

SEGMENTASI CITRA DIGITAL PALEOGRAFI ARSIP VOC MENGUNAKAN METODE *THRESHOLDING*

Bandot Sumardiyono

Arsip Nasional Republik Indonesia
Jl. Ampera Raya No.7, Jakarta Selatan 12560
email : bandot.sumardiyono@anri.go.id

ABSTRAK

Segmentasi merupakan teknik untuk membagi suatu citra menjadi berapa daerah (region) dimana setiap daerah memiliki kemiripan atribut. Thresholding merupakan salah satu teknik segmentasi yang menghasilkan citra biner, yaitu citra yang memiliki dua tingkat keabuan, yaitu hitam dan putih. Segmentasi citra digital paleografi arsip VOC dilakukan untuk memisahkan antara karakter paleografi dengan backgroundnya. Tahapan yang dilakukan adalah dengan mengkonversi citra asli menjadi citra greyscale kemudian dengan menggunakan metode thresholding citra greyscale dikonversi menjadi citra biner. Tujuan pemisahan objek dan background pada citra digital paleografi arsip VOC adalah untuk memperjelas karakter-karakter pada citra digital paleografi sehingga arsip VOC lebih mudah dibaca. Dari hasil penelitian nilai ambang $T=120$ pada metode thresholding binary menghasilkan segmentasi citra digital paleografi yang terbaik.

Kata kunci : Thersholding, Segmentasi Citra, Paleografi.

ABSTRACT

Segmentation is a technique to divide an image into how many regions (regions) where each region has similar attributes. Thresholding is a segmentation technique that produces binary images, namely images that have two levels of gray, namely black and white. The segmentation of digital paleographic images of the VOC archives was carried out to separate the paleographic characters from the background. The steps taken are to convert the original image into a greyscale image then by using the thresholding method the greyscale image is converted into a binary image. The purpose of separating objects and backgrounds in digital paleographic images of the VOC archives is to clarify the characters in the paleographic digital images so that the VOC archives are easier to read. From the research results, the threshold value of $T=120$ in the binary thresholding method produces the best paleographic digital image segmentation.

Keywords: Thersholding, Image Segmentation, Paleography.

I. PENDAHULUAN

Arsip Nasional Republik Indonesia (ANRI) adalah Lembaga Pemerintah Non Kementrian (LPNK) yang memiliki tugas dan fungsi sebagai lembaga konservasi arsip dan sebagai lembaga pembina kearsipan tingkat nasional. Koleksi arsip yang ada di ANRI dibagi kedalam beberapa kategori yaitu arsip media konvensional yaitu berupa kertas dan arsip media baru berupa arsip foto, suara dan film. Koleksi arsip yang disimpan di ANRI mulai dari tahun 1602 hingga sekarang, sedangkan arsip yang tertua adalah koleksi arsip VOC (*Vereenigde Oost-Indische Compagnie*) berangka tahun 1602 yaitu sejak berdirinya VOC di Nusantara.

Arsip-arsip VOC yang ada di seluruh dunia sudah dijadikan sebagai warisan budaya dunia berupa “*Memory of The World*” oleh UNESCO pada tahun 2003. Arsip-arsip VOC ditulis dengan menggunakan huruf Paleografi, bahasa yang digunakan dalam arsip VOC adalah bahasa Belanda Kuno.

Arsip VOC yang sudah berumur ratusan tahun telah mengalami kerusakan yang disebabkan antara lain oleh korosi tinta, akibat serangga serta kelembaban tempat penyimpanannya sehingga hasil alihmedia citra digital paleografi arsip VOC memiliki noise.

Citra digital arsip VOC yang memiliki noise diproses menggunakan metode *thresholding* untuk memisahkan karakter paleografi dengan backgroundnya sehingga didapatkan hasil citra yang lebih baik.

Manfaatnya adalah pembaca lebih mudah membacanya dan arsip kertas paleografi VOC menjadi terjaga karena pengguna yang meminjam dapat menggunakan file elektronik sehingga tidak merusak arsip VOC yang aslinya.

II. TINJUAN PUSTAKA

Penelitian ini fokus pada segmentasi citra digital hasil alihmedia. Beberapa penelitian sebelumnya yang meneliti mengenai segmentasi citra adalah “*Segmentasi Citra Digital Bentuk Daun Pada Tanaman Di Politan Samarinda Menggunakan Metode Thresholding*”[1]. Untuk mengenal jenis-jenis bentuk tulang daun, telah dibuat sistem untuk membandingkan garis tulang daun beserta garis tepi daun menggunakan metode

thresholding. Prosesnya dimulai dengan menginput citra digital daun atau tanaman, selanjutnya dikonversi ke citra greyscale. Kemudian dilakukan proses segmentasi terhadap citra *greyscale*. Selanjutnya, dipilih hasil segmentasi dan ditandai dengan proses penajaman garis tepi menggunakan operator LOG. Proses terakhir adalah membuat histogram terhadap hasil proses penajaman garis tepi. Hasil segmentasi berhasil membandingkan dan menggolongkan bentuk tulang daun yang diambil menggunakan kamera ponsel pada tanaman di Politeknik Pertanian Negeri Samarinda dengan cara *thresholding* dengan hasil segmentasi citra daun yang telah digolongkan berdasarkan bentuk tulang daunnya dengan cara *thresholding* pula.

Dalam penelitian ini penulis menggunakan metode *thresholding*, karena berdasarkan pada penelitian sebelumnya metode *thresholding* terbukti mampu melakukan segmentasi dengan memisahkan objek dan background.

III. METODOLOGI PENELITIAN

Dalam penelitian ini menggunakan bahasa pemrograman C++ dan *library* OpenCV (*Open Source Computer Vision Library*) yaitu *library* fungsi pemrograman yang terutama ditujukan untuk *real-time computer vision*[2].

Langkah-langkah dalam pengolahan citra untuk mendapatkan segmentasi citra digital paleografi arsip VOC dalam penelitian ini adalah sebagai berikut:

3.1. Akuisisi Citra

Akuisisi citra adalah tahap awal untuk mendapatkan citra digital[3]. Tujuan akuisisi citra adalah untuk menentukan data yang diperlukan dan memilih metode perekaman citra digital. Tahap ini dimulai dari objek yang akan diambil gambarnya, persiapan alat-alat, sampai pada pencitraan

3.2. Preprocessing

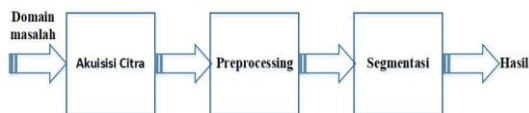
Tahapan ini diperlukan untuk menjamin kelancaran pada proses-proses berikutnya[3]. Hal-hal penting yang dilakukan pada tingkatan ini di antaranya adalah:

- Peningkatan kualitas citra
- Menghilangkan noise

- Perbaikan citratransformasi, dan
- Menentukan bagian citra yang akan diobservasi

3.3. Segmentasi

Tahapan segmentasi bertujuan untuk mempartisi citra menjadi bagian-bagian pokok yang mengandung informasi penting[3]. Misalnya objek dan latar belakang (*background*)



Gambar 3.1 Tahap-tahap segmentasi citra[3]

3.3 Citra Biner

Citra biner adalah citra digital yang hanya memiliki dua kemungkinan nilai pixel yaitu hitam dan putih[4]. Citra biner juga disebut sebagai citra B&W (*black and white*) atau citra monokrom. Hanya dibutuhkan 1bit untuk mewakili nilai setiap pixel dari citra biner. Citra biner seringkali muncul sebagai hasil dari proses pengolahan citra seperti segmentasi, *thresholding*, morfologi ataupun *dithering*.

3.4 Citra Greyscale

Cahaya yang tidak berwarna disebut cahaya monokromatik (atau akromatik). Satu-satunya atribut cahaya monokromatik adalah intensitasnya.

Karena intensitas cahaya monokromatik dianggap bervariasi dari hitam ke abu-abu dan akhirnya putih, istilah tingkat keabuan biasanya digunakan untuk menunjukkan intensitas monokromatik.

Rentang nilai cahaya monokromatik dari hitam ke putih biasanya disebut skala keabuan (*greyscale*), dan citra monokromatik sering disebut sebagai citra *greyscale*[5].

3.5 Metode Histogram

Histogram citra adalah grafik yang menggambarkan penyebaran nilai-nilai intensitas pixel dari suatu citra atau bagian tertentu di dalam citra[6].

Misalkan citra digital memiliki L derajat keabuan, yaitu dari nilai 0 sampai $L-1$ (misalnya pada citra dengan kuantisasi derajat

keabuan 8-bit, nilai derajat keabuan dari 0 sampai 255). Secara matematis histogram citra dihitung dengan rumus:

$$h_i = \frac{n_i}{n}, i = 0, 1, \dots, L-1 \quad (3.1)$$

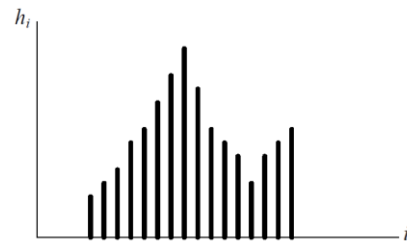
Dalam hal ini:

n_i = jumlah *pixel* yang memiliki derajat keabuan i .

n = jumlah seluruh *pixel* di dalam citra.

Plot h_i versus f_i dinamakan *histogram*. Gambar 3.2 adalah contoh sebuah *histogram* citra. Secara grafis *histogram* ditampilkan dengan diagram batang.

Perhatikan dari persamaan (3.1) bahwa nilai n_i telah dinormalkan dengan membaginya dengan n . Nilai h_i berada di dalam selang 0 sampai 1.



Gambar 3.2 Histogram citra[6]

Misalkan matriks di bawah ini menyatakan nilai digital yang berukuran 8×8 pixel dengan derajat keabuan dari 0 sampai 15 (ada 16 buah derajat keabuan)

3	7	7	8	10	12	14	10
2	0	0	0	1	8	15	15
14	6	5	9	8	10	9	12
12	12	11	8	8	10	11	1
0	2	3	4	5	13	10	14
4	5	0	0	1	0	2	2
15	13	11	10	9	9	8	7
2	1	0	10	11	14	13	12

Tabulasi perhitungan histogramnya ditunjukkan pada Tabel 3.1[6]. Mudah dilihat bahwa semakin besar nilai n_i maka semakin besar pula nilai h_i .

i	n_i	$h_i = n_i/n \ (n = 64)$
0	8	0.125
1	4	0.0625
2	5	0.078125

3	2	0.03125
4	2	0.03125
5	3	0.046875
6	1	0.015625
7	3	0.046875
8	6	0.09375
9	3	0.046875
10	7	0.109375
11	4	0.0625
12	5	0.078125
13	3	0.046875
14	4	0.0625
15	3	0.046875

Tabel 3.1 Tabel perhitungan *histogram*[6]

3.6 Metode Thresholding

Metode thresholding adalah metode segmentasi paling sederhana. Pada metode thresholding memisalkan wilayah gambar yang sesuai dengan objek yang ingin kita analisis. Pemisahan ini didasarkan pada variasi intensitas antara piksel objek dan piksel latar belakang.

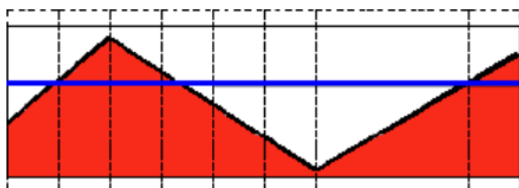
Untuk membedakan piksel yang objek dari yang lain (yang pada akhirnya akan ditolak), dilakukan perbandingan setiap nilai intensitas piksel sehubungan dengan ambang batas (ditentukan sesuai dengan masalah yang akan dipecahkan).

Setelah memisahkan piksel objek dengan benar, kemudian untuk mencari citra digital yang paling sesuai (paling baik/bagus) dapat diatur dengan nilai yang ditentukan untuk mengidentifikasinya (yaitu, nilai 0 (hitam) sampai dengan nilai 255 (putih)).

3.6.1 Tipe Thresholding

Dalam OpenCV ada fungsi `cv::threshold` untuk melakukan operasi *thresholding*. Ada 5 jenis operasi *thresholding* dengan fungsi ini.

Untuk mengilustrasikan bagaimana proses *thresholding* ini bekerja, mari kita pertimbangkan bahwa kita memiliki gambar sumber dengan piksel dengan nilai intensitas asal $src(x, y)$. Gambar 3.2 di bawah menggambarkan hal ini. Garis biru horizontal mewakili ambang batas (tetap).

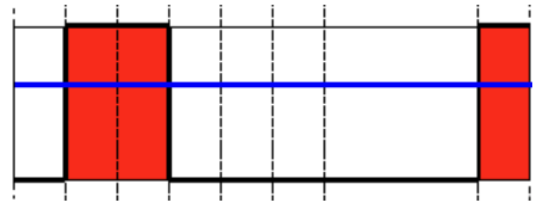
Gambar 3.2 Threshold dengan piksel nilai intensitas asal $src(x, y)$ [7].

3.6.1.1. Threshold Binary

Operasi thresholding dapat dijelaskan sebagai berikut:

$$dst(x, y) = \begin{cases} \text{maxVal} & \text{if } src(x, y) > \text{thresh} \\ 0 & \text{otherwise} \end{cases}$$

Jadi, jika intensitas piksel asal $src(x, y)$ lebih tinggi dari ambang, maka intensitas piksel baru diatur ke MaxVal . Jika tidak, piksel diatur ke 0.



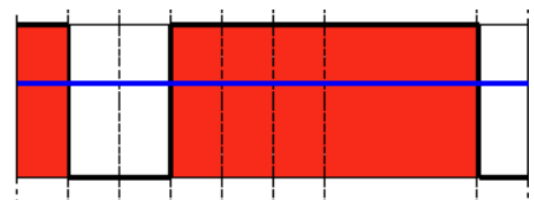
Gambar 3.3 Threshold Binary [7]

3.6.1.2. Threshold Binary, Inverted

Operasi *thresholding* dapat dijelaskan sebagai berikut:

$$dst(x, y) = \begin{cases} 0 & \text{if } src(x, y) > \text{thresh} \\ \text{maxVal} & \text{otherwise} \end{cases}$$

Jika intensitas piksel asal $src(x, y)$ lebih tinggi dari ambang, maka intensitas piksel baru diatur ke 0. Jika tidak, diatur ke MaxVal .



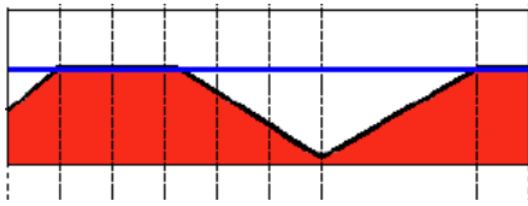
Gambar 3.4 Threshold Binary, Inverted[7]

3.6.1.3. Threshold Truncate

Operasi thresholding dapat dijelaskan sebagai berikut:

$$dst(x, y) = \begin{cases} \text{threshold} & \text{if } src(x, y) > \text{thresh} \\ src(x, y) & \text{otherwise} \end{cases}$$

Nilai intensitas maksimum untuk piksel adalah ambang, jika asal $src(x, y)$ lebih besar, maka nilainya terpotong (*truncated*). Lihat gambar di bawah ini:



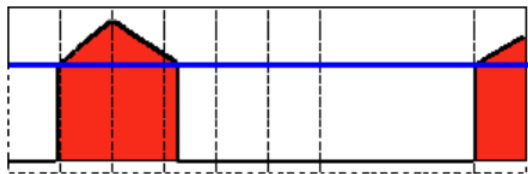
Gambar 3.5 Threshold Truncate[7]

3.6.1.4. Threshold to Zero

Operasi thresholding dapat dijelaskan sebagai berikut:

$$\text{dst}(x, y) = \begin{cases} \text{src}(x, y) & \text{if } \text{src}(x, y) > \text{thresh} \\ 0 & \text{otherwise} \end{cases}$$

Jika asal $\text{src}(x, y)$ lebih rendah dari ambang, nilai piksel baru akan disetel ke 0.



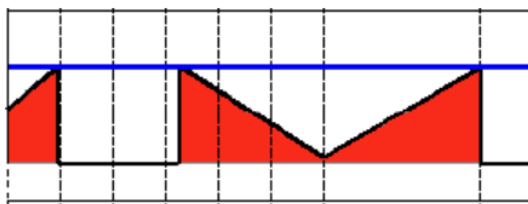
Gambar 3.6 Threshold to Zero[7]

3.6.1.5. Threshold to Zero, Inverted

Operasi *thresholding* dapat dijelaskan sebagai berikut:

$$\text{dst}(x, y) = \begin{cases} 0 & \text{if } \text{src}(x, y) > \text{thresh} \\ \text{src}(x, y) & \text{otherwise} \end{cases}$$

Jika asal $\text{src}(x, y)$ lebih besar dari ambang, nilai piksel baru akan disetel ke 0.

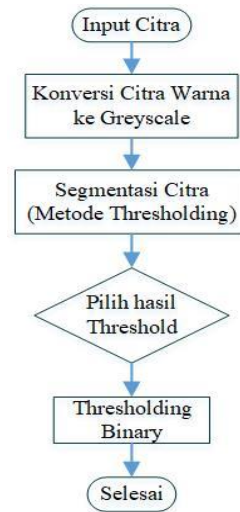


Gambar 3.7 Threshold to Zero, Inverted[7]

IV. HASIL DAN PEMBAHASAN

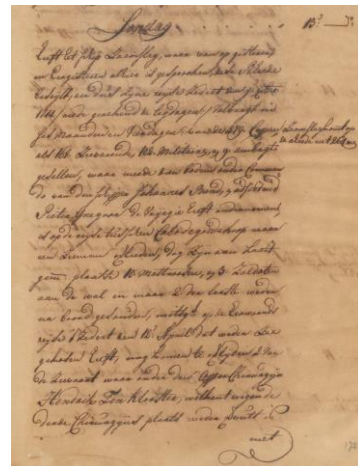
Penelitian ini mengikuti *flowchart* penelitian seperti ditunjukkan pada Gambar 4.1 terhadap citra digital paleografi arsip VOC dengan model warna RGB (Red, Green, Blue) dengan

ekstensi JPG (Joint Photographic Experts Group).

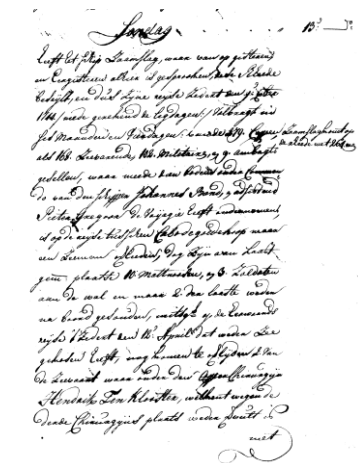


Gambar 4.1 Flowchart Penelitian

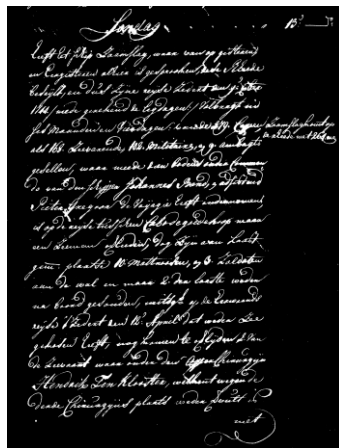
4.1. Hasil



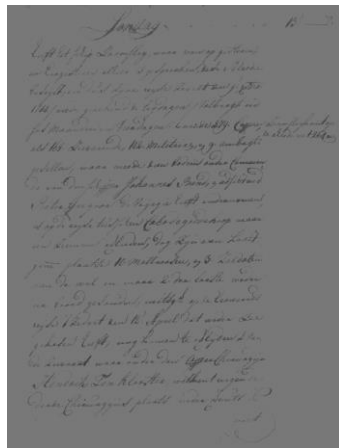
Gambar 4.2 Citra Paleografi asli



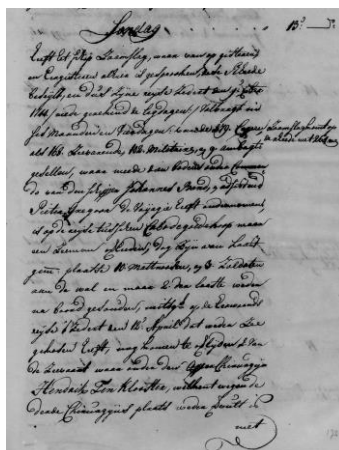
Gambar 4.3 Citra Threshold Binary



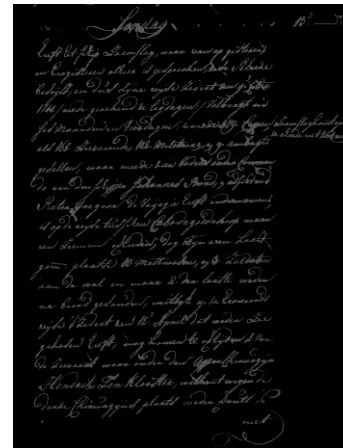
Gambar 4.4 Citra Threshold Binary Inverted



Gambar 4.5 Citra Threshold Truncate



Gambar 4.6 Citra Threshold To Zero

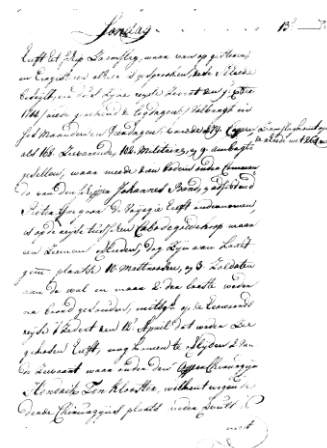


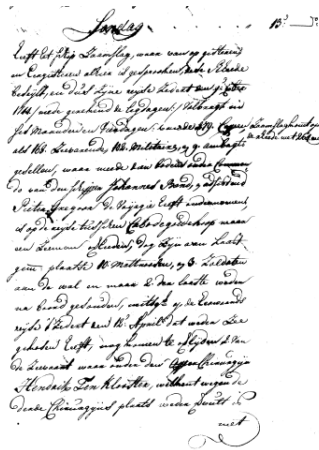
Gambar 4.7 Citra Threshold To Zero Inverted

4.2 Pembahasan

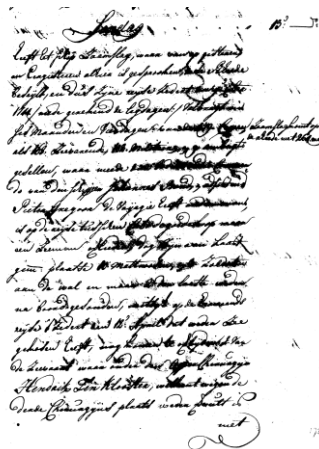
Segmentasi citra digital paleografi arsip VOC menggunakan metode *thresholding binary* menghasilkan segmentasi citra digital paleografi yang paling baik. Kemudian hasil segmentasi citra diuji kembali dengan melakukan perubahan pada nilai ambang $T=100$, $T=120$ dan $T=140$. Hasil dari percobaan dapat dilihat pada gambar-gambar dibawah.

Dari hasil penelitian nilai ambang $T=120$ menghasilkan segmentasi citra digital paleografi VOC yang paling baik.

Gambar 4.8 Citra Threshold Binary $T=100$



Gambar 4.9 Citra Threshold Binary T=120



Gambar 4.10 Citra Threshold Binary T=140

Kesimpulan

Penelitian ini telah memaparkan perbandingan hasil dari metode *thresholding* untuk segmentasi citra digital paleografi arsip VOC. *Thresholding binary* menghasilkan segmentasi citra digital yang paling baik karena dapat memisahkan objek tulisan paleografi dengan backgroundnya. Nilai $T=120$ merupakan segmentasi citra digital paleografi terbaik diantara nilai ambang T yang lainnya.

DAFTAR PUSTAKA

- [1] E. Maria, Y. Yulianto, Y. P. Arinda, J. Jumiatty, and P. Nobel, "Segmentasi Citra Digital Bentuk Daun Pada Tanaman Di Politani Samarinda Menggunakan Metode Thresholding," *J. Rekayasa Teknol. Inf.*, vol. 2, no. 1, p. 37, 2018, doi: 10.30872/jurti.v2i1.1377.
- [2] A. Kaehler and G. Bradski, *Learning OpenCV 3 Computer Vision in C++ with the OpenCV Library*. Sebastopol: O'Reilly Media, Inc, 2017.
- [3] T. Sutoyo, E. Mulyanto, V. Suhartono, O. D. Nurhayati, and Wijanarto, "Teori Pengolahan Citra Digital," Yogyakarta: Penerbit Andi, 2009.
- [4] Darma Putra, *Pengolahan Citra Digital*. Yogyakarta: Penerbit Andi, 2010.
- [5] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, 4th ed. New York: Pearson, 2018.
- [6] Rinaldi Munir, *Pengolahan Citra Digital dengan pendekatan Algoritmik*. Bandung: Informatika, 2004.
- [7] "OpenCV." [Online]. Available: https://docs.opencv.org/4.x/db/d8e/tutorial_threshold.html.