

KOMBINASI METODE NILAI AMBANG LOKAL DAN GLOBAL UNTUK RESTORASI DOKUMEN JAWI KUNO

Khairun Saddami¹, Fitri Arnia^{*2}, Yuwaldi Away³, Khairul Munadi⁴

^{1,2,3,4}Universitas Syiah Kuala

Email: ¹khairun91@mhs.unsyiah.ac.id, ²f.arnia@unsyiah.ac.id, ³yuwaldi@unsyiah.ac.id,

⁴khairul.munadi@unsyiah.ac.id

^{*}Penulis Korespondensi

(Naskah masuk: 01 Februari 2019, diterima untuk diterbitkan: 14 Januari 2020)

Abstrak

Dokumen Jawi kuno merupakan warisan budaya yang berisi informasi penting tentang peradaban masa lalu yang dapat dijadikan pedoman untuk masa sekarang ini. Dokumen Jawi kuno telah mengalami penurunan kualitas yang disebabkan oleh beberapa faktor seperti kualitas kertas atau karena proses penyimpanan. Penurunan kualitas ini menyebabkan informasi yang terdapat pada dokumen tersebut menghilang dan sulit untuk diakses. Artikel ini mengusulkan metode binerisasi untuk membangkitkan kembali informasi yang terdapat pada dokumen Jawi kuno. Metode usulan merupakan kombinasi antara metode binerisasi berbasis nilai ambang lokal dan global. Metode usulan diuji terhadap dokumen Jawi kuno dan dokumen uji standar yang dikenal dengan nama *Handwritten Document Image Binarization Contest* (HDIBCO) 2016. Citra hasil binerisasi dievaluasi menggunakan metode: *F-measure*, *pseudo F-measure*, *peak signal-to-noise ratio*, *distance reciprocal distortion*, dan *misclassification penalty metric*. Secara rata-rata, nilai evaluasi *F-measure* dari metode usulan mencapai 88,18 dan 89,04 masing-masing untuk dataset Jawi dan HDIBCO-2016. Hasil ini lebih baik dari metode pembandingan yang menunjukkan bahwa metode usulan berhasil meningkatkan kinerja metode binerisasi untuk dataset Jawi dan HDIBCO-2016.

Kata kunci: binerisasi dokumen kuno, kombinasi metode global dan lokal, dokumen Jawi kuno, binerisasi berbasis nilai ambang

COMBINATION LOCAL AND GLOBAL THRESHOLDING METHOD FOR BINARIZING ANCIENT JAWI DOCUMENT

Abstract

Ancient Jawi document is a cultural heritage, which contains knowledge of past civilization for developing a better future. Ancient Jawi document suffers from severe degradation due to some factors such as paper quality or poor retention process. The degradation reduces information on the document and thus the information is difficult to access. This paper proposed a binarization method for restoring the information from degraded ancient Jawi document. The proposed method combined a local and global thresholding method for extracting the text from the background. The experiment was conducted on ancient Jawi document and Handwritten Document Image Binarization Contest (HDIBCO) 2016 datasets. The result was evaluated using F-measure, pseudo F-measure, peak signal-to-noise ratio, distance reciprocal distortion, dan misclassification penalty metric. The average result showed that the proposed method achieved 88.18 and 89.04 of F-measure, for Jawi and HDIBCO-2016, respectively. The proposed method resulted in better performance compared with several benchmarking methods. It can be concluded that the proposed method succeeded to enhance binarization performance.

Keywords: document image binarization, combination of local and global thresholding, ancient Jawi document

1. PENDAHULUAN

Dokumen Jawi kuno merupakan warisan nenek moyang di Nusantara yang berisi informasi penting seperti obat-obatan herbal, hukum-hukum masyarakat dan lain-lain. Nusantara adalah daerah yang sekarang terdiri dari beberapa negara berupa

Indonesia, Malaysia, Singapura, Brunei Darussalam, Thailand Selatan, Filipina Selatan dan Myanmar Selatan (Yahaya, 2016).

Dokumen Jawi kuno ditulis menggunakan aksara Arab dengan bahasa yang digunakan adalah bahasa Melayu (Gallop et al., 2015). Dokumen Jawi

kuno tersebut mengalami degradasi kualitas seperti terkena rembesan air, munculnya tulisan dari halaman belakang dokumen yang disebabkan oleh tinta yang banjir serta munculnya bercak-bercak kuning disebabkan oleh kualitas kertas. Penurunan kualitas ini dapat menyebabkan informasi pada dokumen tersebut memudar atau bahkan hilang (Arnia and Munadi, 2017). Gambar 1 menunjukkan contoh dokumen Jawi kuno yang telah mengalami penurunan kualitas.

Salah satu cara untuk membangkitkan kembali informasi dari dokumen kuno adalah dengan melakukan proses analisis terhadap dokumen tersebut. Tahapan awal dari metode analisis dokumen untuk membangkitkan kembali informasi yang ada pada suatu dokumen adalah tahapan binerisasi. Binerisasi adalah metode yang mengubah citra dokumen kuno menjadi citra dalam dua representasi warna yaitu putih sebagai latar dan hitam sebagai teks.

Menurut Wen, ada tiga jenis metode binerisasi yaitu metode berbasis penentuan nilai ambang (*thresholding*), metode binerisasi berbasis kluster, dan metode gabungan (Wen, Li and Sun, 2013). Metode binerisasi berbasis nilai ambang merupakan metode binerisasi yang sederhana dan cepat, sehingga sangat mudah untuk diimplementasikan secara praktis. Metode binerisasi berbasis nilai ambang menghitung suatu nilai ambang sebagai pembatas dalam memisahkan piksel citra menjadi hitam atau putih.

Ada dua jenis metode binerisasi berbasis ambang, yaitu binerisasi berbasis nilai ambang global dan berbasis nilai lokal. Metode berbasis nilai ambang global adalah metode yang mencari suatu nilai ambang untuk keseluruhan nilai piksel pada citra. Sedangkan metode berbasis nilai ambang lokal mencari nilai ambang untuk piksel tertentu berdasarkan informasi dari tetangga piksel tersebut pada area tertentu.

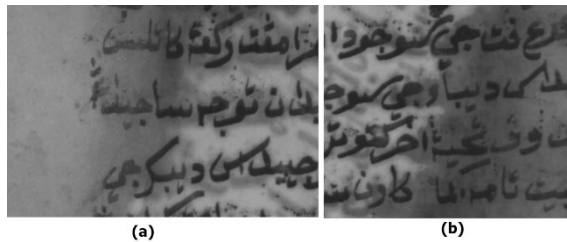
Metode Otsu merupakan metode binerisasi yang berbasis nilai ambang yang sangat populer. Metode Otsu adalah metode penentuan nilai ambang secara global (Otsu, 1979). Metode Otsu dirumuskan sebagai berikut.

$$\sigma_w^2(t) = \argmin [w_0(t) \sigma_0^2(t) + w_1(t) \sigma_1^2(t)] \quad (1)$$

dimana w_0 dan w_1 masing-masing adalah probabilitas dua kelas yang dipisahkan oleh nilai ambang t , dan σ_0^2 dan σ_1^2 adalah varian dari masing-masing kelas.

Namun metode Otsu mempunyai beberapa kekurangan diantaranya adalah metode ini gagal jika degradasi yang terdapat pada suatu dokumen tidak seragam. Oleh karena itu, pada tahun 1985, Niblack memperkenalkan konsep binerisasi berbasis nilai ambang lokal (Saxena, 2017). Metode Niblack dirumuskan sebagai berikut.

$$T_n = m + k\sigma \quad (2)$$



Gambar 1. Dokumen Jawi kuno yang mengalami degradasi yang sangat parah

dimana T_n adalah nilai ambang Niblack, m adalah nilai rata-rata dalam suatu lokal area yang disebut dengan *window*, σ adalah nilai standar deviasi dari piksel yang ada di dalam *window*, dan k adalah faktor yang mempunyai nilai -0,2. Kelemahan metode Niblack adalah citra hasil binerisasi membangkitkan banyak bintik-bintik hitam yang bukan teks, namun metode ini menjadi dasar dari binerisasi berbasis nilai ambang lokal.

Sauvola mengusulkan metode perbaikan dari Niblack yang dirumuskan sebagai persamaan (3) berikut (Sauvola and Pietikäinen, 2000).

$$T_s = m \left[1 + k \left(\frac{\sigma}{R} - 1 \right) \right] \quad (3)$$

dimana T_s adalah nilai ambang Sauvola, R adalah nilai maksimum dari standar deviasi pada setiap *window*, dan nilai k adalah 0,5. Metode Sauvola mempunyai kinerja yang lebih baik dari metode Niblack, namun metode ini gagal jika nilai piksel teks dan latar pada dokumen berdekatan. Oleh karena itu, Wolf mengusulkan metode nilai ambang baru yang merupakan modifikasi dari metode Sauvola. Metode Wolf dirumuskan sebagai (Wolf and Jolion, 2004).

$$T_w = (1 - k)m + kM + k \frac{\sigma}{R} (m - M) \quad (4)$$

dimana T_w adalah nilai ambang Wolf, M adalah nilai minimum yang terdapat dalam suatu *window* dan nilai k adalah 0,5.

Beberapa peneliti telah melakukan kajian binerisasi untuk dokumen Jawi kuno, seperti Bataineh (Bataineh, Abdullah and Omar, 2015), Razak (Rzak, 2016), Muchallil (Muchallil et al., 2015), iNICK (Saddami et al., 2017), dan Local Otsu (Saddami et al., 2019). Hasil kajian menunjukkan bahwa tidak ada metode yang mempunyai performansi terbaik untuk dokumen Jawi kuno. Walaupun iNICK menunjukkan performansi yang memuaskan, namun masih dibutuhkan peningkatan kinerja binerisasi berbasis nilai ambang untuk dokumen kuno.

Hasil penelitian oleh (Saddami et al., 2017, 2019) menunjukkan bahwa metode Otsu dan Niblack mengekstraksi banyak piksel hitam yang bukan teks pada binerisasi dokumen Jawi kuno. Sedangkan metode Wolf gagal dalam mengekstraksi piksel yang seharusnya menjadi teks. Oleh karena itu, permasalahan yang dialami oleh ketiga metode ini menjadi perhatian dalam membangun suatu kombinasi metode binerisasi baru berbasis nilai

ambang. Beberapa penelitian telah dilakukan dengan cara mengkombinasikan metode binerisasi lokal dan global seperti (Sari, Kefali and Bahi, 2014; Lech, 2015), namun hanya berhasil pada dokumen tertentu saja.

Artikel ini mengusulkan metode binerisasi baru dengan mengkombinasikan antara nilai ambang lokal dan nilai ambang global. Metode ini menggabungkan metode Niblack, Wolf dan Otsu dengan proses penurunan persamaan matematis untuk meningkatkan hasil binerisasi dokumen Jawi kuno. Walaupun beberapa metode binerisasi telah dikembangkan, namun untuk implementasi secara praktis, metode binerisasi berbasis nilai ambang tetap menjadi perhatian karena kecepatan proses dan kemudahan dalam implementasi.

2. METODE PENELITIAN

Penelitian ini dilakukan untuk mengembangkan persamaan baru berdasarkan penurunan persamaan matematis dari metode Niblack dan Wolf yang merupakan metode binerisasi berbasis nilai ambang lokal dan metode Otsu yang merupakan metode binerisasi berbasis nilai ambang global. Selanjutnya, dilakukan simulasi komputasi berdasarkan persamaan hasil penggabungan tersebut. Simulasi yang dimaksud adalah membinaerisasi dokumen-dokumen kuno baik dokumen Jawi maupun non-Jawi, kemudian dihitung nilai evaluasi menggunakan metode evaluasi dari kompetisi *Document Image Binarization Contest* (DIBCO).

2.1 Metode usulan

Metode usulan merupakan gabungan dari metode binerisasi lokal Niblack dan Wolf, dengan metode global Otsu. Tahapan pertama dari penelitian ini adalah kombinasi antara metode Niblack dan Wolf.

$$T_b = \frac{T_{niblack} + T_{wolf}}{2} \quad (5)$$

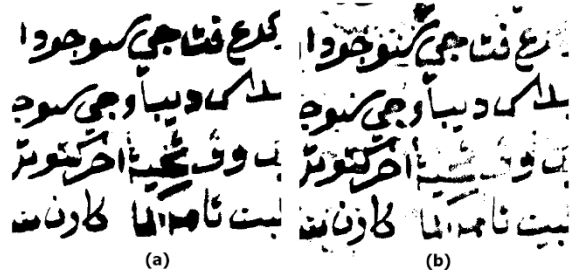
dimana $T_{niblack}$ adalah nilai ambang Niblack dan T_{wolf} adalah nilai ambang Wolf.

$$T_b = \frac{m + k_{niblack}\sigma + \{(1 - k_{wolf})m + Mk_{wolf} + k_{wolf}R\sigma(m - M)\}}{2} \quad (6)$$

dimana m adalah nilai rata-rata dalam suatu jendela lokal (*window*), $k_{niblack}$ adalah nilai k pada metode Niblack yang bernilai -0,2, σ adalah standar deviasi dari jendela lokal, M adalah nilai minimum dari citra, k_{wolf} adalah nilai k pada metode wolf yang bernilai 0,5, dan R adalah nilai maksimum dari standar deviasi dari citra masukan yang bernilai 128. Oleh karena itu, formula gabungan Niblack dan Wolf dirumuskan sebagai berikut.

$$T_b = c_1m + c_3m\sigma + c_3\sigma \quad (7)$$

dimana m adalah rata-rata *window*, σ adalah standar deviasi citra. Berdasarkan hasil pengujian yang berulang-ulang, nilai yang tepat untuk c_1 , c_2 , dan c_3 adalah $c_1 = \frac{3}{4}$, $c_2 = \frac{1}{256}$, dan $c_3 = \frac{1}{20}$.

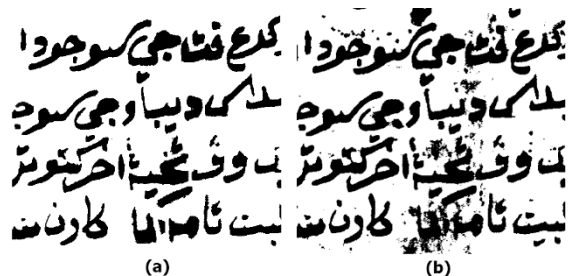


Gambar 2. Citra biner dari dokumen Jawi kuno pada Gambar 1b: (a) Citra Ground-Truth (GT), (b) Hasil binerisasi dokumen pada Gambar 1b menggunakan persamaan (7)

Metode penentuan nilai ambang pada persamaan (7) tersebut diuji terhadap dokumen Jawi kuno. Gambar 2 menunjukkan hasil binerisasi terhadap dokumen Jawi kuno yang mengalami penurunan kualitas disebabkan oleh tumpahan air, tinta yang memudar, dan pencahayaan yang tidak merata. Untuk meningkatkan kinerja metode usulan terhadap dokumen seperti ini, maka metode usulan diperbarui dengan mengkombinasikan metode usulan dengan metode penentuan nilai ambang Otsu dan standar deviasi.

$$T_{clg} = \frac{T_b + T_{otsu} - SD}{2} \quad (8)$$

dimana T_{clg} adalah nilai ambang baru hasil penggabungan metode binerisasi berbasis nilai ambang lokal dan global, T_b adalah nilai ambang pada persamaan (7), T_{otsu} adalah nilai ambang Otsu, dan $SD = \frac{\sigma}{2}$. Setelah dikombinasikan metode lokal dan global, maka hasil binerisasi menjadi lebih baik seperti ditunjukkan gambar 3 yang dibinaerisasi menggunakan persamaan (8).

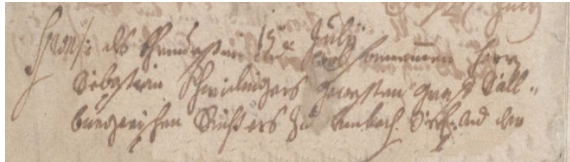


Gambar 3. Citra biner dari dokumen Jawi kuno pada Gambar 1b: (a) Citra Ground-Truth (GT), (b) Hasil binerisasi dokumen pada Gambar 1b menggunakan persamaan (8).

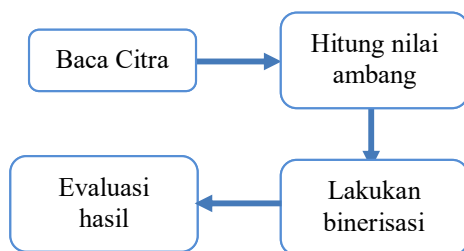
2.2 Simulasi

Simulasi dilakukan terhadap dokumen Jawi kuno dan dokumen non-Jawi yang berupa citra *Handwritten Document Image Binarization Contest*

(HDIBCO) 2016 (Pratikakis et al., 2016). Dokumen HDIBCO 2016 adalah dokumen kuno yang bersifat terbuka yang digunakan pada kompetisi binerisasi tingkat internasional. Gambar 4 menunjukkan contoh dokumen HDIBCO 2016.



Gambar 4. Contoh dokumen HDIBCO 2016



Gambar 5. Proses tahapan simulasi

Tahapan simulasi dari proses binerisasi menggunakan metode usulan ditunjukkan pada Gambar 5. Tahapan pertama adalah proses membaca citra. Kemudian dilanjutkan dengan menghitung nilai ambang untuk tahapan binerisasi. Selanjutnya melakukan proses binerisasi menggunakan nilai ambang yang telah didapatkan pada langkah sebelumnya. Tahapan terakhir dari simulasi adalah melakukan evaluasi citra terbinerisasi.

Selain metode usulan, simulasi juga dilakukan terhadap beberapa metode binerisasi berbasis nilai ambang (*thresholding*) lainnya berupa Otsu, Niblack, Sauvola, Wolf, dan NICK (Khurshid et al., 2009) sebagai pembandingan untuk mengetahui kinerja dari metode usulan.

Dalam membangun simulasi dari metode binerisasi Niblack, Sauvola, Wolf dan NICK, beberapa parameter harus diatur secara manual. Ukuran *window* yang digunakan pada metode Niblack, Sauvola, dan Wolf adalah 15x15. Nilai k untuk metode Niblack, Sauvola dan Wolf masing-masing adalah -0,2 0,5, dan 0,5 (Sezgin and Sankur, 2004). Sedangkan untuk metode NICK, ukuran *window* yang digunakan adalah 19x19 dan $k = -0,2$, sesuai dengan yang disarankan oleh pengarang.

Parameter evaluasi yang digunakan pada penelitian ini adalah F-measure (FM), pseudo F-measure (FMps), peak signal-to-noise ratio (PSNR), misclassification penalty metric (MPM) (Gatos, Ntirogiannis and Pratikakis, 2011) dan distance reciprocal distortion (DRD) (Lu, Kot and Shi, 2004). Semakin tinggi nilai evaluasi FM, FMps, dan PSNR, maka semakin bagus suatu metode binerisasi. Hal sebaliknya berlaku pada metode DRD dan MPM,

semakin rendah nilai evaluasi DRD dan MPM, maka semakin bagus metode tersebut.

3. HASIL DAN PEMBAHASAN

3.1 Analisis terhadap kualitas citra terbinerisasi

Pada bagian ini dipaparkan hasil dan pembahasan terhadap kualitas citra terbinerisasi. Gambar 5 dan 6 menunjukkan hasil binerisasi terhadap dokumen Jawi kuno yang terdapat pada Gambar 1. Selanjutnya, Gambar 7 menunjukkan contoh hasil binerisasi citra dari *dataset* HDIBCO 2016.

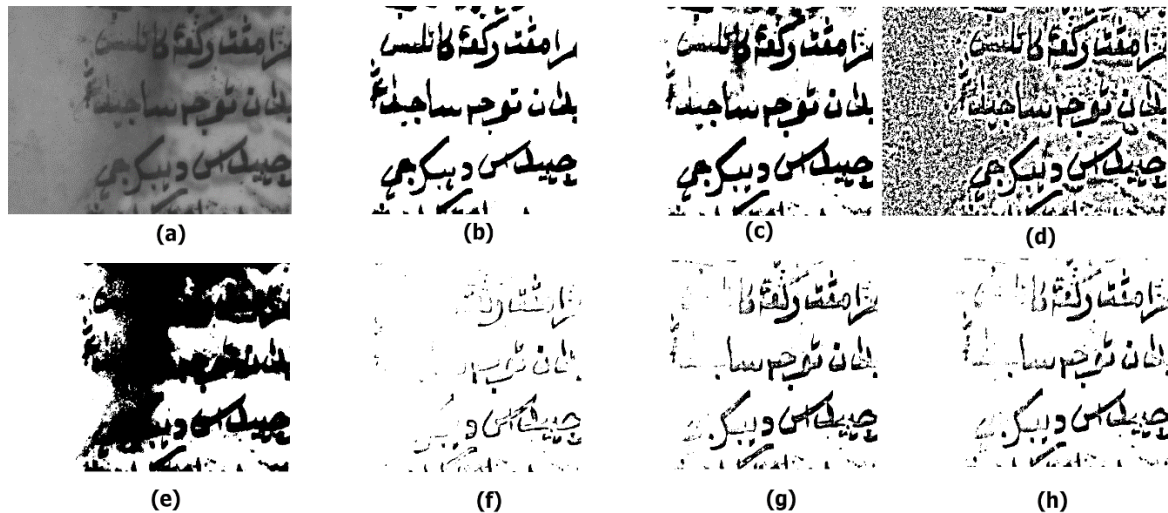
Gambar 5 menunjukkan perbandingan hasil binerisasi pada dokumen Jawi kuno. Gambar 5b merupakan citra ground-truth (GT) yang merupakan citra referensi sebagai hasil binerisasi yang terbaik. Gambar 5c, yang merupakan hasil binerisasi menggunakan metode usulan, menggambarkan hasil yang paling mirip dengan citra GT. Gambar 5d merupakan hasil binerisasi menggunakan metode Niblack. Gangguan (*noise*) yang dihasilkan oleh metode Niblack membuat teks sulit dikenali. Sedangkan Gambar 5e merupakan hasil binerisasi menggunakan metode Otsu. Kedua metode tersebut menghasilkan citra biner yang mempunyai banyak gangguan atau *noise* disekitar teks. Gambar 5f, 5g, dan 5h merupakan hasil binerisasi menggunakan metode Sauvola, Wolf dan NICK. Ketiga metode ini gagal membangkitkan sebagian besar teks, terutama metode Sauvola. Metode Wolf dan NICK hanya berhasil membangkitkan sebagian teks yang ada pada citra masukan, sedangkan sisanya gagal dibangkitkan sebagai teks. Hasil ini menunjukkan bahwa metode usulan mempunyai performansi yang bagus dalam membangkitkan teks dengan tepat.

Gambar 6 mempunyai kualitas yang hampir sama dengan Gambar 5. Hasil menunjukkan bahwa Gambar 6c yang merupakan hasil binerisasi menggunakan metode usulan mempunyai kinerja yang baik dalam membangkitkan citra terbinerisasi. Gambar 6d, dan 6e merupakan hasil binerisasi metode Otsu dan Niblack. Kedua metode ini membangkitkan banyak *noise* disekitar teks bahkan menutupi teks seperti yang diperlihatkan pada Gambar 6e. *Noise* disekitar teks ini membuat informasi berupa teks menjadi kabur dan sulit dikenali. Sebaliknya, Gambar 6f, 6g, dan 6h gagal membangkitkan sebagian besar informasi berupa teks, terutama pada Gambar 6f. Gambar 6f, 6g, dan 6h masing-masing merupakan hasil binerisasi menggunakan metode Sauvola, Wolf, dan NICK. Gambar 6f hanya membangkitkan sedikit teks dari yang seharusnya dibangkitkan. Metode Wolf dan NICK mempunyai kinerja yang sedikit lebih baik dari metode Sauvola.

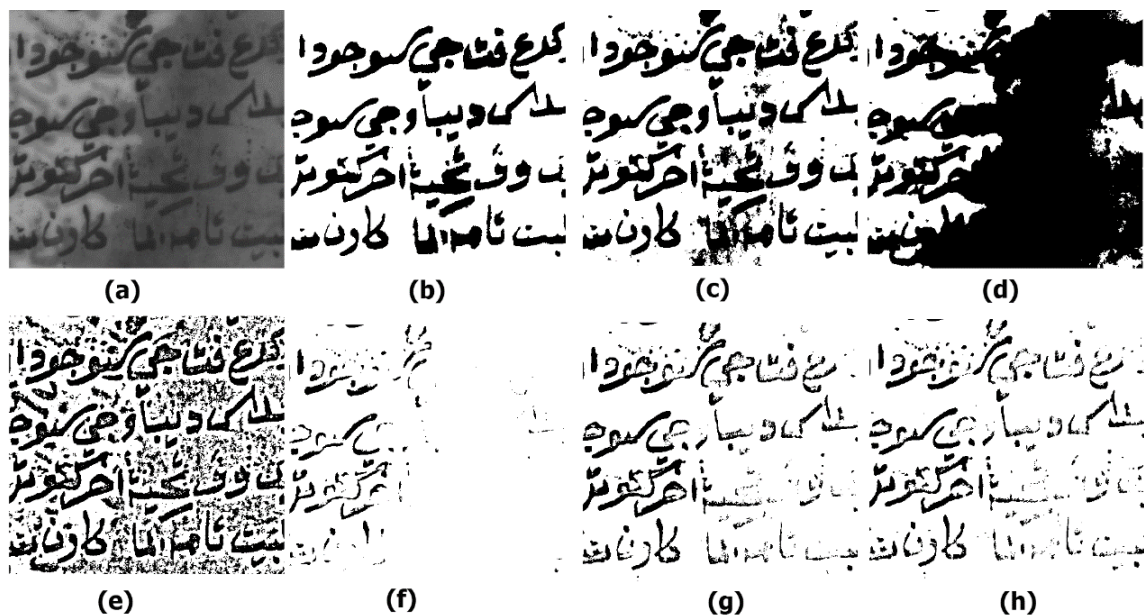
Gambar 7 menunjukkan perbandingan hasil binerisasi terhadap dokumen HDIBCO 2016. Gambar 7c menunjukkan hasil binerisasi dengan menggunakan metode usulan. Gambar 7d dan 7e

merupakan citra hasil binerisasi menggunakan metode Otsu dan Niblack. Metode Otsu dan Niblack menghasilkan banyak piksel hitam yang mengganggu teks yang asli bahkan menutupi teks yang asli. Sedangkan Gambar 7f, 7g, dan 7h merupakan hasil binerisasi menggunakan metode Sauvola, Wolf, dan

NICK. Hasil binerisasi menggunakan metode Sauvola, Wolf, dan NICK gagal membangkitkan piksel yang berupa teks, terutama metode Sauvola dan NICK. Metode Wolf menghasilkan citra terbinerisasi yang paling mendekati metode usulan.



Gambar 5. Perbandingan hasil binerisasi terhadap dokumen Jawi pada Gambar 1a: (a) Citra dokumen Jawi kuno, (b) Citra GT, (c) Hasil binerisasi metode usulan, (d) Hasil binerisasi metode Niblack, (e) Hasil binerisasi metode Otsu, (f) Hasil binerisasi metode Sauvola, (g) Hasil binerisasi metode Wolf, (h) Hasil binerisasi metode NICK.



Gambar 6. Perbandingan hasil binerisasi terhadap dokumen Jawi pada Gambar 1b: (a) Citra dokumen Jawi kuno, (b) Citra GT, (c) Hasil binerisasi metode usulan, (d) Hasil binerisasi metode Niblack, (e) Hasil binerisasi metode Otsu, (f) Hasil binerisasi metode Sauvola, (g) Hasil binerisasi metode Wolf, (h) Hasil binerisasi metode NICK.

Berdasarkan contoh dari Gambar 5, 6, dan 7, dapat dilihat bahwa metode Niblack dan Otsu membangkitkan banyak piksel berupa teks dan juga noise, sedangkan pada metode Wolf banyak piksel

berupa teks atau informasi yang gagal dibangkitkan. Penggabungan tiga metode tersebut menunjukkan hasil yang memuaskan terhadap binerisasi dokumen kuno terutama dokumen Jawi kuno.



Gambar 7. Perbandingan hasil binerisasi terhadap dokumen Jawi pada Gambar 4: (a) Citra dokumen HDIBCO 2016, (b) Citra GT, (c) Hasil binerisasi metode usulan, (d) Hasil binerisasi metode Otsu, (e) Hasil binerisasi metode Niblack, (f) Hasil binerisasi metode Sauvola, (g) Hasil binerisasi metode Wolf, (h) Hasil binerisasi metode NICK.

3.2 Analisis terhadap parameter evaluasi

Tabel 1 dan 2 menunjukkan perbandingan hasil evaluasi terhadap dataset dokumen Jawi kuno dan HDIBCO 2016. Tabel 1 menunjukkan perbandingan hasil binerisasi metode usulan dengan metode Otsu, Niblack, Sauvola, Wolf, dan NICK terhadap dokumen Jawi kuno. Hasil evaluasi terbaik ditandai dengan tulisan hitam tebal. Berdasarkan Tabel 1, metode usulan memiliki nilai terbaik untuk metode evaluasi FM, FMps, PSNR, DRD, dan MPM dengan nilai masing-masing adalah 88,18, 92,29, 15,81, 5,52, dan 6,13. Metode yang paling mendekati hasil metode usulan adalah metode Wolf. Metode Wolf memiliki nilai 80,16 untuk FM, 83,56 untuk FMps, 13,63 untuk PSNR, 9,25 untuk DRD, dan 7,45 untuk MPM. Metode yang mendapatkan hasil paling rendah adalah metode Niblack.

Berdasarkan Tabel 1, terlihat jelas bahwa metode usulan lebih unggul dibandingkan dengan metode berbasis ambang lainnya. Metode usulan berhasil membangkitkan teks dengan baik serta tidak banyak menghasilkan piksel-piksel yang bukan teks. Hal ini ditandai dengan rendahnya nilai DRD dan MPM. Nilai FM, FMps, dan PSNR menunjukkan bahwa metode usulan berhasil membangkitkan teks dengan baik dibandingkan metode pembandingan lainnya. Sedangkan nilai DRD dan MPM menunjukkan bahwa metode usulan sangat baik dalam menemukan piksel teks atau bukan teks.

Tabel 2 memaparkan perbandingan hasil binerisasi metode usulan dengan metode Otsu, Niblack, Sauvola, Wolf, dan NICK, serta dengan masing-masing juara 1, 2 dan 3 pada kompetisi binerisasi HDIBCO 2016 terhadap dokumen HDIBCO 2016. Berdasarkan Tabel 2, metode usulan memperoleh nilai 89,04, 92,51, 18,85, 3,90, dan 2,18 untuk metode evaluasi FM, FMps, PSNR, DRD, dan

MPM secara berturut-turut. Hasil yang paling mendekati dengan metode usulan adalah juara 2 dari kompetisi tersebut. Metode Niblack memperoleh nilai paling rendah dari keseluruhan metode pembandingan.

Tabel 1. Perbandingan hasil binerisasi pada dokumen Jawi kuno dalam nilai rata-rata

Metode	FM	FMps	PSNR	DRD	MPM
Otsu	71,51	71,56	11,04	25,30	42,68
Niblack	61,58	62,79	7,20	30,77	155,49
Sauvola	62,55	64,44	12,25	14,76	10,30
Wolf	80,16	83,56	13,63	9,25	7,45
NICK	71,86	79,84	12,33	11,35	8,80
Usulan	88,18	92,29	15,81	5,52	6,13

Tabel 2. Perbandingan hasil binerisasi pada dokumen HDIBCO 2016 dalam nilai rata-rata

Metode	FM	FMps	PSNR	DRD	MPM
Otsu	86,59	89,92	17,79	5,58	5,52
Niblack	34,07	34,77	5,77	133,57	173,50
Sauvola	74,98	82,21	16,57	7,33	1,30
Wolf	80,21	88,42	16,89	6,22	1,21
NICK	75,93	85,14	16,24	7,56	2,67
Juara 3	88,47	91,71	18,29	3,93	-
Juara 2	88,72	91,84	18,45	3,86	-
Juara 1	87,61	91,28	18,11	5,21	-
Usulan	89,04	92,51	18,85	3,90	2,18

4. KESIMPULAN

Penelitian ini mengusulkan metode binerisasi berbasis nilai ambang yang merupakan gabungan metode nilai ambang lokal dan global. Metode yang diusulkan merupakan kombinasi metode Niblack, Wolf, dan Otsu. Metode usulan diuji terhadap dokumen Jawi kuno dan dataset HDIBCO 2016 yang digunakan pada kompetisi binerisasi tahun 2016. Hasil menunjukkan bahwa metode usulan mempunyai kinerja yang lebih baik dari metode Otsu, Niblack, Sauvola, Wolf dan NICK. Bahkan hasil metode usulan lebih baik dari juara pada kompetisi binerisasi tahun 2016. Penelitian ke depan akan

mengintegrasikan metode usulan dengan metode binerisasi berbasis *framework*.

PENGHARGAAN

Penelitian ini dibiayai oleh Kementerian RistekDikti dibawah skema PMDSU.

DAFTAR PUSTAKA

ARNIA, F. AND MUNADI, K., 2017. Binarization of Ancient Document Images based on Multipeak Histogram Assumption. *Telkomnika*, 15(3).

BATAINEH, B., ABDULLAH, S.N.H.S. AND OMAR, K., 2015. Adaptive binarization method for degraded document images based on surface contrast variation. *Pattern Analysis and Applications*, pp.1–14.

GALLOP, A.T., MAMAT, W.A.W., AKBAR, A., BRAGINSKY, V., TENGAH, A.H.B. BIN A.H., CALDWELL, I., CHAMBERT-LOIR, H., CORDELL, H., DENISOVA, T.A., YAHYA, F. and others, 2015. A Jawi sourcebook for the study of Malay palaeography and orthography. *Indonesia and the Malay World*, 43(125), pp.13–171.

GATOS, B., NTIROGIANNIS, K. AND PRATIKAKIS, I., 2011. DIBCO 2009: Document image binarization contest. *International Journal on Document Analysis and Recognition*, 14(1), pp.35–44.

KHURSHID, K., SIDDIQI, I., FAURE, C. AND VINCENT, N., 2009. Comparison of Niblack inspired Binarization methods for ancient documents. In: *IS&T/SPIE Electronic Imaging*. p.72470U--72470U.

LECH, P., 2015. Binarization of document images using the modified local-global Otsu and Kapur algorithms. *Przegląd Elektrotechniczny*, [online] 1(2), pp.73–76. Available at: <<http://sigma-not.pl/publikacja-89353-2015-2.html>>.

LU, H., KOT, A.C. AND SHI, Y.Q., 2004. Distance-reciprocal distortion measure for binary document images. *IEEE Signal Processing Letters*, 11(2), pp.228–231.

MUCHALLIL, S., ARNIA, F., MUNADI, K. and FARDIAN, 2015. Performance comparison of denoising methods for historical documents. *Jurnal Teknologi*, 77(22), pp.137–143.

OTSU, N., 1979. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, [online] 9(1), pp.62–66. Available at: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4310076>>.

PRATIKAKIS, I., ZAGORIS, K., BARLAS, G. AND GATOS, B., 2016. ICFHR2016 Handwritten Document Image Binarization Contest (H-DIBCO 2016). In: *Frontiers in Handwriting Recognition*

(ICFHR), 2016 15th International Conference on. pp.619–623.

RZAK, Z., 2016. *Old Jawi manuscript: digital recognition*. University of Malaya.

SADDAMI, K., MUNADI, K., AWAY, Y. AND ARNIA, F., 2019. Improvement of Binarization Performance using Local Otsu Thresholding. *International Journal of Electrical and Computer Engineering*, 9(1), pp.264–272.

SADDAMI, K., MUNADI, K., MUCHALLIL, S. AND ARNIA, F., 2017. Improved Thresholding Method for Enhancing Jawi Binarization Performance. In: *Document Analysis and Recognition, 2017. ICDAR'17. 14th International Conference on*. pp.1108–1113.

SARI, T., KEFALI, A. AND BAHI, H., 2014. Text extraction from historical document images by the combination of several thresholding techniques. *Advances in Multimedia*, 2014.

SAUVOLA, J. AND PIETIKÄINEN, M., 2000. Adaptive document image binarization. *Pattern recognition*, 33(2), pp.225–236.

SAXENA, L.P., 2017. Niblack's binarization method and its modifications to real-time applications: a review. *Artificial Intelligence Review*, pp.1–33.

SEZGIN, M. AND SANKUR, B., 2004. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1), pp.146–165.

WEN, J., LI, S. AND SUN, J., 2013. A new binarization method for non-uniform illuminated document images. *Pattern Recognition*, [online] 46(6), pp.1670–1690. Available at: <<http://dx.doi.org/10.1016/j.patcog.2012.11.027>>.

WOLF, C. AND JOLION, J.-M., 2004. Extraction and recognition of artificial text in multimedia documents. *Pattern Analysis & Applications*, 6(4), pp.309–326.

YAHAYA, D.M.H., 2016. The Jawi Manuscript: Its HISTORY, ROLE, AND FUNCTION IN THE MALAY ARCHIPELAGO.2016. *Journal of Islamic Studies and Culture*, [online] 4(1), pp.52–61. Available at: <<http://jiscnet.com/vol-4-no-1-june-2016-abstract-7-jisc>>.

Halaman ini sengaja dikosongkan