Optimizing computer vision through adaptive thresholding-based image segmentation

M. Riyaz Mohammed[*1], Dr. S. Abdul Saleem[2], Dr. K. Syed Kousar Niasi[3]

*[*1]Assistant Professor, Department of Computer Science, Jamal Mohamed College, (Affiliated to Bharathidasan University), Tiruchirappalli-20, Tamil Nadu, India*

[2,3]Associate Professor, Department of Computer Science, Jamal Mohamed College, (Affiliated to Bharathidasan University), Tiruchirappalli-20, Tamil Nadu, India

DOI: 10.5281/zenodo.23053813

Abstract:
Image segmentation separates foreground structures from background so that later analysis can use a simpler representation. A single global threshold is efficient, but uneven lighting and low contrast can make one cutoff unsuitable across an image. This focused review examines adaptive thresholding, which estimates a threshold from a local neighborhood for each pixel. It compares mean- and Gaussian-based approaches with variance-aware and integral-image variants; describes preprocessing, window and offset selection, and post-processing; and sets out an evaluation protocol for document, biomedical, and inspection images. The source paper did not supply a dataset, ground-truth masks, measured results, or a reproducible experiment. Accordingly, this review reports methodological comparisons and expected trade-offs rather than claiming numerical performance. Adaptive thresholding is most useful when foreground/background separation is primarily an intensity problem with spatially varying illumination; it is less reliable when texture, noise, and foreground intensities overlap locally.

Keywords:
Adaptive thresholding, image segmentation, binarization, document image binarization, local threshold estimation, Otsu thresholding, Niblack method, Sauvola method, integral image, uneven illumination, degraded documents, histogram analysis, morphological post-processing, foreground extraction, evaluation metrics, computer vision.

1. Introduction:

Thresholding assigns pixels to foreground or background by comparing intensity with a cutoff. Otsu's global method selects one cutoff from an image histogram [1]. It is a useful baseline, but one value cannot reflect every region when a page has a shadow, a specimen has uneven illumination, or a surface has a bright reflection. Adaptive thresholding instead computes a local threshold T(x,y) from neighboring pixels [2–4].

This paper concentrates on that local decision: how the neighborhood is prepared, how a threshold is estimated, and how the resulting mask should be assessed. It is a review and method outline, not a report of a newly run experiment. Figure 1 maps the processing sequence; the results section makes qualitative, conditional comparisons only. Related surveys place adaptive thresholding among a wider family of global and local segmentation techniques [10,20].

2. Literature survey:

Otsu's method is a reference point for global threshold selection: it maximizes separation between two histogram classes [1]. Where illumination is uniform and foreground and background are well separated, this single cutoff is simple and fast. Spatially changing illumination violates that assumption, motivating local statistics.

Niblack estimates a threshold from neighborhood mean and standard deviation [2]. The variance term helps respond to local contrast, but noisy backgrounds can also generate spurious foreground. Sauvola and Pietikäinen adapt the variance-aware approach for document binarization, reducing the tendency to mark

relatively uniform background as text [3]. These methods require choices of window size and sensitivity parameters, so performance depends on image scale and background variation.

Bradley and Roth use an integral image to obtain local sums efficiently [4]. The integral-image method is valuable when many overlapping windows must be evaluated or low-latency processing is needed. Mean and Gaussian adaptive thresholds are widely used practical baselines: the former treats window pixels equally, while the latter emphasizes the center of the window. Neither weighting scheme should be called universally superior without measurements on the target images.

Contrast-limited adaptive histogram equalization (CLAHE) addresses local contrast rather than performing segmentation itself [5]. Used cautiously before thresholding, it can expose faint structures; excessive enhancement can strengthen noise. The methods share a common limitation: they act on grayscale intensity and local context, not semantic meaning. This review therefore treats enhancement, local estimation, and cleanup as distinct stages, and does not equate a visually clean binary image with validated segmentation quality.

2.1. Global and histogram-based thresholding:

Beyond Otsu's between-class variance criterion [1], several global selectors recur in the thresholding literature. Kapur, Sahoo, and Wong choose the cutoff that maximizes the entropy of the resulting foreground and background histograms [8], while Kittler and Illingworth model the histogram as a mixture of two classes and select a minimum-error threshold [7]. Surveys by Sezgin and Sankur [10] and by Pal and Pal [20] compare these global selectors across many image categories and find that no single criterion dominates; performance tracks how closely the underlying histogram model matches the image. These methods remain useful baselines because they are cheap and parameter-light, but all of them still produce one cutoff for the whole frame, which is exactly the assumption that uneven illumination breaks.

2.2. Local and document-oriented binarization:

Local methods predate many of the global refinements. Bernsen computes a threshold from the midpoint of local minimum and maximum intensities within a window [6], and Wellner's DigitalDesk work applies an adaptive local average to camera images of paper lying on a desk, where lighting cannot be controlled [9]. For degraded documents, Gatos, Pratikakis, and Perantonis combine background-surface estimation with a Sauvola-style local decision [11]; Su, Lu, and Tan construct a contrast image so that local maxima separate text stroke edges from background [14]; and Howe estimates the binarization parameters automatically from image statistics so that tuning does not rely on hand-chosen constants [15]. Bataineh, Abdullah, and Omar address the fixed-window weakness directly by selecting dynamic window sizes for each region [18]. Each of these methods keeps the same core structure as Niblack and Sauvola [2,3]: a local estimate, one or two sensitivity parameters, and a per-pixel comparison rule.

2.3. Benchmarks, efficiency, and application studies:

Public benchmark contests have made document binarization results comparable across research groups: the DIBCO series, beginning with the ICDAR 2009 contest [17], and the evaluation methodology of Ntirogiannis, Gatos, and Pratikakis [16] define matched test sets and measures, so that a reported score can be interpreted against shared ground truth. On the efficiency side, Shafait, Keysers, and Breuel show that Sauvola-quality binarization can be computed with integral images in time close to a global threshold, independent of window size [12]; this extends the engineering argument of Bradley and Roth [4]. Application studies show the same pattern in other domains: Phansalkar, More, Sabale, and Joshi adapt local thresholding to unevenly stained cytology images by operating in color spaces matched to the stain [13]. Standard image-processing texts such as Gonzalez and Woods [19] present thresholding as one segmentation tool among several, which matches the position taken in this review.

Table. 1: Local thresholding methods and their main trade-offs

Method	Local rule	Main consideration
Mean − C	Uniform neighborhood average minus offset	Fast baseline; sensitive to bright/dark contamination
Gaussian − C	Center-weighted neighborhood average minus offset	Can follow local variation; weighting and window matter
Niblack	Mean + sensitivity × local standard deviation	Responds to contrast; noisy background risk

Sauvola	Mean × normalized standard-deviation correction	Useful for degraded documents; parameter tuning
Bradley–Roth	Mean via integral image	Efficient local sums; scale/offset selection

3. Methodology:

3.1. Image and evaluation material:

The source is a review paper and supplies no original image set, labels, per-image conditions, or numerical scores. A reproducible future study should assemble images with and without illumination gradients, plus representative noise, text or object scales, and backgrounds. Foreground masks drawn and checked by independent annotators would be needed for pixel-level evaluation. Keep acquisition and annotations fixed across all compared methods; divide any tuning and final test material before selecting parameters.

3.2. Data preprocessing:

Convert color images to grayscale when the target is separable by luminance. Inspect the intensity histogram and representative dark/bright regions before applying enhancement. A small median filter may suppress impulse noise; Gaussian smoothing may reduce sensor noise but can soften narrow edges. If local contrast is weak, apply CLAHE with conservative limits [5]. Preprocessing settings must be chosen on development images only, then held constant on the test set.

Image dimensions and foreground width guide window selection. A local window should be large enough to represent both object and nearby background but small enough to follow changes in lighting. Windows such as 15 × 15, 21 × 21, and 31 × 31 appear as illustrative candidates in the source paper, not optimized values. Boundary handling must be specified because windows around image edges extend beyond available pixels.

3.3. Local threshold model:

For an intensity image I, define the binary mask $B(x,y) = 255$ if $I(x,y) > T(x,y)$, and 0 otherwise. Invert this decision when the desired foreground is darker than its background. For the mean variant, $T(x,y)$ = mean of intensities within window $W(x,y) - C$. For the Gaussian variant, replace the uniform mean with a Gaussian-weighted local average. C is an offset and must be tuned to the intensity scale and foreground polarity.

Variance-aware methods introduce local standard deviation to distinguish low-contrast regions from background [2,3]. They are related variants, not drop-in equivalents to mean/Gaussian rules. Apply each method to exactly the same preprocessed images when comparing approaches; otherwise an improvement cannot be attributed to the threshold estimator.

3.4. Parameter selection and post-processing:

Choose neighborhood size and offset using a prespecified validation subset. A small window tracks rapid lighting changes but can fragment wide structures and react to noise; a large window stabilizes the estimate but may cross unrelated regions. Sweep a modest grid of odd-valued windows and offsets, then record both selected settings and failures. Compare foreground polarity explicitly. Avoid tuning on final test masks.

After binarization, use morphological opening to remove isolated specks or closing to bridge small gaps only if they match the object scale. Report the structuring element and size. Aggressive morphology can erase thin strokes or join adjacent objects, so retain the raw mask for comparison.

3.5 Algorithm (pseudocode):

```
BEGIN
Read and verify input image
Convert to grayscale; optionally denoise or enhance
For each pixel, calculate local mean or Gaussian-weighted value
Subtract offset C and compare pixel intensity with local threshold
Build binary mask with declared foreground polarity
Optionally apply bounded morphological cleanup
Save mask, settings, and quality measures
END
```

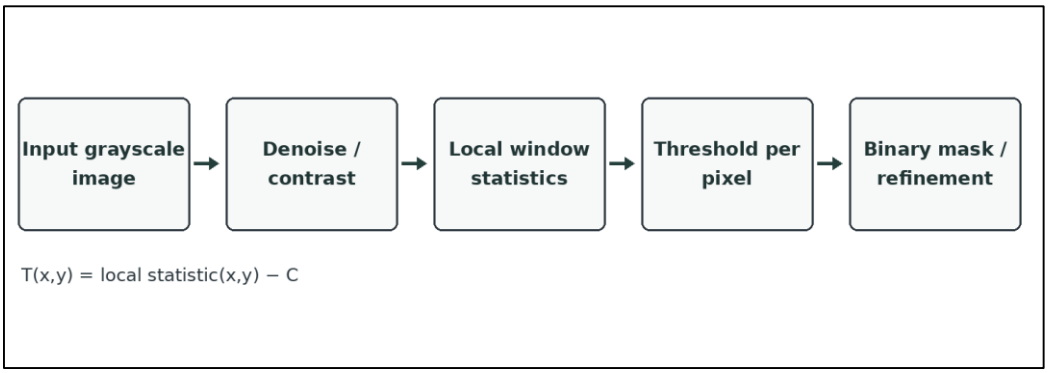


Figure. 1: Adaptive thresholding workflow (method schematic)

The diagram describes processing order only; it is not an experimental output. When reproducibility matters, save the input version, preprocessing configuration, threshold variant, window, C value, polarity, border policy, and cleanup operation with each output mask.

3.6. Evaluation protocol:

A fair comparison requires the same images, ground-truth masks, polarity, and preprocessing policy. Include a global threshold baseline [1] and at least mean and Gaussian local thresholds; include Niblack, Sauvola, or integral-image methods where the application calls for them [2–4]. Select settings using training or validation images, then measure on held-out test images. Stratify results by lighting condition and object scale rather than averaging away difficult cases.

Pixel-level counts use true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). When foreground is sparse, accuracy can remain high even with poor object recovery; foreground precision, recall, F1, intersection-over-union (IoU), and Dice are more informative. State whether empty-mask cases are excluded or assigned a defined score. Record runtime and memory under the same hardware and image dimensions if efficiency is part of the claim.

Table. 2: Metrics for a future labeled evaluation

Measure	Definition	Interpretation
Precision	TP / (TP + FP)	False foreground control
Recall	TP / (TP + FN)	Foreground recovery
F1-score	2 × precision × recall / (precision + recall)	Balance of precision and recall
IoU	TP / (TP + FP + FN)	Foreground overlap
Dice	2TP / (2TP + FP + FN)	Foreground similarity
Accuracy	(TP + TN) / (TP + TN + FP + FN)	Overall pixels; can hide class imbalance

For any metric with a zero denominator, publish a consistent convention. Dice and IoU are related but not interchangeable. Structural similarity (SSIM) measures image appearance and may be reported as a separate visual-quality measure; it does not replace mask-based overlap against ground truth.

3.7. Reporting requirements:

A complete experimental report should identify image sources and permissions, sample counts, annotation procedure, preprocessing, parameter grids, held-out splits, and uncertainty across images. Show representative successes and failures at matched display scales. The present source has none of these experimental records, so Table II defines the evaluation design rather than presenting scores.

4. Results and discussion:

No measured segmentation results were supplied in the source attachment. Accordingly, there are no defensible accuracy percentages, runtime comparisons, confusion matrices, or claims that a particular adaptive variant won. The comparison below states expected behavior from each threshold rule; it is a guide to experiments, not evidence of measured improvement on a specific dataset.

4.1. Qualitative comparison:

A fixed global cutoff is appropriate when foreground and background have stable intensity separation across the frame. Under a slow lighting gradient, that cutoff may miss dark-side foreground or turn bright-side background into foreground. A local cutoff can follow the gradient, but its window may also follow texture or noise. Mean and Gaussian variants differ in how they weight pixels within the window; their relative quality depends on object width, blur, offset, and illumination pattern.

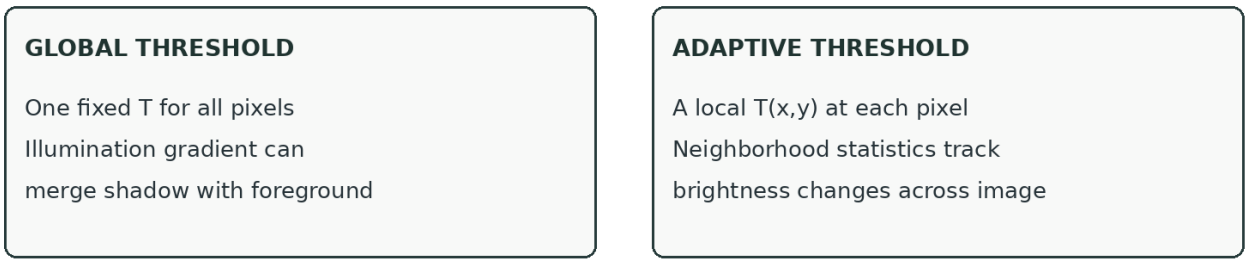


Figure. 2: Conceptual distinction between global and local decisions

Figure 2 is a schematic of threshold selection, not a side-by-side segmentation result. In degraded documents, Sauvola-style variance correction may help when local background statistics vary [3]. An integral-image implementation can reduce the repeated cost of local means [4], but speed depends on the full pipeline and hardware. Claims about effectiveness or real-time suitability require measurements on the intended application.

4.2. Error analysis:

Inspect false foreground in textured backgrounds and missed foreground in low-contrast or very narrow structures. Compare raw and morphologically refined masks to reveal whether cleanup concealed threshold errors. In images where color or texture carries the target distinction but grayscale intensity overlaps, local binarization alone cannot solve the underlying ambiguity.

4.3. Application-specific considerations:

For document binarization, uneven page lighting, faded text, and thin characters motivate local statistics and careful handling of stroke width [3]. For biomedical and industrial images, local contrast may vary, but a binary intensity split should not be treated as a semantic diagnosis or defect label. Validate the segmented boundaries with domain-specific annotations before using masks downstream. In all cases, an attractive-looking mask is not a substitute for quantified error.

Large images invite efficient implementation, including integral images for rectangular local means [4]. Memory use, edge policy, filter cost, and optional morphology can dominate the complete pipeline. A practical implementation should log per-image failure cases and latency as well as overlap metrics; comparing only the threshold operation would understate system cost.

4.4. Related situations and benchmark context:

The situations covered by the additional references share a common trigger: spatial variation that a single cutoff cannot follow. Camera-acquired documents under ordinary desk lighting motivated early local averages [9]; degraded and historical manuscripts motivated background-surface and contrast-based estimators [11,14]; stained cytology slides motivated color-space-aware local thresholds [13]; and large document collections motivated integral-image implementations that keep local statistics affordable [4,12]. In each situation the authors report the same practical lesson: the window, polarity, and sensitivity constants must match the image scale, and the method should be scored on a shared test set such as the DIBCO benchmarks [17] under an explicit evaluation protocol [16].

These related studies also clarify what this review does not claim. Published benchmark scores apply to the specific datasets and parameter settings in those papers; they do not transfer automatically to new cameras, stains, or surfaces. The honest way to use the cited results here is as guidance for what to try and how to measure it, not as evidence that any one adaptive rule will win on an unseen image set.

4.5. Limitations and future work:

Adaptive thresholding does not automatically choose a correct foreground class, a good window, or a suitable offset. Highly textured regions, severe noise, tiny objects, and foreground/background intensity overlap remain difficult. A future study could evaluate multi-scale windows, variance-aware estimators, and cautious enhancement on annotated images under controlled illumination shifts. Hybrid methods may be worth testing where intensity-only masks fail, but they should be compared under a shared test protocol rather than assumed superior.

5. Conclusion:

Adaptive thresholding replaces one image-wide cutoff with a threshold conditioned on nearby pixels. That choice directly addresses spatial variation in brightness while keeping the decision relatively simple. The most defensible use is a well-specified pipeline with declared polarity, reproducible window and offset choices, a fixed evaluation set, and inspection of both gains and failures. This review preserves the source paper's adaptive-thresholding methods and evaluation concepts, while withholding numerical outcome claims because no original dataset or experimental measurements were supplied. The added references situate adaptive thresholding among global histogram selectors, document binarization benchmarks, and application studies, and they point to a consistent evaluation practice: fixed ground truth, declared parameters, and held-out test images [10,16,17].

6. References:

(1) N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-9, no. 1, pp. 62–66, 1979.
(2) W. Niblack, An Introduction to Digital Image Processing. Prentice Hall, 1986.
(3) J. Sauvola and M. Pietikäinen, "Adaptive Document Image Binarization," Pattern Recognition, vol. 33, no. 2, pp. 225–236, 2000.
(4) D. Bradley and G. Roth, "Adaptive Thresholding Using the Integral Image," Journal of Graphics Tools, vol. 12, no. 2, pp. 13–21, 2007.
(5) A. M. Reza, "Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement," Journal of VLSI Signal Processing Systems, vol. 38, pp. 35–44, 2004.
(6) J. Bernsen, "Dynamic thresholding of grey-level images," in Proc. 8th Int. Conf. on Pattern Recognition (ICPR), Paris, France, 1986, pp. 1251–1255.
(7) J. Kittler and J. Illingworth, "Minimum error thresholding," Pattern Recognition, vol. 19, no. 1, pp. 41–47, 1986.
(8) J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," Computer Vision, Graphics, and Image Processing, vol. 29, no. 3, pp. 273–285, 1985.
(9) P. D. Wellner, "Adaptive Thresholding for the DigitalDesk," Xerox EuroPARC Technical Report EPC-93-110, 1993.
(10) M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," Journal of Electronic Imaging, vol. 13, no. 1, pp. 146–165, 2004.
(11) B. Gatos, I. Pratikakis, and S. J. Perantonis, "Adaptive degraded document image binarization," Pattern Recognition, vol. 39, no. 3, pp. 317–327, 2006.
(12) F. Shafait, D. Keysers, and T. M. Breuel, "Efficient implementation of local adaptive thresholding techniques using integral images," in Proc. SPIE 6815, Document Recognition and Retrieval XV, 2008.
(13) N. Phansalkar, S. More, A. Sabale, and M. Joshi, "Adaptive local thresholding for detection of nuclei in diversity stained cytology images," in Proc. Int. Conf. on Communications and Signal Processing (ICCSP), 2011.
(14) B. Su, S. Lu, and C. L. Tan, "Robust document image binarization technique for degraded document images," IEEE Transactions on Image Processing, vol. 22, no. 4, pp. 1408–1417, 2013.
(15) N. R. Howe, "Document binarization with automatic parameter tuning," International Journal on Document Analysis and Recognition, vol. 16, no. 3, pp. 247–258, 2013.
(16) K. Ntirogiannis, B. Gatos, and I. Pratikakis, "Performance evaluation methodology for historical document image binarization," IEEE Transactions on Image Processing, vol. 22, no. 2, pp. 595–609, 2013.
(17) I. Pratikakis, B. Gatos, and K. Ntirogiannis, "ICDAR 2009 document image binarization contest (DIBCO 2009)," in Proc. 10th Int. Conf. on Document Analysis and Recognition, 2009, pp. 1375–1382.

(18) B. Bataineh, S. N. H. S. Abdullah, and K. Omar, “An adaptive local binarization method for document images based on a novel thresholding method and dynamic windows,” Pattern Recognition Letters, vol. 32, no. 14, pp. 1805–1813, 2011.
(19) R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3rd ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2008.
(20) N. R. Pal and S. K. Pal, “A review on image segmentation techniques,” Pattern Recognition, vol. 26, no. 9, pp. 1277–1294, 1993.