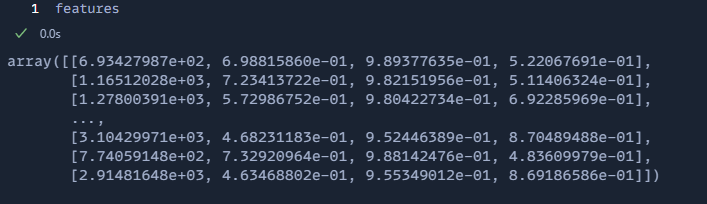
1. Lalu setelah itu sesuai rumus saya ekstrak 4 fitur dari matrix



1. Maka untuk setiap image dalam array saya lakukan segmentasi dan saya ekstraksi fiturnya

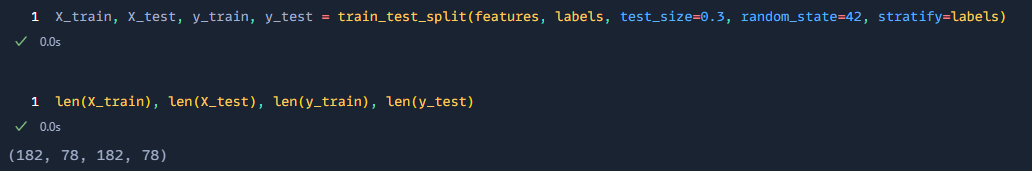


1. Maka didapat array fitur untuk contrast, energy, homogeneity, dan entropy untuk setiap gambar

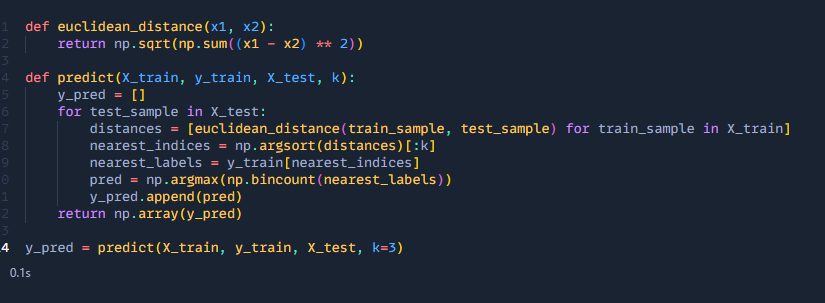


1. Modelling

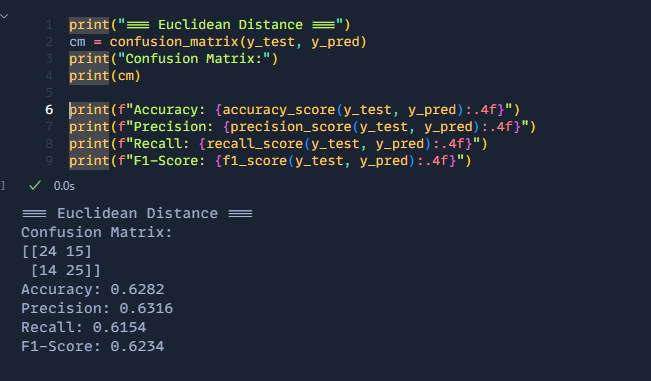
Train test spit saya bagi 70% train 30% test



Klasifikasi saya mencoba sederhanda dengan euclidean distance



Dengan hasil evaluasi sebagai berikut



Format default sklearn.metrics.confusion\_matrix

[[TN FP]

[FN TP]]