Presisi: Jurnal Teknologi Rekayasa dan Inovasi Teknik
Volume: 1, Nomor: 3, (2026), September: 202 - 220
Website: https://ejournal.ojs-lppm.com/index.php/Presisi/article/view/560
E-ISSN: 3164-4694

Analisis Matematis dan Validasi Numerik Konvolusi 2D Operator Sobel untuk Ekstraksi Fitur Tepi pada Citra Digital

Ahmad Robi Faro'id[1]
Universitas Jember, Indonesia [1]
Corresponding Author: *rrobbyxx@gmail.com*[1*]

Info Artikel
Received: *02 September 2026*
Revised : *15 September 2026*
Accepted: *23 September 2026*
Published: *30 September 2026*

Keywords: *2D convolution, Sobel operator, edge detection, gradient magnitude, feature map, digital image.*

Abstract

Edge detection is a fundamental step in digital image processing and computer vision. This study mathematically analyzes the 2D convolution operation using the Sobel operator to extract edge features in latte art images. The color image is cropped into a region of interest (ROI) measuring 530×380 pixels, converted to grayscale, and then convolved with a 3×3 Sobel kernel in both the horizontal and vertical directions without padding and with a stride of 1, resulting in a feature map measuring 378×528. Edge strength is calculated using the gradient magnitude. Manual calculations on a 9×9 sample yielded Gx = −80, Gy = −130, and G ≈ 152.64. Validation at three points within the ROI image confirmed consistency with the program, and a comparison of the entire feature map against the two-loop convolution implementation resulted in a maximum absolute error of 3.4×10^{-13}. In the coffee surface area, the gradient magnitude has a mean of 19.312, a median of 12.927, a standard deviation of 20.159, and a maximum value of 289.744. A 95th percentile-based threshold was used to mark the strongest edge responses, with a value of 59.862 in the coffee surface area and 79.891 across the entire ROI. The results show that most pixels exhibit low intensity changes, while high responses are concentrated in a small fraction of pixels. This study is descriptive in nature, based on a single image, and does not evaluate the accuracy of edge detection against a reference edge.

Kata Kunci*:* konvolusi 2D, operator Sobel, deteksi tepi, magnitude gradient, feature map, citra digital.

Abstrak

Deteksi tepi merupakan tahap dasar dalam pengolahan citra digital dan visi komputer. Penelitian ini menganalisis secara matematis operasi konvolusi 2D menggunakan operator Sobel untuk mengekstraksi fitur tepi pada citra latte art. Citra berwarna dipotong menjadi region of interest (ROI) berukuran 530×380 piksel, dikonversi ke grayscale, lalu dikonvolusi dengan kernel Sobel 3×3 arah horizontal dan vertikal tanpa padding dengan stride 1, sehingga menghasilkan feature map berukuran 378×528. Kekuatan tepi dihitung dengan magnitude gradient. Perhitungan manual pada sampel 9×9 menghasilkan Gx = −80, Gy = −130, dan G ≈ 152,64. Validasi pada tiga titik citra ROI menunjukkan kesesuaian dengan program, dan perbandingan seluruh feature map terhadap implementasi konvolusi dua loop menghasilkan galat absolut maksimum $3{,}4\times10^{-13}$. Pada area permukaan kopi, magnitude gradient memiliki rata-rata 19,312, median 12,927, simpangan baku 20,159, dan nilai maksimum 289,744. Threshold berbasis persentil ke-95 digunakan untuk menandai respons tepi terkuat, dengan nilai 59,862 pada area permukaan kopi dan 79,891 pada seluruh ROI. Hasil menunjukkan bahwa sebagian besar piksel memiliki perubahan intensitas rendah, sedangkan respons

tinggi terkonsentrasi pada sebagian kecil piksel. Penelitian ini bersifat deskriptif pada satu citra dan tidak mengevaluasi akurasi deteksi terhadap tepi acuan.

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.
Publisher: Lembaga Penerbit Penelitian Multidisipliner

PENDAHULUAN

Citra digital menjadi sumber informasi utama di banyak bidang, seperti pengenalan objek, pencitraan medis, robotika, dan kendaraan otonom. Agar dapat dianalisis secara otomatis, informasi visual perlu diringkas menjadi fitur yang bermakna. Salah satu fitur paling mendasar adalah tepi (*edge*), yaitu lokasi tempat intensitas citra berubah secara tajam [1]. Tepi umumnya berkaitan dengan batas objek, perubahan orientasi permukaan, perbedaan pencahayaan, atau perubahan tekstur, sehingga menjadi dasar bagi segmentasi, deteksi kontur, dan analisis bentuk [1], [2]. Teori deteksi tepi telah dirumuskan sejak awal oleh Marr dan Hildreth [3] dan kemudian diperkuat oleh kriteria optimalitas Canny [11][13]. Sejak itu, ekstraksi tepi dipandang sebagai tahap awal yang penting dalam pengolahan citra dan visi komputer.

Citra grayscale 8-bit dapat dipandang sebagai fungsi intensitas diskrit I(x,y) dengan nilai pada rentang 0 sampai 255, dengan 0 untuk hitam dan 255 untuk putih [12]. Bila citra diperoleh dari konversi warna dengan bobot luminans, nilainya dapat berupa bilangan desimal kecuali dibulatkan. Pada wilayah homogen, intensitas hampir konstan sehingga turunannya mendekati nol. Pada batas dua wilayah berintensitas berbeda, turunannya bernilai besar [10]. Karena citra bersifat diskrit, turunan diaproksimasi dengan selisih antar piksel bertetangga, dan aproksimasi ini dapat dinyatakan sebagai konvolusi 2D antara citra dan sebuah kernel kecil. Untuk citra I dan kernel K berukuran k×k, konvolusi 2D didefinisikan sebagai [1]

$$(I * K)(i,j) = \sum_{m}\sum_{n} I\,(i-m,\, j-n)\, K(m,n).$$

Persamaan tersebut setara dengan menggeser kernel yang telah diputar $180°$ pada citra, lalu menjumlahkan hasil kali elemen pada setiap posisi. Perlu dicatat bahwa banyak pustaka pengolahan citra dan pembelajaran mesin sebenarnya mengimplementasikan *cross-correlation*, yaitu operasi yang sama tanpa pembalikan kernel, tetapi tetap menyebutnya konvolusi [8]. Untuk kernel Sobel, perbedaan ini hanya membalik tanda respons gradien dan tidak memengaruhi besarnya.

Bila citra berukuran $H \times W$ dikonvolusi dengan kernel $K \times K$, dengan padding P dan stride S, ukuran keluarannya adalah

$$H_{out} = \frac{H - K + 2P}{S} + 1, \qquad W_{out} = \frac{W - K + 2P}{S} + 1.$$

Pada konvolusi *valid* (P = 0, S = 1), ukuran keluaran mengecil sebesar K−1 pada tiap dimensi karena kernel tidak boleh keluar dari batas citra [1]. Biaya komputasi konvolusi langsung untuk citra M×N dan kernel K×K adalah $O(M \cdot N \cdot K^2)$.

Di antara berbagai kernel penaksir turunan, operator Sobel [4], [5] termasuk yang paling luas digunakan, bersama operator Prewitt [6] yang serupa. Operator ini memakai dua kernel 3×3 untuk menaksir gradien pada arah horizontal dan vertikal:

$$K_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \qquad K_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

Setiap kernel dapat diuraikan menjadi perkalian antara vektor diferensiasi [−1 0 1] dan vektor penghalus $[1\ 2\ 1]^T$ [9]. Dengan demikian Sobel menghitung selisih intensitas sekaligus merata-ratakan piksel tetangga secara berbobot, sehingga lebih tahan terhadap *noise* dibandingkan selisih piksel sederhana. Sifat terpisahkan ini juga memungkinkan komputasi diturunkan dari $O(M \cdot N \cdot K^2)$ menjadi $O(M \cdot N \cdot K)$. Respons kedua arah dinyatakan sebagai Gx = I * Kx dan Gy = I * Ky, lalu digabungkan menjadi magnitude gradient

$$G = \sqrt{G_x^2 + G_y^2},$$

yang menyatakan kekuatan tepi pada suatu piksel. Magnitude ini dapat didekati dengan $G \approx |Gx| + |Gy|$ untuk komputasi yang lebih sederhana [1], [2]. Arah gradien diperoleh dari θ = atan2 (Gy, Gx). Pada konvolusi murni, tanda Gx dan Gy berkebalikan dengan konvensi gradien yang lazim, sehingga θ bergeser 180°, sedangkan G tidak berubah. Feature map G selanjutnya dapat diberi ambang (*threshold*) untuk memisahkan piksel tepi dari piksel non-tepi.

Sobel hanya memakai turunan orde pertama tanpa penipisan tepi, sehingga tepi yang dihasilkan cenderung tebal dan sensitif terhadap tekstur halus. Metode yang lebih kompleks seperti Canny [7] menambahkan penghalusan Gaussian, *non-maximum suppression*, dan *hysteresis thresholding* untuk mengatasi keterbatasan ini.

Meskipun Sobel sangat umum, banyak tutorial dan implementasi hanya memanggilnya sebagai fungsi pustaka tanpa menjabarkan dasar matematisnya, yaitu definisi konvolusi, peran pembalikan kernel, perhitungan ukuran keluaran, dan hubungan antara nilai piksel dengan respons gradien. Akibatnya, pengguna sulit memastikan bahwa keluaran program benar-benar sesuai dengan definisinya, termasuk dalam hal tanda dan ukuran [11]. Penelitian ini memakai foto latte art sebagai

objek uji karena citra tersebut memuat kombinasi permukaan halus (foam dan kopi) dengan batas tajam (tepi cangkir dan pola), sehingga distribusi respons gradient yang beragam dapat diamati pada satu citra [9].

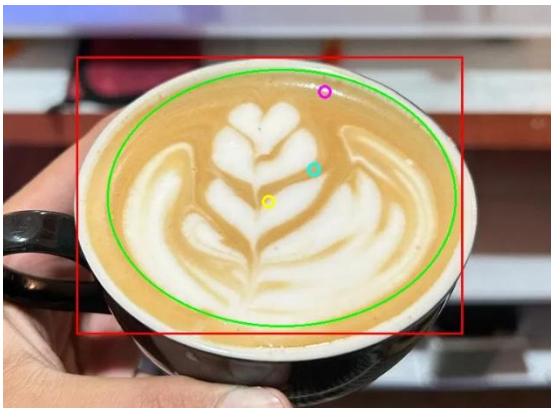
Berdasarkan hal tersebut, penelitian ini bertujuan untuk (1) menjabarkan operasi konvolusi 2D dan hubungannya dengan operator Sobel untuk ekstraksi fitur tepi; (2) menunjukkan perhitungan konvolusi secara manual dari matriks piksel citra nyata; (3) memvalidasi implementasi program terhadap perhitungan manual dengan melaporkan selisih numeriknya; dan (4) menganalisis distribusi magnitude gradient serta menentukan piksel tepi kuat dengan threshold berbasis persentil. Penelitian ini bersifat deskriptif pada satu citra dan tidak mengevaluasi akurasi deteksi terhadap tepi acuan.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan eksperimen komputasional yang dilengkapi perhitungan manual. Perhitungan manual dipakai untuk menunjukkan dasar matematis konvolusi, sedangkan program dipakai untuk memproses seluruh citra. Hasil keduanya dibandingkan secara numerik untuk validasi.

Data Citra

Data yang digunakan adalah satu foto latte art berwarna berukuran 960 × 1280 piksel (lebar × tinggi, orientasi potret). Dari foto tersebut diambil *region of interest* (ROI) berukuran 530 × 380 piksel (lebar × tinggi) yang berfokus pada cangkir dan latte art. Sudut kiri-atas ROI berada pada koordinat $(x, y) = (250, 470)$pada foto asli, sehingga ROI mencakup $x = 250$–779dan $y = 470$–849. Pemotongan ini bertujuan memusatkan analisis pada objek utama. Letak ROI ditunjukkan pada Gambar 1.



Tahapan Penelitian

Penelitian dilakukan dalam urutan berikut: (1) citra asli, (2) pemotongan ROI, (3) konversi grayscale, (4) representasi matriks intensitas, (5) konvolusi Sobel untuk memperoleh $G_x$dan G_y, (6) perhitungan *magnitude gradient*, (7) perhitungan manual dan validasi terhadap program, (8) analisis statistik, (9) penentuan *threshold* persentil ke-95, dan (10) pembentukan *strong-edge map*.

Konversi Grayscale

Citra berwarna dikonversi menjadi grayscale dengan pembobotan luminans standar [1]:

$$I(x,y) = 0{,}299R(x,y) + 0{,}587G(x,y) + 0{,}114B(x,y)$$

Hasil konversi disimpan sebagai bilangan desimal presisi ganda (*float64*) tanpa pembulatan, sehingga nilai intensitas berupa pecahan pada rentang 0–255. Pembulatan tidak dilakukan agar tidak menambah galat sebelum konvolusi. Bilangan bulat pada sampel 9 × 9 di bagian Hasil hanyalah pembulatan untuk keperluan penyajian. Dengan konversi ini, citra direpresentasikan sebagai satu matriks intensitas sehingga analisis dapat difokuskan pada konvolusi 2D tanpa memproses tiga kanal warna.

Konvolusi Sobel

Matriks grayscale dikonvolusi dengan kernel Sobel $K_x$dan $K_y$berukuran 3 × 3 seperti pada Pendahuluan. Konvolusi dilakukan secara valid, yaitu tanpa *padding* $(P = 0)$dan dengan *stride* 1 $(S = 1)$.

Untuk ROI dengan tinggi $H = 380$dan lebar $W = 530$, ukuran keluarannya adalah:

$$H_{\text{out}} = \left\lfloor \frac{380 - 3 + 0}{1} \right\rfloor + 1 = 378$$

$$W_{\text{out}} = \left\lfloor \frac{530 - 3 + 0}{1} \right\rfloor + 1 = 528$$

Dengan demikian, setiap *feature map* $(G_x, G_y,$ dan $G)$berukuran 378 × 528 atau sebanyak:

$$378 \times 528 = 199.584$$

piksel.

Konvolusi mengikuti definisi matematis dengan pembalikan kernel 180°, sehingga kernel yang dipakai dalam perhitungan adalah:

$$K_x^* = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

$$K_y^* = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

Karena:

$$K_x^* = -K_x$$

dan

$$K_y^* = -K_y$$

hasil konvolusi murni sama dengan hasil *cross-correlation* dengan tanda berlawanan. *Magnitude gradient* dihitung dengan:

$$G = \sqrt{G_x^2 + G_y^2}$$

Kompleksitas komputasi langsung adalah:

$$O(M \cdot N \cdot K^2)$$

dengan $M = 380$, $N = 530$, dan $K = 3$.

Konvensi Koordinat

Seluruh indeks berbasis nol, yaitu baris dan kolom pertama bernomor 0. Posisi (r, c)pada *feature map* menyatakan hasil konvolusi dengan pusat kernel di piksel $(r, +1c + 1)$pada ROI karena konvolusi valid kehilangan satu piksel di tiap tepi.

Piksel tersebut berkorespondensi dengan piksel:

$$(x, y) = (c, +1 + 250\, r + 1 + 470)$$

pada foto asli.

Sebagai contoh, titik $(r, c) = (47, 339)$pada *feature map* berpusat di piksel $(48, 340)$pada ROI dan berkorespondensi dengan koordinat:

$$(x, y) = (590, 518)$$

pada foto asli.

Perhitungan Manual

Perhitungan manual dilakukan pada dua tingkat.

1. Sampel 9 × 9

Sampel matriks 9 × 9 diambil dari area *foam latte art*, dengan sudut kiri-atas pada baris 180

dan kolom 274 di ROI. Posisi tersebut berkorespondensi dengan $x = 524$dan $y = 650$pada foto asli.

Kernel digeser pada seluruh matriks sehingga menghasilkan *feature map* 7 × 7 karena:

$$9 - 3 + 1 = 7$$

Posisi tengah, yaitu $(3, 3)$pada *feature map* 7 × 7, dihitung secara rinci. Tabel G_x, G_y, dan Guntuk seluruh 49 posisi disajikan pada Hasil.

2. Titik pada Citra ROI

Perhitungan manual dilakukan pada tiga titik pada *feature map* ROI yang dipilih sengaja agar mewakili respons yang berbeda, yaitu (a) titik respons tinggi pada $(47, 339)$, yaitu peralihan dari bibir cangkir yang terang ke permukaan kopi yang lebih gelap, (b) titik area halus pada $(153, 324)$, dan (c) titik area peralihan pada $(197, 262)$.

Pada setiap titik, submatriks 3 × 3 diambil dari matriks grayscale, lalu G_x, G_y, dan Gdihitung dengan tangan menggunakan K_x^* dan K_y^*.

Implementasi Program

Seluruh proses pada citra dijalankan menggunakan Python 3.12 dengan pustaka NumPy, SciPy (scipy.signal.convolve2d, mode valid), dan Pillow [sesuaikan dengan versi yang Anda pakai]. Fungsi convolve2d mengimplementasikan konvolusi murni dengan pembalikan kernel, sehingga tanda hasilnya sesuai dengan K_x^*dan K_y^*. Fungsi lain seperti cv2.filter2D dan cv2.Sobel bekerja sebagai *cross-correlation* dan menghasilkan tanda berlawanan.

Validasi

Validasi dilakukan pada tiga tingkat.

1. Titik tunggal

Nilai G_x, G_y, dan Gdari program pada posisi yang sama dengan perhitungan manual dibandingkan.

2. Seluruh Feature Map

Keluaran pustaka dibandingkan elemen demi elemen dengan implementasi konvolusi dua *loop* sederhana yang ditulis langsung dari definisi. Galat absolut maksimum dilaporkan untuk G_x, G_y, dan G:

$$\varepsilon = \max\left|G_{\text{pustaka}}(i,j) - G_{\text{loop}}(i,j)\right|$$

3. Konsistensi Konvensi

Hasil konvolusi murni dibandingkan dengan *cross-correlation* untuk memeriksa bahwa:

$$G_{x,\text{konv}} = -G_{x,\text{korelasi}}$$

$$G_{y,\text{konv}} = -G_{y,\text{korelasi}}$$

sedangkan magnitude gradient memiliki nilai identik:

$$G_{\text{konv}} = G_{\text{korelasi}}$$

Kesesuaian dinyatakan tercapai bila:

$$\varepsilon < 10^{-9}$$

Toleransi tersebut dianggap wajar untuk aritmetika *float64*.

Analisis Statistik dan Threshold

Distribusi *magnitude gradient* diringkas menggunakan nilai minimum, maksimum, rata-rata, median, simpangan baku populasi, serta persentil ke-90, ke-95, dan ke-99.

Rata-rata dihitung menggunakan:

$$\mu_G = \frac{1}{N}\sum_{i=1}^{N} G_i$$

Simpangan baku populasi dihitung menggunakan:

$$\sigma_G = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(G_i - \mu_G)^2}$$

Persentil dihitung dengan interpolasi linear menggunakan bawaan numpy.percentile.

Statistik dihitung pada dua lingkup.

Lingkup A: Area Permukaan Kopi

Dipakai mask elips pada koordinat *feature map* dengan pusat:

$$(r_0, c_0) = (192, 280)$$

Mask memiliki semi-sumbu vertikal 176 piksel dan semi-sumbu horizontal 239 piksel. Piksel (r, c)masuk mask bila memenuhi:

$$\left(\frac{r - r_0}{176}\right)^2 + \left(\frac{c - c_0}{239}\right)^2 \leq 1$$

Mask memuat 132.127 piksel. Mask ini dibuat agar statistik tidak didominasi tangan, meja, dan latar belakang. Titik $(47, 339)$berada di dalam mask, dekat tepi atas elips, sehingga mask masih memuat sebagian peralihan bibir cangkir.

Lingkup B: Seluruh Feature Map ROI

Seluruh *feature map* ROI berjumlah 199.584 piksel dan digunakan untuk membentuk *strong-edge map*.

Piksel tepi kuat ditentukan dengan *threshold* berbasis persentil:

$$T = P_{95}$$

Piksel dengan:

$$G > T$$

dikategorikan sebagai tepi kuat.

Threshold persentil dipilih karena nilainya menyesuaikan distribusi *gradient* pada citra, bukan ditetapkan sebagai angka tetap. Karena P_{95}dihitung pada dua lingkup yang berbeda, nilainya juga berbeda. Oleh karena itu, lingkup perhitungan selalu disebutkan bersama nilai *threshold* pada Hasil dan Pembahasan.

Nilai P_{95}menetapkan sekitar 5% piksel dengan nilai *gradient* tertinggi sebagai tepi kuat. Oleh karena itu, *threshold* ini mengatur proporsi piksel tepi dan tidak mengukur banyaknya tepi sebenarnya.

Visualisasi

Feature map $G_x$dan $G_y$dapat bernilai negatif. Untuk penayangan pada Gambar 4 dan 5, nilainya dinormalisasi secara linear ke rentang 0–255 dengan nilai nol dipetakan ke abu-abu tengah.

Magnitude gradient pada Gambar 6 dinormalisasi dengan membagi nilai terhadap nilai maksimum *feature map*. Normalisasi hanya digunakan untuk tampilan dan tidak memengaruhi statistik maupun *threshold*.

HASIL DAN PEMBAHASAN

Perhitungan Manual pada Sampel 9 × 9

Sampel matriks intensitas 9 × 9 diambil dari area *foam latte art* pada baris 180 dan kolom 274 di ROI, yang berkorespondensi dengan $x = 524$dan $y = 650$pada foto asli. Nilai intensitas dibulatkan ke bilangan bulat hanya untuk keperluan penyajian:

$$I = \begin{bmatrix} 170 & 172 & 174 & 176 & 179 & 179 & 180 & 180 & 183 \\ 172 & 177 & 179 & 179 & 176 & 179 & 183 & 187 & 192 \\ 173 & 178 & 178 & 174 & 175 & 181 & 189 & 197 & 204 \\ 174 & 178 & 175 & 171 & 184 & 191 & 200 & 209 & 214 \\ 179 & 182 & 184 & 185 & 203 & 207 & 214 & 218 & 220 \\ 184 & 189 & 197 & 204 & 218 & 220 & 223 & 223 & 223 \\ 194 & 200 & 206 & 211 & 218 & 219 & 220 & 222 & 224 \\ 200 & 204 & 210 & 214 & 219 & 220 & 221 & 223 & 225 \\ 208 & 211 & 215 & 218 & 221 & 222 & 223 & 224 & 226 \end{bmatrix}$$

Dengan konvolusi valid, ukuran keluaran adalah:

$$9 - 3 + 1 = 7$$

sehingga diperoleh *feature map* berukuran 7 × 7. Posisi tengah, yaitu $(3, 3)$pada *feature map* 7 × 7, dihitung secara rinci. Submatriks 3 × 3 pada posisi tersebut adalah:

$$I_{(3,3)} = \begin{bmatrix} 171 & 184 & 191 \\ 185 & 203 & 207 \\ 204 & 218 & 220 \end{bmatrix}$$

Dengan kernel terbalik K_x^*, respons arah horizontal adalah:

$$G_x = (171 - 191) + 2(185 - 207) + (204 - 220)$$

$$G_x = -20 - 44 - 16 = -80$$

Dengan kernel terbalik K_y^*, respons arah vertikal adalah:

$$G_y = (171 + 2 \cdot 184 + 191) - (204 + 2 \cdot 218 + 220)$$

$$G_y = 730 - 860 = -130$$

Magnitude gradient pada posisi tersebut adalah:

$$G = \sqrt{(-80)^2 + (-130)^2}$$

$$G = \sqrt{6400 + 16900}$$

$$G = \sqrt{23300} \approx 152{,}64$$

Tanda negatif pada $G_x$dan $G_y$muncul karena pembalikan kernel. Pada submatriks ini, intensitas meningkat dari kiri-atas ke kanan-bawah. Bila digunakan *cross-correlation* tanpa pembalikan kernel, diperoleh $G_x = 80$dan $G_y = 130$, sedangkan magnitude tetap sebesar 152,64. Dengan demikian, perbedaan antara konvolusi dan *cross-correlation* hanya memengaruhi tanda, bukan kekuatan tepi.

Program yang dijalankan pada matriks bulat menghasilkan nilai yang sama persis. Bila program menggunakan nilai desimal asli (*float64*) tanpa pembulatan, diperoleh $G_x = -79{,}841$, $G_y = -129{,}619$, dan $G = 152{,}236$. Selisih kecil tersebut hanya berasal dari pembulatan yang digunakan pada penyajian matriks.

Hasil penggeseran kernel pada seluruh matriks 9 × 9 disajikan pada Tabel 1 sampai Tabel 3.

Tabel 1. *Feature map $G_x(7 \times 7)$*

	0	1	2	3	4	5	6
0	−23	−4	4	−10	−29	−33	−36
1	−18	13	0	−34	−51	−58	−53
2	−12	15	−34	−69	−57	−63	−49
3	−24	−14	−68	−80	−43	−43	−26
4	−43	−44	−73	−62	−23	−20	−10
5	−47	−47	−54	−38	−11	−12	−12
6	−39	−38	−36	−24	−8	−11	−15

Tabel 2. *Feature map $G_y(7 \times 7)$*

	0	1	2	3	4	5	6
0	−19	−12	4	8	−9	−37	−64
1	0	15	12	−20	−49	−68	−83
2	−20	−27	−56	−93	−105	−97	−83
3	−54	−88	−122	−130	−115	−89	−60
4	−73	−88	−89	−68	−45	−28	−18
5	−59	−51	−34	−12	1	4	0
6	−45	−36	−26	−16	−12	−11	−9

Tabel 3. *Magnitude gradient $G(7 \times 7)$*

	0	1	2	3	4	5	6
0	29,83	12,65	5,66	12,81	30,36	49,58	73,43
1	18,00	19,85	12,00	39,45	70,72	89,38	98,48
2	23,32	30,89	65,51	115,80	119,47	115,66	96,38
3	59,09	89,11	139,67	152,64	122,78	98,84	65,39
4	84,72	98,39	115,11	92,02	50,54	34,41	20,59
5	75,43	69,35	63,81	39,85	11,05	12,65	12,00
6	59,55	52,35	44,41	28,84	14,42	15,56	17,49

Nilai pada posisi $(3, 3)$pada ketiga tabel, yaitu $G_x = -80$, $G_y = -130$, dan $G = 152{,}64$, sama

dengan hasil perhitungan manual. Magnitude tertinggi pada matriks 7 × 7 adalah 152,64, disusul 139,67 dan 122,78, yang berada di sekitar peralihan dari *foam* gelap ke *foam* terang. Nilai terendah adalah 5,66, 11,05, dan 12,00, yang berada pada area dengan intensitas relatif seragam.

Semua nilai $G_x$dan $G_y$bernilai negatif atau mendekati nol karena intensitas pada sampel ini secara umum meningkat ke arah kanan-bawah. Pola tersebut sesuai dengan prinsip dasar deteksi tepi, yaitu semakin besar perubahan intensitas dalam lingkungan piksel, semakin besar respons operator Sobel [1].

Ukuran Feature Map pada Citra Asli

Dengan ROI berukuran 530 × 380 piksel, kernel 3 × 3, $P = 0$, dan $S = 1$, *feature map* yang dihasilkan berukuran 378 × 528 atau sebanyak 199.584 piksel, sesuai dengan:

$$H_{\text{out}} = \left\lfloor \frac{380 - 3 + 0}{1} \right\rfloor + 1 = 378$$

$$W_{\text{out}} = \left\lfloor \frac{530 - 3 + 0}{1} \right\rfloor + 1 = 528$$

Ukuran tersebut mengecil dua piksel pada tiap dimensi karena kernel tidak dapat keluar dari batas citra. Bila pinggiran citra perlu dipertahankan, dapat digunakan *padding* $P = 1$sehingga ukuran keluaran sama dengan ukuran masukan.

Validasi pada Citra Asli

Validasi dilakukan pada tiga titik yang dipilih untuk mewakili respons berbeda seperti ditunjukkan pada Tabel 4.

Tabel 4. *Perhitungan manual pada tiga titik citra ROI*

Titik (r, c)	Letak pada foto asli (x, y)	Lokasi objek	Gx	Gy	G
A (47, 339)	(590, 518)	Peralihan bibir cangkir ke permukaan kopi	10,839	289,541	289,744
B (153, 324)	(575, 624)	Permukaan krema yang halus	0,005	1,747	1,747
C (197, 262)	(513, 668)	Batas foam putih dan krema	−23,440	−55,784	60,509

Titik A (respons tinggi). Submatriks grayscale 3 × 3 di sekitar titik ini adalah:

$$\begin{bmatrix} 243{,}510 & 248{,}417 & 245{,}140 \\ 197{,}278 & 196{,}278 & 188{,}022 \\ 172{,}861 & 172{,}089 & 178{,}904 \end{bmatrix}$$

Perhitungan manual dengan K_x^*memberikan:

$$G_x = (243{,}510 - 245{,}140) + 2(197{,}278 - 188{,}022) + (172{,}861 - 178{,}904)$$

$$G_x = 10{,}839$$

Dengan K_y^*:

$$G_y = (243{,}510 + 2 \cdot 248{,}417 + 245{,}140) - (172{,}861 + 2 \cdot 172{,}089 + 178{,}904)$$

$$G_y = 289{,}541$$

Magnitudenya adalah:

$$G = \sqrt{10{,}839^2 + 289{,}541^2} \approx 289{,}744$$

Titik ini memiliki $| G_y | \gg | G_x |$. Intensitas turun tajam dari baris atas, sekitar 243–248, menuju baris bawah, sekitar 172–178, sedangkan perubahan antar kolom relatif kecil. Artinya, perubahan intensitas dominan pada arah vertikal sehingga terbentuk tepi yang membentang horizontal. Pada citra, titik ini berada pada peralihan bibir cangkir yang terang menuju permukaan kopi yang lebih gelap. $G_y$bertanda positif karena pada konvolusi murni intensitas yang menurun ke arah bawah menghasilkan respons positif.

Titik B (area halus). Submatriks 3 × 3 adalah:

$$\begin{bmatrix} 177{,}991 & 179{,}350 & 179{,}350 \\ 181{,}094 & 180{,}795 & 180{,}795 \\ 178{,}735 & 178{,}795 & 177{,}969 \end{bmatrix}$$

Dari perhitungan diperoleh:

$$G_x = (177{,}991 - 179{,}350) + 2(181{,}094 - 180{,}795) + (178{,}735 - 177{,}969) = 0{,}005$$

Untuk arah vertikal:

$$G_y = 716{,}041 - 714{,}294 = 1{,}747$$

Magnitudenya adalah:

$$G = \sqrt{0{,}005^2 + 1{,}747^2} \approx 1{,}747$$

Intensitas hampir konstan, yaitu sekitar 178–181, sehingga respons gradient mendekati nol. Kondisi tersebut sesuai dengan prinsip bahwa wilayah homogen memiliki turunan yang kecil.

Titik C (area peralihan). Submatriks 3 × 3 adalah:

$$\begin{bmatrix} 200{,}904 & 204{,}278 & 209{,}022 \\ 209{,}908 & 212{,}880 & 215{,}624 \\ 217{,}108 & 218{,}080 & 220{,}998 \end{bmatrix}$$

Dari perhitungan diperoleh:

$$G_x = (200{,}904 - 209{,}022) + 2(209{,}908 - 215{,}624) + (217{,}108 - 220{,}998) = -23{,}440$$

Untuk arah vertikal:

$$G_y = 818{,}482 - 874{,}266 = -55{,}784$$

Magnitudenya adalah:

$$G = \sqrt{(-23{,}440)^2 + (-55{,}784)^2} \approx 60{,}509$$

Intensitas meningkat dari kiri-atas menuju kanan-bawah sehingga kedua respons bertanda negatif. Nilai Gpada titik ini sedikit di atas P_{95}area permukaan kopi, yaitu 59,862. Oleh karena itu, titik C tergolong tepi kuat pada lingkup tersebut, tetapi bukan pada lingkup seluruh ROI.

Validasi program. Nilai program pada ketiga titik sama dengan perhitungan manual, dengan selisih tidak lebih dari 0,0005 yang berasal dari pembulatan tiga desimal pada perhitungan manual. Untuk seluruh *feature map*, program dibandingkan dengan implementasi konvolusi dua *loop* yang ditulis langsung dari definisi. Galat absolut maksimumnya adalah:

$$\varepsilon_{G_x} = 2{,}3 \times 10^{-13}$$

$$\varepsilon_{G_y} = 3{,}1 \times 10^{-13}$$

$$\varepsilon_{G} = 3{,}4 \times 10^{-13}$$

Seluruh nilai tersebut jauh di bawah toleransi 10^{-9}. Hasil konvolusi murni juga sama dengan negatif hasil *cross-correlation*, dengan selisih maksimum $3{,}4 \times 10^{-13}$untuk G_x, G_y, dan G. Hal tersebut mengonfirmasi bahwa perbedaan keduanya hanya terdapat pada tanda.

Statistik Magnitude Gradient

Mask elips memuat 132.127 piksel atau 66,2% dari seluruh *feature map*. Statistik pada area permukaan kopi disajikan pada Tabel 5.

Tabel 5. *Statistik magnitude gradient pada area permukaan kopi (mask elips)*

Parameter	Nilai
Minimum	0,000
Maksimum	289,744
Rata-rata	19,312
Median	12,927

Simpangan baku	20,159
Persentil 90	44,021
Persentil 95	59,862
Persentil 99	91,707

Rata-rata sebesar 19,312 jauh di bawah nilai maksimum 289,744, sedangkan median sebesar 12,927 lebih kecil daripada rata-rata. Kondisi tersebut menunjukkan distribusi yang miring ke kanan, yaitu sebagian besar piksel memiliki perubahan intensitas rendah, sedangkan respons tinggi hanya muncul pada sebagian kecil piksel.

Persentil ke-99 sebesar 91,707 masih jauh di bawah nilai maksimum 289,744 sehingga respons yang sangat tinggi tergolong jarang. Simpangan baku sebesar 20,159 yang sedikit lebih besar daripada rata-rata juga menunjukkan sebaran nilai yang cukup lebar. Kondisi ini wajar pada citra latte art karena permukaan kopi dan *foam* umumnya relatif halus, sedangkan perubahan intensitas yang tajam muncul pada batas pola.

Nilai maksimum pada mask sebesar 289,744 merupakan titik A yang berada di dekat tepi atas elips pada peralihan bibir cangkir. Oleh karena itu, nilai maksimum tersebut mencerminkan batas cangkir yang masih terjangkau oleh mask, bukan pola latte art.

Threshold dan Strong-Edge Map

Threshold ditentukan menggunakan $T = P_{95}$. Dua lingkup analisis menghasilkan nilai threshold yang berbeda.

Pada area permukaan kopi:

$$T_{\text{mask}} = P_{95} = 59{,}862$$

Sebanyak 6.607 piksel atau sekitar 5% berada di atas ambang tersebut.

Pada seluruh *feature map* ROI:

$$T_{\text{ROI}} = P_{95} = 79{,}891$$

Dengan jumlah piksel sebanyak 199.584, sebanyak 9.980 piksel atau tepat 5% ditandai sebagai tepi kuat.

Nilai P_{95} seluruh ROI lebih tinggi karena ROI memuat tepi cangkir, tangan, dan latar yang kaya detail. Nilai maksimum seluruh ROI mencapai 920,98 pada posisi $(307, 31)$, yaitu di luar mask pada tepi kiri-bawah cangkir. Nilai tersebut hampir tiga kali lipat maksimum pada mask.

Dari 9.980 piksel tepi kuat pada seluruh ROI, sebanyak 2.463 piksel atau 24,7% berada di dalam mask, sedangkan 7.517 piksel atau 75,3% berada di luar mask. Dengan demikian, threshold seluruh

ROI hanya menandai sekitar 1,9% piksel pada area permukaan kopi, sedangkan sebagian besar tepi kuat berasal dari tepi cangkir dan latar.

Karena lingkup perhitungannya berbeda, kedua nilai threshold tidak dapat dibandingkan secara langsung. *Strong-edge map* pada seluruh ROI dibentuk menggunakan:

$$T = 79{,}891$$

Hasil visual tahapan pemrosesan ditunjukkan pada Gambar 3 sampai Gambar 8.



Gambar 3. *Citra grayscale ROI*



Gambar 4. *Feature map* G_x



Gambar 5. *Feature map* G_y



Gambar 6. *Magnitude gradient*



Gambar 7. *Strong-edge map (threshold* $P_{95} = 79{,}891$*)*



Gambar 8. *Overlay strong-edge pada citra asli ROI*

Feature map $G_x$pada Gambar 4 merespons perubahan intensitas pada arah horizontal sehingga menonjolkan tepi yang membentang vertikal. Sebaliknya, $G_y$pada Gambar 5 menonjolkan tepi yang membentang horizontal. *Magnitude gradient* pada Gambar 6 menggabungkan kedua respons sehingga tepi dari berbagai orientasi dapat ditampilkan secara bersamaan. *Strong-edge map* pada Gambar 7 hanya mempertahankan 5% respons tertinggi sehingga tepi dominan lebih mudah diamati, sedangkan *overlay* pada Gambar 8 memperlihatkan posisi tepi tersebut terhadap objek pada citra asli.

Pembahasan

Pertama, kesesuaian hasil manual dan program, baik pada sampel 9 × 9 maupun pada tiga titik di citra ROI, menunjukkan bahwa implementasi konvolusi sesuai dengan definisi matematisnya, termasuk aturan ukuran keluaran dan perlakuan tanda akibat pembalikan kernel. Validasi pada seluruh *feature map* dengan galat maksimum $3{,}4 \times 10^{-13}$memperkuat kesimpulan tersebut karena tidak hanya titik terpilih yang diperiksa. Hasil ini penting karena banyak pustaka menggunakan *cross-*

correlation, dan penelitian ini menunjukkan bahwa perbedaannya hanya memengaruhi tanda, bukan *magnitude gradient* yang digunakan untuk menilai kekuatan tepi.

Kedua, distribusi *magnitude* yang miring ke kanan menunjukkan bahwa sebagian besar piksel permukaan kopi memiliki gradient rendah, sementara respons tinggi terkonsentrasi pada batas pola dan tepi cangkir. Kondisi ini menjadi alasan pemilihan threshold berbasis persentil yang menyesuaikan diri dengan sebaran nilai pada citra, sedangkan ambang tetap harus ditentukan tanpa mempertimbangkan distribusi aktual. Penelitian ini belum membandingkan threshold persentil dengan ambang tetap atau metode otomatis seperti Otsu sehingga keunggulannya belum diuji. Selain itu, P_{95}menetapkan jumlah piksel tepi kuat secara relatif tetap, yaitu sekitar 5%, terlepas dari jumlah tepi yang sebenarnya terdapat pada citra. Threshold tersebut menunjukkan posisi relatif respons tertinggi dan tidak secara langsung menunjukkan jumlah tepi sebenarnya.

Ketiga, perbedaan P_{95}antara area permukaan kopi sebesar 59,862 dan seluruh ROI sebesar 79,891 memperlihatkan bahwa statistik gradient sangat bergantung pada lingkup analisis. Sebanyak 75,3% piksel tepi kuat pada seluruh ROI berada di luar mask sehingga *strong-edge map* seluruh ROI terutama menandai tepi cangkir dan latar, sedangkan pola latte art memiliki respons yang lebih rendah. Bila tujuan analisis adalah mengekstraksi pola pada permukaan kopi, threshold sebaiknya dihitung pada lingkup permukaan kopi dengan $T = 59{,}862$. Pemilihan lingkup perlu disesuaikan dengan tujuan analisis dan dilaporkan secara eksplisit.

Keempat, terdapat beberapa keterbatasan yang perlu dicatat. Operator Sobel hanya menggunakan turunan orde pertama tanpa penghalusan Gaussian dan tanpa penipisan tepi sehingga tepi yang dihasilkan cenderung lebih tebal dan sensitif terhadap tekstur halus serta noise pada latar [1], [2]. Metode seperti Canny [7] mengatasi kondisi tersebut melalui *non-maximum suppression* dan *hysteresis thresholding*. Penelitian ini hanya menggunakan satu citra tanpa *ground truth* tepi sehingga evaluasi bersifat deskriptif terhadap distribusi gradient, bukan terhadap akurasi deteksi. Mask elips juga ditentukan secara manual dan masih memuat sebagian bibir cangkir sehingga statistik pada Tabel 5 sedikit dipengaruhi oleh tepi cangkir. Penelitian lanjutan dapat menggunakan beberapa citra, membandingkan Sobel dengan Prewitt [6] dan Canny [7], serta menggunakan metrik evaluasi terhadap tepi acuan.

SIMPULAN

Pertama, konvolusi 2D dengan kernel Sobel 3×3 tanpa padding dan dengan stride 1 mengubah ROI 530×380 piksel menjadi feature map 378×528 piksel, sesuai dengan H_out = ⌊(H − K + 2P)/S⌋

+ 1. Pada sampel 9×9, feature map yang dihasilkan berukuran 7×7.

Kedua, perhitungan manual sesuai dengan keluaran program. Pada sampel 9×9 diperoleh Gx = −80, Gy = −130, dan G ≈ 152,64. Pada tiga titik citra ROI, nilai manual sama dengan program, dan pada seluruh feature map galat maksimum antara pustaka dan implementasi loop adalah $3{,}4\times10^{-13}$. Perbedaan antara konvolusi murni dan cross-correlation hanya membalik tanda respons gradient, sedangkan magnitude tidak berubah.

Ketiga, distribusi magnitude gradient pada area permukaan kopi miring ke kanan, dengan rata-rata 19,312, median 12,927, simpangan baku 20,159, dan maksimum 289,744. Sebagian besar piksel memiliki perubahan intensitas rendah, sedangkan respons tinggi hanya muncul pada sebagian kecil piksel.

Keempat, threshold persentil ke-95 menandai 5% piksel dengan respons tertinggi pada lingkup yang dipilih. Pada seluruh ROI, T = 79,891 menandai 9.980 dari 199.584 piksel, dan pada area permukaan kopi, T = 59,862. Karena nilai threshold bergantung pada lingkup, lingkup perhitungan harus selalu dilaporkan bersama nilainya. Penelitian ini tidak menguji ketepatan threshold tersebut terhadap tepi acuan.

Keterbatasan: analisis hanya memakai satu citra tanpa ground truth, mask ditentukan manual, dan Sobel menghasilkan tepi yang tebal serta sensitif terhadap tekstur halus. Penelitian selanjutnya dapat memakai beberapa citra, membandingkan Sobel dengan Prewitt dan Canny, serta memakai metrik evaluasi terhadap tepi acuan.

DAFTAR PUSTAKA

R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. New York: Pearson, 2018.

R. Szeliski, *Computer Vision: Algorithms and Applications*, 2nd ed. Cham: Springer, 2022.

D. Marr and E. Hildreth, "Theory of edge detection," *Proc. R. Soc. Lond. B*, vol. 207, no. 1167, pp. 187–217, 1980.

R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.

I. Sobel, "An isotropic 3×3 image gradient operator," in *Machine Vision for Three-Dimensional Scenes*, H. Freeman, Ed. San Diego: Academic Press, 1990, pp. 376–379.

J. M. S. Prewitt, "Object enhancement and extraction," in *Picture Processing and Psychopictorics*, B. S. Lipkin and A. Rosenfeld, Eds. New York: Academic Press, 1970, pp. 75–149.

J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, 1986.

I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.

A. Flores-Vidal, J. Castro, and D. Gómez, "Postprocessing of edge detection algorithms with machine learning techniques," *Mathematical Problems in Engineering*, vol. 2022, Art. no. 9729343, 2022, doi: 10.1155/2022/9729343.

K. M. R. A. Utama, R. Umar, and A. Yudhana, "Edge detection comparative analysis using Roberts, Sobel, Prewitt, and Canny methods," *Jurnal Teknologi dan Sistem Komputer*, vol. 10, no. 2, pp. 67–71, 2022, doi: 10.14710/jtsiskom.2022.14209.

R. A. A. S. Remya and S. Gopalan, "Comparative analysis of eight direction Sobel edge detection algorithm for brain tumor MRI images," *Procedia Computer Science*, vol. 201, pp. 487–494, 2022, doi: 10.1016/j.procs.2022.03.063.

Q. Chang, X. Li, Y. Li, and J. Miyazaki, "Multi-directional Sobel operator kernel on GPUs," *Journal of Parallel and Distributed Computing*, vol. 177, pp. 160–170, 2023, doi: 10.1016/j.jpdc.2023.03.004.

M. Mahalingam *et al.*, "Parallel processing of Sobel edge detection on FPGA: Enhancing real-time image analysis," *Sensors*, vol. 25, no. 12, Art. no. 3649, 2025, doi: 10.3390/s25123649.