The Original

GENHAND: A HYBRID EVOLUTIONARY ADVERSARIAL FRAMEWORK FOR STRUCTURE-PRESERVING CHARACTER LEVEL GENERATION

Manikanta Srinivasula[1], P. Vidyullatha[2*], Sunita M[3], P.Anil kumar[4], G Ramana Murthy[5], Swetha Kodali[6]

[1, 2*]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Green Fields, Guntur, AP, India.Manikantasrinivasula960@gmail.com[1], latha22pellakuri@gmail.com[2]
[3]Associate professor, Department of Computer Science and Engineering (Data Science), Marri Laxman Reddy Institute of technology and management sunita.m@mlritm.ac.in
[4]Associate Professor, ECE Department, Kommuri Pratap Reddy Institute of Technology, Telangana anilpole403@gmail.com
[5]Alliance School of Applied Engineering, Alliance University, Bangalore, India gajularm@gmail.com
[6]Assistant Professor, Department of Computer Science and Engineering (AIML), Sridevi Women's Engineering College, Hyderabad skswethakodali7@gmail.com

ABSTRACT
The Focus of the proposed research is on the generation of Handwritten Text, which has been a difficult problem because of style variations, overlapping strokes and the problem of preserving structural integrity at the character level. This research work focusses on a hybrid evolutionary-adversarial framework that can be used to solve such issues by integrating object detection, evolutionary optimisation and generative modelling. To train a YOLOv8 detector was created a custom dataset of 28,830 annotated English alphanumeric characters, and Particle Swarm Optimisation (PSO) and Covariance Matrix Adaptation Evolution Strategy (CMA-ES) are applied to perform the refinement of the bounding boxes and correct their orientations. To produce high-quality handwriting, a two-GAN architecture, ContourGAN to generate edge fidelity and FullMaskGAN for structural completeness that is jointly stabilized with genetic algorithm is proposed in this work. A Lightweight Transformer Module is developed to processes the output after they have been refined to enhance the perceptual quality. The resulting pipeline allows end-to-end conversion of typed text to a variety of, structure-preserving handwriting styles, with several applications in data augmentation, OCR enhancement, and digital archiving.

KEYWORD Covariance Matrix Adaptation Evolution Strategy, Evolutionary Generative Adversarial Networks, Peak Signal-to-Noise Ratio, Particle Swarm Optimization with Region Proposal Network, You Only Look Once object detection model.

1. INTRODUCTION

Handwriting remains a critical medium of human communication and cultural heritage, evident in archived manuscripts, medical prescriptions, administrative forms, and personal notes. Despite rapid digitization, vast volumes of handwritten material must still be segmented, recognized, and restored for downstream applications such as optical character recognition (OCR), digital archiving, and handwriting imitation [1,23]. A key prerequisite for these tasks is character-level segmentation process made especially challenging by diverse writing styles, overlapping strokes, skew, noise, and irregular spacing [2]. Character segmentation also plays a vital role in synthetic handwriting generation, where high-quality, mask-based rendering can augment training datasets for recognition systems, particularly for underrepresented styles or rare characters [3]. Accurate segmentation further enables applications such as handwriting imitation [4], font synthesis [5], document restoration [6], and digital preservation [7]. However, traditional heuristic methods and basic object detectors often fail under the variability inherent in freehand writing [8,29]. While deep learning approaches have advanced the field, they still degrade in performance on complex layouts, degraded documents, or inconsistently written text [9]. Similarly, generative adversarial networks (GANs) used for handwriting synthesis frequently suffer from mode collapse or require large amounts of labeled data, limiting their scalability and generalization [10,24]. These shortcomings underscore the need for an adaptive, minimally supervised pipeline capable of extracting, refining, and reconstructing handwritten characters while preserving structural integrity. Handwriting poses unique challenges compared to printed text: strokes frequently overlap, edges are ambiguous, boundaries are fuzzy, and annotated data at the character level is scarce across many languages and scripts. Real-world degradations such as noise, paper texture, and lighting variations-further complicate segmentation. To address these issues, robust object detection must be combined with edge-aware cropping and evolutionary post-processing. Methods such as Particle Swarm Optimization (PSO), Region Proposal Networks (RPN), and Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [11,25,26,27,28,29], when integrated with GAN-based mask generation, offer a principled path toward precise and scalable character segmentation. Existing techniques remain limited: rule-based heuristics rely on rigid geometric assumptions and fail under natural variability, contour- and projection-based methods often under-crop or merge adjacent characters and most GAN-based solutions prioritize style transfer or super-resolution over accurate, mask-level segmentation[30,31]. Moreover, many generative models are trained end-to-end with minimal structural supervision, making them prone to mode collapse and poor performance on edge cases. To bridge these gaps, we propose an evolution-guided pipeline for handwritten character segmentation and mask generation. Our framework integrates YOLOv8 with PSO to refine bounding box alignment and edge precision, and employs CMA-ES to correct rotation inconsistencies. For synthesis, we introduce a dual-GAN architecture: *ContourGAN*, which captures edge fidelity, and *FullMaskGAN*, which ensures structural completeness. Genetic operators are incorporated to stabilize training and enrich stylistic diversity. Unlike conventional pipelines, our modular design is robust to noisy, low-label inputs and highly scalable for data augmentation and document restoration. The proposed system has broad applicability. By enabling precise character isolation and structure-preserving generation, it enhances OCR accuracy, supports digital document repair facilitates large-scale synthetic dataset creation [12], and aids in low-resource language preservation. Beyond recognition, it can be applied to educational tools signature modeling, and font design. More broadly, our work advances the vision of end-to-end handwritten text understanding by bridging the gap between segmentation accuracy and generative realism-without reliance on heavily curated data or rigid rule-based systems.

2. RELATED WORK

The recent developments in text detection have employed the use of YOLO-based architectures to obtain robust performance to handle complex text detection situations. Sun et al. (2024) [13] proposed RA-YOLOv8 by modifying RFEMA as the fore feature extraction and curved text detection to have high accuracy in the presence of clutter in the background. **Kundu** and **Bhattacharya** (2024) [14] implemented the A* pathfinding algorithm to the YOLO-based text line segmentation method to effectively deal with skew and noise in degraded historical manuscripts. Such techniques are effective at line- or region-level detection but are not effective at unconstrained handwriting where character-level segmentation is required. This work develops an extension of YOLOv8 that includes edge-informed cropping and evolutionary refinements that leads to highly precise isolation of characters in highly variable handwriting. Evolutionary Optimization of detection and segmentation algorithms has also been used in a different context where evolutionary algorithms may be used to create more resilient bounding box localization, under a variable geometry aspect. **Bhanurangarao** et al. (2025) [15] made PSO orientation-aware by changing the fitness function to incorporate aspect ratio, orientation, and overlap which made PSO more robust to rotation and scale. **Borckmans** and others (2010) [16] directly implemented PSO on rotation group to find optimal oriented 3D bounding boxes at high precision. These methods, however, are perceptual to all object detection instead of the irregular shapes of the handwritten characters. We generalize such principles to handwriting using PSO-refined bounding boxes and CMA-ES-mapped rotation correction, adapted to the noise and variability of freehand handwriting.

Generative adversarial networks have contributed a lot in handwriting synthesis. HiGAN [17,18] make it possible to generate indefinitely long text with style-content disentanglement and allow the one-shot writer imitation. GANwriting adapts this to the use of few-shot style conditioning, whereas ScrabbleGAN [19] combines the use of transformers with semi-supervised learning on variable length text. The stroke level: SSGAN adopts attention-enhanced ResU-Net, which can maintain calligraphic form. Aesthetically pleasing, these models can be highly style-transfer or word-level synthesis oriented where pixel-aligned said character masks are not a focus. Conversely, our pipeline experiments with two-coordinate style-transfer architecture, which consist of ContourGAN (edge fidelity) and FullMaskGAN (structural completeness) with the goal of preserving structural integrity at character level. Hybrid networks consisting of GANs and auxiliary models have been demonstrated to have potential in document analysis as well. U-Net and PatchGAN were use by Ji and Chen (2019) [20] to treat the task of extracting text lines as image-to-image translation, whereas tsegGAN adopted adversarial learning to segment touching characters with a modified separation loss. As an alternative, Qian et al. (2020) [21] combined super-resolution and classification into a three-player GAN, and Yang et al. (2022) [22] presented the GAN-based handwriting style transfer that showed high visual realism. Although they have worked well, most of these pipelines are oriented to layout-based modeling or word-level adaptation of style. In contrast to this work, we incorporate modular, evolutionarily stabilized modules, combine PSO and CMA-ES to improve stability through more character-based segmentation and synthesis.

The current approaches to handwriting segmentation and synthesis tend to investigate the imitation of style only on global levels such as words or even global layouts leaving behind the necessity to attain character-level precision. Evolutionary optimization, however, has been used in geometric localization but not yet in handwriting pipelines. The main contributions based on research gaps are.

1. Character-level extension of YOLOv8 with evolutionary refinements for precise segmentation.
2. Initial application of PSO + CMA-ES to handwriting synthesis to allow powerful correction of bounding-boxes and rotation errors.
3. An innovation based on the structure-preserving generate technique, a novel dual-GAN architecture was stabilized by evolutionary strategies.
4. Lightweight transformer-based refinement for perceptual realism.
5. A scalable pipeline with the capability to scale data augmentation, OCR improvement, and digital archiving.

3. METHODOLOGY

This research work suggests the entire pipeline of deep learning and evolutionary-based generative modelling that optimizes the production of high-fidelity handwritten character masking. The data are gathered, labelled and fine tuned to a detector YOLOv8 to attain a noteworthy localization of the characters, and the rotations are carried out on CMA-ES that helps to vary within-class. The data is restricted to grayscale English characters so far, and augmentation mitigates writer variability to some extent (more on that in the Conclusion). The pipeline consists of a model based on the dual-gans architecture are FullMaskGan with dense structural masks and ContourGan with edge-based masks, both optimised with evolutionary optimisation and transformer refinement.

3.1. Dataset Preparation and Augmentation

To design an effective handwritten character recognition system, we created 93 high-resolution composite images with 640 x 640 pixels. The role of character regions was annotated with the help of the Roboflow platform that stored labels of bounding boxes compatible with YOLO. The composite images that contained these annotated characters were separated into and classified into 62 separate directories that reflected each of the classes. This results in a total of 28,830 individual character instances available for training and augmentation. Random affine transformations were applied to each character five times to make the diversity of datasets higher and improve the generalization. Using n_b=93 base samples per class, a augmentation factor=5, C classes =62 and we have total number of images as. This total is calculated using (1), which ensures a balanced augmentation across all classes.

$$N_{\text{total}} = C \cdot (n_b \cdot a) = 62 \cdot (93 \cdot 5) = 28{,}830 \tag{1}$$

This equation shows how data augmentation increases the total number of samples, ensuring sufficient representation for each character. Affine transformations can be expressed by a 2×3 matrix A. The structure of these transformations is described in (2), combining rotation, scaling, shear, and translation. where s is the scale, θ is rotation, ϕ is shear, and (tx, ty) are translations. This matrix represents geometric perturbations used to simulate handwriting variations.

$$A = \begin{bmatrix} s \cdot \cos\theta & -s \cdot \sin\theta + \emptyset & t_x \\ s \cdot \sin\theta + \emptyset & s \cdot \cos\theta & t_y \end{bmatrix} \tag{2}$$

3.2. YOLOv8m-Based Detection and Classification

The YOLOv8m model, character detection and classification is used to predicts bounding boxes $[x_i, y_i, w_i, h_i]$ and class probabilities $P(\, y_i \mid x\,)$. The final predicted class is given by. This prediction is made by choosing the class with the highest confidence as shown in (3) that identifies the class label with the highest confidence score.

$$\hat{y} = \arg\max_i P\,(\,y_i \mid x\,) \tag{3}$$

To evaluate bounding box accuracy, the Intersection-over-Union (IoU) metric. The IoU calculation in (4) helps quantify how well predicted boxes match ground truth annotations. where A and B are the ground-truth and predicted bounding boxes. IoU is critical for computing performance metrics such as mAP.

$$\text{IoU}(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{4}$$

3.3. Character Extraction, PSO Refinement, and CMA-ES-Based Rotation

Initial segmentation is achieved through Canny edge detection, where the gradient magnitude is computed as. This edge enhancement is based on the gradient formulation in (5).

$$G(x, y) = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \tag{5}$$

with I(x,y) being the grayscale input image. This step enhances edge detection for character boundaries. Particle Swarm Optimization (PSO) refines the bounding boxes. For each box B_i, the fitness function is defined as. The multi-objective fitness is expressed in (6), considering both edge alignment and shape compactness.

$$\mathcal{F}(B_i) = \lambda_1 \cdot \mathcal{E}_{\text{edge}}(B_i) + \lambda_2 \cdot C_{Compact}(B_i) \tag{6}$$

where $\mathcal{E}_{\text{edge}}$ evaluates alignment with detected edges and $C_{Compact}$penalizes irregular shapes. For binarization, adaptive thresholding is used. This dynamic binarization follows the formulation given in (7).

$$T_{adaptive}(x) = \begin{cases} 1, & x > \mu(x_{nxn}) - \delta \\ 0, & otherwise \end{cases} \tag{7}$$

where $\mu(x_{nxn})$ is the local mean intensity in an n×n neighborhood. To correct rotation, Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is applied. The fitness function for rotation is As shown in (8), it measures how close the rotated output is to the target angular set. Where $t = [-8°, -4°, 4°, 8°]$ defines the desired set of rotation angles.

$$\text{Fitness}(\theta) = \frac{1}{4}\sum_{i=1}^{4}|\theta_i - t_i| \tag{8}$$

3.4. ContourGAN: Edge-Based Mask Synthesis

ContourGAN is built using a U-Net generator G and a PatchGAN discriminator D to synthesize binary edge masks $\hat{y} = G(x) \in [0,1]^{64\times64}$. The adversarial loss is captured by the binary cross-entropy loss shown in (9).

$$L_{adv} = \text{BCE}(D(x,\hat{y}),1) \tag{9}$$

and the reconstruction loss is This L1-based term encourages fidelity to the ground truth as in (10). The total generator loss is weighted combination of losses is defined in (11). With λL1=50. This ensures both realism and accuracy of the generated masks.

$$\mathcal{L}_{\text{L}} = |\hat{y} - y|_1 \tag{10}$$

$$\mathcal{L}_{\mathcal{G}} = \lambda_{\text{adv}} \cdot \mathcal{L}_{adv} + \lambda_{\text{L1}} \cdot \mathcal{L}_{\text{L}} \tag{11}$$

The Fréchet Inception Distance (FID) is used to assess similarity between real and generated samples. he computation of FID is presented in (12). ContourGAN is a U-Net and a PatchGAN to learn the generator and discriminator in such a way that it would generate a binary edge mask $\hat{y} = G(x)$. Generator training makes use of a weighted sum of the adversarial loss function $\mathcal{L}_{adv}$ reconstruction loss (the combination is known as an adversarial L1 loss) as demonstrated in Equation (9)-(11), where FID (Equation 12) is used to score the quality of generations.

$$\text{FID} = |\mu_r - \mu_g|^2 + \text{Tr}\left(\Sigma_r + \Sigma_g - 2\left(\Sigma_r\Sigma_g\right)^{1/2}\right) \tag{12}$$

The entire architectural specifics of the Generator and the Discriminator employed in ContourGAN are provided in Table 1. The Generator encoder-decoder U-Net architecture with skip linkages and comprising layered Conv2D, BatchNorm, ReLU, as well as ConvTranspose2D functions has more than 20 million arguments; it can be trained with a loss of 0.31. The Discriminator uses a patchGAN-like architecture, which has down-sampling Conv2D layers with progressively symmetrical convolutions and completion Sigmoid activation, which basically tells the difference between real and generated edge masks.

Table 1: Architectural details of ContourGAN's **U-Net-based Generator** and **PatchGAN Discriminator**

Layer (Type)	Generator Output Shape	Generator Params	Discriminator Output Shape	Discriminator Params
Conv2d	[-1, 64, 32, 32]	1,024	[-1, 64, 32, 32]	2,048
LeakyReLU	[-1, 64, 32, 32]	0	[-1, 64, 32, 32]	0
Conv2d	[-1, 128, 16, 16]	131,072	[-1, 128, 16, 16]	131,072
BatchNorm2d	[-1, 128, 16, 16]	256	[-1, 128, 16, 16]	256
LeakyReLU	[-1, 128, 16, 16]	0	[-1, 128, 16, 16]	0
Conv2d	[-1, 256, 8, 8]	524,288	[-1, 256, 8, 8]	524,288
BatchNorm2d	[-1, 256, 8, 8]	512	[-1, 256, 8, 8]	512
LeakyReLU	[-1, 256, 8, 8]	0	[-1, 256, 8, 8]	0
Conv2d	[-1, 512, 4, 4]	2,097,152	[-1, 512, 7, 7]	2,097,152
BatchNorm2d	[-1, 512, 4, 4]	1,024	[-1, 512, 7, 7]	1,024
LeakyReLU	[-1, 512, 4, 4]	0	[-1, 512, 7, 7]	0
AdaptiveAvgPool2d	-	-	[-1, 512, 1, 1]	0
Linear	-	-	[-1, 1]	513
Sigmoid	-	-	[-1, 1]	0

3.5. FullMaskGAN: Dense Mask Generation

This GAN variant uses a shallow U-Net and evolves parameters via mutation-only genetic operations. A parameter θk is updated as: This mutation-based parameter update is shown in (13).

$$\theta'_k(x) = \begin{cases} \theta_k + \varepsilon, & \epsilon \sim N(0,\sigma^2) \\ \theta_k, & otherwise \end{cases} \tag{13}$$

where σ=0.022 and p=0.01.

Final masks are thresholded using Otsu's method. This optimal thresholding criterion is defined in (14).

$$\theta^* = \arg\min_{\theta}[\omega_1(\theta)\sigma_1^2(\theta) + \omega_2(\theta)\sigma_2^2(\theta)] \tag{14}$$

where ω1, ω2and σ1, σ2 represent class probabilities and variances. The protein also diminished representations that underlie genetic algorithms to educate a shallow U-Net and give rise to FullMaskGAN. Parameters are being updated in accordance with the use of Gaussian perturbation (Equation 13) with the Otsu method (Equation 14) used to binarize the final mask to obtain the optimal thresholding. To post-process and subtly improve perceived realism of the created masks, a light optimisation based on the Transformer is applied. Mi has a slow attending-ish update. This will represent what will show the formula with the equation number (15), aligned to the right.

$$M_i^{\text{final}} = M_i + \alpha \cdot \text{Refine}(M_i) \tag{15}$$

Where Refine (·) represents a simple contextual enhancement and α≪1 controls the intensity of refinement. This step preserves the original stroke structure while slightly improving edge consistency, contrast, and overall realism of the masks.

3.6. Text Layout and Rendering Engine

The rendering engine composes text by mapping each character ci to a binary mask Mi. When a character is unavailable, fallback substitutions (e.g., "0" → "O") are applied. For Contour rendering, a Gaussian blur and Canny edge detector are used. The combined operations are defined in (16).

$$G(x) = x * \mathcal{N}(0, \sigma^2), \quad E(x) = \text{Canny}\big(G(x)\big) \tag{16}$$

followed by morphological dilation. The dilation kernel enlarges the edges as shown in (17).

$$D(x) = \text{dilate}(E(x), \mathcal{K}) \tag{17}$$

Text is placed on a virtual A4 canvas (794×1123 pixels) with 50 px margins, 10 px line spacing, and 2 px character spacing. The output is presented through an interactive GUI.

Algorithm 1. End-to-end pipeline for character recognition, detection, mask refinement, and document rendering.

Algorithm 1: End-to-End Character Recognition Pipeline
Input: Composite image I with 62 characters
Output: Final rendered document with refined masks
1: Annotate dataset → bounding boxes BBB
2: Split dataset by class; apply 5× affine transformations
3: Train YOLOv8m on annotated dataset
4: Predict bounding boxes (x,y,w,h) and class probabilities P(y|x)
5: Assign $\hat{y} = \arg\max_i P(y_i \mid x)$; apply IoU filtering
6: Apply Canny edge detection on detected regions
7: Refine bounding boxes with PSO fitness: $\mathcal{F}(B) = \lambda_1 E_{\text{edge}} + \lambda_2 C_{\text{compact}}$
8: Binarize characters using adaptive thresholding
9: Correct rotations with CMA-ES optimization
10: Generate masks with ContourGAN (edge-based)
11: Refine masks with FullMaskGAN + Otsu thresholding
12: Apply Transformer refinement: $M^* = M + \alpha \cdot \text{Refine}(M)$
13: Preprocess masks (Gaussian blur + Canny + dilation)
14: Arrange characters on A4 canvas (794×1123 px)
15: Render final document via GUI

4. RESULTS AND DISCUSSIONS

The second part is an account of the experimental results of our proposed pipeline, including the elaboration on the performance of each component i.e., detection, refinement to mask generation using both quantitative and visual outputs. These annotations were created using the Roboflow interface to ensure diversity and sufficient coverage for both detection and segmentation tasks. While a sample image is not shown here, the annotation process emphasized variability and accuracy across the dataset. Training was conducted using character classes at the character level. Each distinctly identified character was encoded and annotated with color-coded bounding boxes in a YOLOv8-compatible format to support effective detection within the pipeline. Although cropped experiments were carried out with exploratory trials on subset versions, which still made it possible to artificially maximize IoU (>0.99), those were not representative, as the lower diversity and closely truncated bounding boxes decreased them. The full dataset IoU@0.5 mean = 0.839 scales well with the reported mAP of the detections (mAP@0.5 = 0.83, mAP@[0.5:.95] = 0.78), and indicates more realistic estimation of quality of localization. From Table 2 & Figure 2. An Epoch-wise description of the training and performance indicators of the YOLOv8 model. Throughout the 50 epochs, the model improved steadily in mAP@50, precision and recall, to 0.86922, 0.84253, and 0.81657 respectively, but the losses reduced steadily and average IoU reaches the highest part of 0.9952, meaning good localization ability. In Figure 3 we have Plots of performance evaluation curves of the YOLOv8 model, F1-Confidence, Precision-Confidence, Precision-Recall and Recall-Confidence outcomes. The model also had a max F1-score of 0.83 with a confidence of 0.34, and 1 precision with high confidence, and an averaged precision of 0.86

which indicates that the model displays a high and flexible detection performance across the different confidences. In Figure 4 Denser image comparison of characters between traditional and RPN+PSO bounding box. Edge density on the RPN+PSO reflects better character coverage and structure preservation, which are requisites of reproducing high-fidelity masks of handwrites. In Figure 5 Comparison bar charts on the generative quality measures: Inception Score, (IS), FID and KID of the Contour and Full-Mask models. Compared to Full-Mask, Contour gains better FID and KID values, i.e. these values are better-balanced in their distribution and the entire generative consistency.

Table 2: Epoch-wise Training Details of the YOLOv8 Model

epoch	time	train/ box_loss	train/ cls_loss	train/ dfl_loss	metrics/ precision (B)	metrics/ recall(B)	metrics/ mAP50(B)	metrics/mAP50-95(B)
1	172.118	0.49842	3.5848	1.04009	0.34222	0.44289	0.30367	0.30348
10	1563.31	0.21715	1.87053	0.90048	0.72223	0.72897	0.74023	0.7402
20	3097.61	0.17074	1.72187	0.89562	0.75266	0.78505	0.79595	0.79414
30	4631.71	0.14487	1.60999	0.89268	0.79245	0.79457	0.82004	0.81865
40	6163.65	0.12302	1.52181	0.88878	0.82183	0.78828	0.83363	0.83251
50	7730.72	0.05375	0.44522	0.8974	0.84253	0.81657	0.86922	0.86911

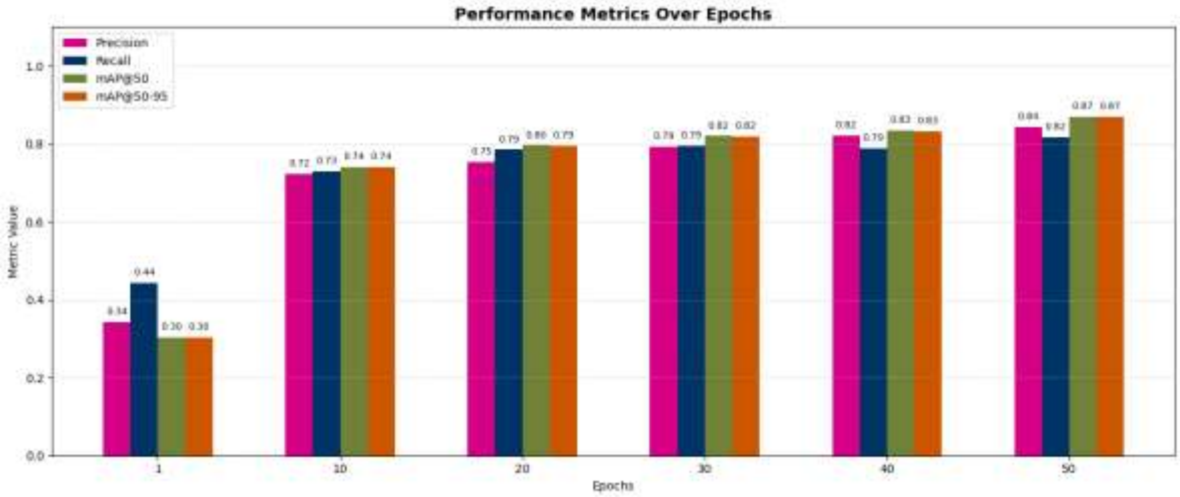

Figure 2. Epoch-wise training progression and performance metrics of the YOLOv8 model

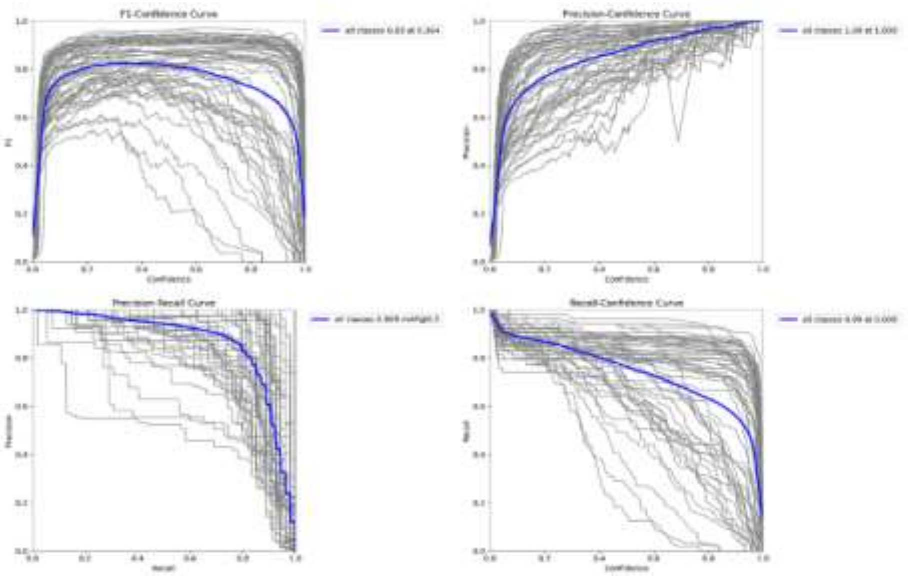

Figure 3: Performance curves of YOLOv8 showing robust detection accuracy.

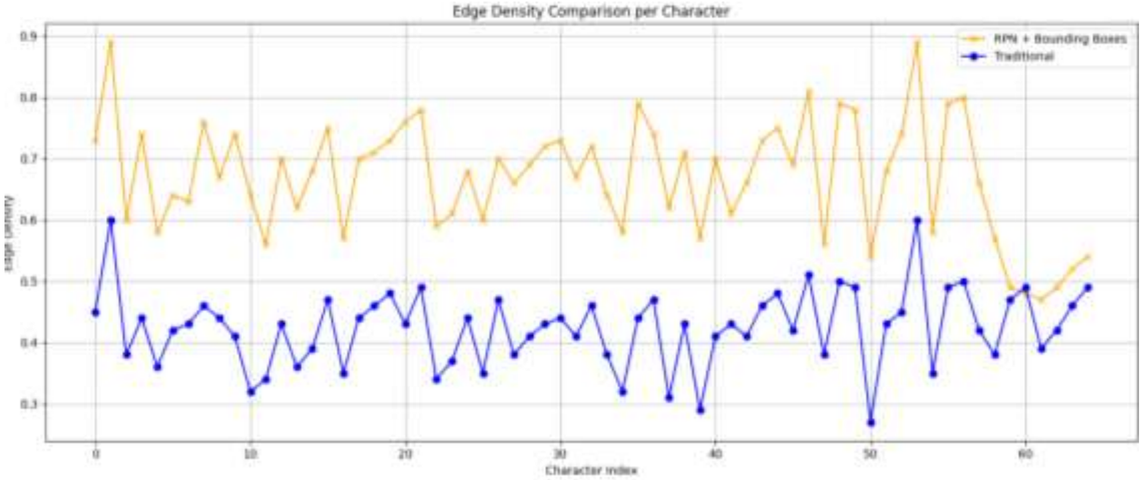

Figure 4: Comparison of RPN+PSO

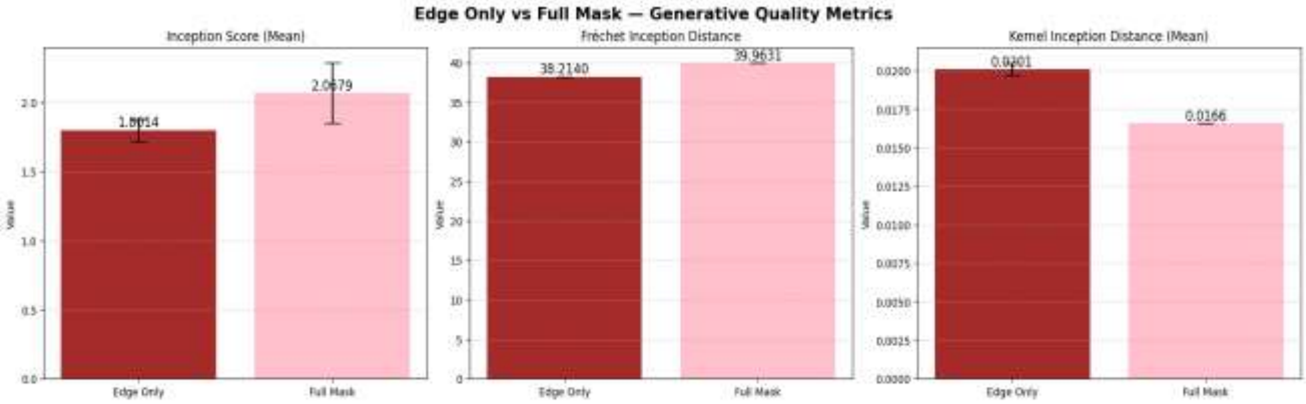

Figure 5: Comparative bar charts of IS, FID, and KID between Contour and FullMask models

CONCLUSION

This research work presented the concept of a single evolutionary-adversarial pipeline of handwriting synthesis, inclusive of YOLOv8-based detection, PSO and CMA-ES motivated refinements, and a dual-GAN formulation (ContourGAN to integrity of edges and FullMaskGAN to integrity of structure), which uses a lightweight transformer further refined. Combining these elements enables character-level accuracy and perceptual fidelity and allows processing type information to generate structure-preserving handwritten text in a paragraph scale. Int this work, an optimization gain, a stronger level of segmentation fidelity can be achieved through evolutionary optimization to consistently improve the fidelity of the generative model. But there are two main limitations with this study. First, it is based on a small grayscale English data set, which constrains the variety of styles and scripts. Second, the originality in the novel is especially the combination of evolutionary and adversarial methods, as opposed to the establishment of completely new theoretical frameworks, which might restrict its perceived contribution to the basic research in computer vision. The future work will address these limitations by extending the framework both to multi-writer, multi-script data through a more fully-fledged modelling of stylistics and to more theoretically directed work on the general theory of evolutionary search and generative models.

REFERENCES

1. Banavatu T, Parthasarathy G. Heuristic aided fusion serial cascaded deep network for handwritten character recognition from handwritten images using optimization strategy. Appl Soft Comput. 2025.
2. Aparna C, Rajchandar K. A robust solution for recognizing accurate handwritten text extraction using quantum convolutional neural network and transformer models. Comput Electr Eng. 2024;120(Part C):109794.
3. Huang Y, Zhang Z, Jiao A, Ma Y, Cheng R. A comparative visual analytics framework for evaluating evolutionary processes in multi-objective optimization. IEEE Trans Vis Comput Graph. 2024;30(1):661-71.
4. Nomura M, Shibata M. cmaes: a simple yet practical Python library for CMA-ES. arXiv preprint arXiv:2402.01373. 2024.
5. Pandey S. Neuroevolution of a multi-generator GAN (Student Abstract). Proc AAAI Conf Artif Intell. 2024;38(21):23608-9.
6. Kang L, Riba P, Wang Y, Rusiñol M, Fornés A, Villegas M. GANwriting: content-conditioned generation of styled handwritten word images. In: Lecture Notes in Computer Science, vol 12368. Springer; 2020. p. 273-89.
7. Mattick A, Mayr M, Seuret M, Maier A, Christlein V. SmartPatch: improving handwritten word imitation with patch discriminators. In: Lecture Notes in Computer Science, vol 12821. Springer; 2021. p. 268-83.
8. Tsai TH, Huang SA. Refined U-Net: a new semantic technique on hand segmentation. Neurocomputing. 2022;495:1-10.

9. Wang TQ, Jiang X, Liu CL. Query pixel guided stroke extraction with model based matching for offline handwritten Chinese characters. Pattern Recognit. 2022;123:108416.
10. Yu MM, Zhang H, Yin F, Liu CL. An approach for handwritten Chinese text recognition unifying character segmentation and recognition. Pattern Recognit. 2024;151:110373.
11. Zdenek J, Nakayama H. Handwritten text generation with character-specific encoding for style imitation. In: Lecture Notes in Computer Science, vol 14188. Springer; 2023. p. 313-29.
12. Du K, Wang X, Yan Y, Lu Y, Wang H. EGNET: a novel edge guided network for instance segmentation. Proc IEEE Int Conf Image Process (ICIP). 2022:3868-72.
13. Sun H, Tan C, Pang S, Wang H, Huang B. RA YOLOv8: an improved YOLOv8 seal text detection method. Electronics. 2024;13(15):3001.
14. Kundu A, Bhattacharya U. YOLO assisted A* algorithm for robust line segmentation of degraded document images. In: Lecture Notes in Computer Science, vol 14807. Springer; 2024. p. 407-24.
15. Bhanurangarao M, Ramesh M. Adaptive PSO with orientation awareness for robust object localization and bounding box refinement. Tiutic Org [Internet]. 2025 [cited 2025 Jul 12]. Available from: https://www.tiutic.org/pdf/volume/Vol_8/Vol8_Article_5.pdf
16. Borckmans P. Oriented bounding box computation using particle swarm optimization. In: Proceedings of the European Symposium on Artificial Neural Networks (ESANN); 2010
17. Gan J, Wang W. HiGAN: handwriting imitation conditioned on arbitrary-length texts and disentangled styles. Proc AAAI Conf Artif Intell. 2021;35(9):7484-92.
18. Gan J, Wang W, Leng J, Gao X. HiGAN+: handwriting imitation GAN with disentangled representations. ACM Trans Graph. 2023;42(1).
19. Fogel S, Averbuch-Elor H, Cohen S, Mazor S, Litman R. ScrabbleGAN: semi-supervised varying length handwritten text generation. Proc IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR). 2020:4323-32.
20. Ji B, Chen T. Generative adversarial network for handwritten text. arXiv preprint arXiv:1907.11845. 2019.
21. Qian Z, Huang K, Wang QF, Xiao J, Zhang R. Generative adversarial classifier for handwriting characters super resolution. Pattern Recognit. 2020;107:107453.
22. Yang K, Liang X, Liu Q, Wen K. Handwriting imitation with generative adversarial networks. In: Lecture Notes in Electrical Engineering, vol 808. Springer; 2022. p. 189-95.
23. Yogesh, Y., Ghantasala, G. P., & Priya, A. (2023, March). Artificial intelligence based handwriting digit recognition (hdr)-a technical review. In *2023 International Conference on Device Intelligence, Computing and Communication Technologies,(DICCT)* (pp. 275-278). IEEE.
24. Kumar, N. A., Srinadh, T., Ioannou, I., Ghantasala, G. P., Vidyullatha, P., & Vassiliou, V. (2025). Revolutionizing diabetic maculopathy detection with MobileNet, GAN-enhanced imaging, and Graph Neural Networks: A multimodal AI approach for precision ophthalmology. *Informatics in Medicine Unlocked*, *58*, 101687.
25. Pulicherla, P., Ghantasala, G. P., Reddy, A., Kumar, A. R., & Prasuna, A. L. (2026). EARLY PREDICTION OF CERVICAL CANCER RISK USING EXPLAINABLE DEEP NEURAL NETWORKS AND FEATURE ENGINEERING OPTIMIZATION. *Genetics and Molecular Research*.
26. Ghantasala, G. P. (2026, January). Quantum Approximate Optimization Algorithm for Reliable Diagnosis of Benign vs. Malignant Breast Tumors. In *2026 6th International Conference on Image Processing and Capsule Networks (ICIPCN)* (pp. 1611-1615). IEEE.
27. Guruguntla, V., Yuvaraju, B. A. G., Rao, T. S. B., Ghantasala, G. P., Vidyullatha, P., & Pydi, H. P. (2026). Development and optimization of a female-specific Biomechanical model for biodynamic response analysis: a comparison with male biomechanical models. *Scientific Reports*.
28. Ghantasala, G. P., Razvi, S. H., Yashwanth, T., Karthik, M., & Manikanta, P. S. (2025, December). Optimizing Multi-Objective Cardiovascular Disease Prediction Models Using Ant Colony Optimization and Swarm Intelligence Metaheuristics for Enhanced Clinical Decision Support. In *2025 1st International Conference on Advancement in Futuristic Technologies (ICAFT)* (pp. 1-5). IEEE.
29. Ghantasala, G. P., Reddy, A., Sharma, R. R., Vidyullatha, P., Sungheetha, A., & Ananthanagu, U. (2025, October). Metaheuristic Optimization for Feature Reduction in Cerebrovascular Disease Detection: A Butterfly Algorithm Approach. In *2025 1st IEEE Uttar Pradesh Section Women in Engineering International Conference on Electrical Electronics and Computer Engineering (UPWIECON)* (pp. 66-70). IEEE.
30. kanth Padmanabhuni, R., Malathy, V., Kumar, N. A., Ghantasala, G. P., Sridhar, T., & Singh, D. (2025, September). Walrus Optimization Algorithm for Maculopathy Detection: A Hybrid Metaheuristic Approach to Improve Classification Accuracy. In *2025 International Conference on Computing and Communications (COMPUTINGCON)* (pp. 1-8). IEEE.
31. Dilip, K., Ghantasala, G. P., Rathee, M., Kallam, S., & Bathla, P. (2023, October). A Brief Comparative Study of Metaheuristic Approaches for Hyperparameter Optimization of Machine Learning Model. In *2023 International Conference on Computer Science and Emerging Technologies (CSET)* (pp. 1-5). IEEE.